Journal of Convex Analysis
Volume 18 (2011), No. 1, 203–208

A Remark on the Structure of the Busemann Representative of a Polyconvex Function*

J. J. Bevan
Department of Mathematics, University of Surrey,
Guildford, GU2 7XH, UK

Received: June 11, 2008
Revised manuscript received: September 8, 2009

Under mild conditions on a polyconvex function $W : \mathbf{R}^{2\times 2} \to \mathbf{R}$, its largest convex representative, known as the Busemann representative, may be written as the supremum over all affine functions $\phi : \mathbf{R}^5 \to \mathbf{R}$ satisfying $\phi(\xi, \det \xi) \leq W(\xi)$ for all 2×2 matrices ξ. In this paper, we construct an example of a polyconvex $W : \mathbf{R}^{2\times 2} \to \mathbf{R}$ whose Busemann representative is, on an open set, strictly larger than the supremum of all affine functions ϕ as above and which also satisfy $\phi(\xi_0, \det \xi_0) = W(\xi_0)$ for at least one 2×2 matrix ξ_0.

2000 Mathematics Subject Classification: 26B25, 52A40, 47J30

1. Introduction

Polyconvexity was first identified by Morrey in [6] and was later developed by Ball [1] in connection with nonlinear elasticity. A function $W : \mathbf{R}^{N\times n} \to \mathbf{R} \cup \{\infty\}$ is polyconvex if there exists a convex function φ, said to be a convex representative of W, such that

$$W(\xi) = \varphi(R(\xi))$$

for all real $N \times n$ matrices ξ, where $R(\xi)$ is the list of minors of ξ written in some fixed order. Busemann *et al.* pointed out in [4] that there is a largest such convex representative, which, henceforth, we refer to as the Busemann representative and denote by φ_W.

One of the broader aims of the series of papers [4] Busemann *et al.* was to study the restriction of convex functions to non-convex sets. Ball observed in [1] that polyconvexity fits into this framework, and the relationship between the two has since been explored further in [3].

Busemann *et al.* show in [4] that provided the polyconvex function W is bounded linearly below and real-valued then the Busemann representative can be expressed as

$$\varphi_W(X) = \inf\left\{\sum_{i=1}^{d+1} \lambda_j W(\xi_j) : \quad \lambda_j \geq 0,\ \sum_{j=1}^{d+1} \lambda_j = 1 \text{ and } \sum_{j=1}^{d+1} \lambda_j R(\xi_j) = X\right\}. \qquad (1)$$

*This research was supported by an EPSRC Postdoctoral Research Fellowship GR/S29621/01, by the European Research and Training Network MULTIMAT and by an RCUK Academic Fellowship.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

Here, d is the least integer such that $R(\xi) \in \mathbf{R}^d$ for all $\xi \in \mathbf{R}^{N\times n}$ and X lies in $\mathbf{R}^d$. They also show that

$$\varphi_W(X) = \sup\{a(X) : a \in \mathcal{L}\}, \tag{2}$$

where

$$\mathcal{L} = \{\phi \text{ affine} : \ \phi(R(\xi)) \leq W(\xi) \ \forall\, \xi \in \mathbf{R}^{N\times n}\}.$$

The hypothesis that W is bounded linearly below ensures that there is at least one element in $\mathcal{L}$. The graph of any $\phi \in \mathcal{L}$ is a hyperplane, so (2) states that φ_W is built from hyperplanes which lie below the set $G_W := \{(R(\xi), W(\xi)) : \ \xi \in \mathbf{R}^{N\times n}\}$.

The main result in this short note is that there is no redundancy in the expression (2) in the case $N = n = 2$. To be precise, one cannot replace $\mathcal{L}$ in (2) by the smaller class $\mathcal{T}$, where

$$\mathcal{T} = \{\phi \in \mathcal{L} : \exists\, \xi \in \mathbf{R}^{2\times 2} \text{ s.t. } W(\xi) = \phi(R(\xi))\}.$$

Thus $\mathcal{T}$ represents the collection of supporting hyperplanes to G_W which meet G_W in at least one point. We define

$$\varphi_{\mathcal{T}}(X) = \sup\{a(X) : \ a \in \mathcal{T}\}.$$

Note that $\varphi_W \geq \varphi_{\mathcal{T}}$ in view of the inclusion $\mathcal{T} \subset \mathcal{L}$. In the next section we construct a real-valued, non-negative polyconvex W to which (2) applies and which is such that $\varphi_W > \varphi_{\mathcal{T}}$ on a large set. This result is surprising since the set $\{R(\xi) : \ \xi \in \mathbf{R}^{2\times 2}\}$ is large: its convex hull is the whole of $\mathbf{R}^5$. (For a proof of this fact see [1].) Certainly one might expect $\varphi_W = \varphi_{\mathcal{T}}$ to be the case under extra assumptions, which could include super-quadratic growth of W, for example. See [3] for further details.

The result of this note is relevant to [3, Lemma 2.4], where the structure of φ_W is described. We present a version of the lemma below for the reader's benefit; for the proof consult [3].

Lemma 1.1 ([3, Lemma 2.4]). *Let $\mathcal{S} = \{R(\xi) : \ \xi \in \mathbf{R}^{N\times n}\}$ and suppose $W : \mathbf{R}^{N\times n} \to \mathbf{R}$ is polyconvex. Define φ_W by (1). Then for each $X \in \mathbf{R}^d$ either one or both of the following hold:*

(a) *there exists $Y \in \mathcal{S}$ such that $\varphi_W|_{[Y,X]}$ is affine;*
(b) *there exists a unit vector $e \in \mathbf{R}^d$ such that for all $Y \in \mathbf{R}^d$ and all $t \in \mathbf{R}$ the function $t \mapsto \varphi_W(Y + te)$ is constant.*

The dichotomy can be sharp in the sense that (a) and not (b) can hold, as easy examples show, and that (b) and not (a) can hold, which is a consequence of the counterexample constructed below. It is shown in [3] that when (a) holds the differentiability of φ_W on $\mathcal{S}$ implies that φ_W is the unique convex representative. The counterexample below shows that this result is false when (b) holds and (a) does not.

1.1. Notation

We do not distinguish between the inner product of two matrices and the inner product of two vectors in $\mathbf{R}^5$, using $\cdot$ for both. Here, $\mathbf{R}^5$ is shorthand for $\mathbf{R}^{2\times 2} \times \mathbf{R}$, and in this case the inner product of (ξ, s) with (η, t) is given by

$$(\xi, s) \cdot (\eta, t) = \xi \cdot \eta + st,$$

where ξ, η are two matrices in $\mathbf{R}^{2\times 2}$, $s, t \in \mathbf{R}$ and

$$\xi \cdot \eta = \mathrm{tr}(\xi^T \eta).$$

Finally, if $a, b \in \mathbf{R}^2$ then the 2×2 matrix $a \otimes b$ has $(i, j)-$entry $a_i b_j$.

2. Construction of W such that $\varphi_W > \varphi_\tau$ on a large set

We restrict attention to polyconvex functions defined on $\mathbf{R}^{2\times 2}$, so that $R(\xi) = (\xi, \det \xi)$ for each 2×2 matrix ξ. To begin with we recall some basic facts about the subgradients of φ_W (for the definition of the subgradient of a convex function see [7]). When $W : \mathbf{R}^{2\times 2} \to \mathbf{R}$ is polyconvex and differentiable on an open set $U \subset \mathbf{R}^{2\times 2}$ it can be shown that for each $\xi \in U$

$$\partial\varphi_W(R(\xi)) = \{(DW(\xi) - \rho \mathrm{cof}\, \xi, \rho) : \ \rho_{\min}(\xi) \le \rho \le \rho_{\max}(\xi)\}, \tag{3}$$

where the functions $\rho_{\max}$, $\rho_{\min} : \mathbf{R}^{2\times 2} \to \mathbf{R}$ are defined by

$$\rho_{\max}(\xi) = \inf \left\{ \frac{W(\eta + \xi) - W(\xi) - DW(\xi) \cdot \eta}{\det \eta} : \ \det \eta > 0 \right\} \tag{4}$$

$$\rho_{\min}(\xi) = \sup \left\{ \frac{W(\eta + \xi) - W(\xi) - DW(\xi) \cdot \eta}{\det \eta} : \ \det \eta < 0 \right\}. \tag{5}$$

The proof of these assertions can be found in [3, Section 2]. Thus when $\xi \in U$, a sufficient condition for the differentiability of φ_W, and hence of φ_τ (because $\varphi_W \ge \varphi_\tau$ on $\mathbf{R}^5$, and because φ_W and φ_τ agree on $\mathcal{S}$—see [2, Corollary 2.5]), at $R(\xi)$ is that there exists a number $\rho(\xi)$ such that

$$W(\xi + \eta) \ge W(\xi) + DW(\xi) \cdot \eta + \rho(\xi) \det \eta$$

for all 2×2 matrices η.

Now let $[\xi] = \xi - \xi_{11} e_1 \otimes e_1$, where e_1 is the first canonical basis vector in $\mathbf{R}^2$, and define $W(\xi) = |([\xi], \det \xi - y)|$, where $|z|$ is the usual Euclidean norm in $\mathbf{R}^5$ and where y is a fixed positive number. It is easy to see that W is polyconvex and differentiable away from the set $\{\xi : \ W(\xi) = 0\}$, which, since $y \ne 0$, is empty. With the above remarks in mind the following proposition shows that φ_W is differentiable at all points $\mathbf{R}(\xi)$ in $\mathcal{S}$.

Proposition 2.1. *Let $\xi \in \mathbf{R}^{2\times 2}$ and let W be as above. Then for all η*

$$W(\xi + \eta) - W(\xi) - DW(\xi) \cdot \eta \ge \rho(\xi) \det \eta,$$

where $\rho(\xi) = \frac{(\det \xi - y)}{W(\xi)}$.

Proof. The inequality amounts to proving

$$|([\xi+\eta], \det(\xi+\eta)-y)| \geq \frac{1}{W(\xi)}([\xi+\eta]\cdot[\xi] + (\det\xi - y)(\det(\xi+\eta)-y)).$$

But this follows directly from the Cauchy-Schwarz inequality. □

Remark 2.2. *The choice of $\rho(\xi)$ in Proposition 2.1 is by analogy with the following example. Suppose $f(\xi) = |R(\xi)|$ and note that an obvious convex representative of f is $\varphi(\xi,\delta) = |(\xi,\delta)|$. Differentiating this with respect to δ, evaluating at $R(\xi)$, where $\xi \neq 0$, and referring to (3) gives a candidate $\rho(\xi) = \frac{\det\xi}{f(\xi)}$.*

Now consider the line $L := \mathrm{Span}\{e_1 \otimes e_1\}$. Clearly $\det l = 0$ for all $l \in L$. Since $D^2\det(\xi)[\eta,\eta] = 2\det\eta$ for all 2×2 matrices ξ and η, we can assume that the curvature of the graph of the determinant (i.e., the curvature of $\mathcal{S}$) is bounded above uniformly on the set $\{l+\eta : l \in L,\ |\eta| < 1\}$. In particular, we deduce that for sufficiently small $\epsilon > 0$ the (convex) tube

$$T_\epsilon := \{(l+\eta, y) :\ l \in L,\ |\eta| \leq \epsilon\},$$

which lies in $\mathbf{R}^5$, satisfies $\mathrm{dist}\,(T_\epsilon, \mathcal{S}) > 0$. With W as above it is claimed that $\varphi_W > \varphi_\mathcal{T}$ on the tube T_ϵ. Figure 2.1 below is intended as an analogy which may help the reader to visualize the idea behind the proof of Proposition 2.3.

Proposition 2.3. *Let $W(\xi) = |([\xi], \det\xi - y)|$ and assume ϵ has been chosen so that the tube T_ϵ does not meet $\mathcal{S}$. Then $\varphi_W(X) > \varphi_\mathcal{T}(X)$ for all $X \in T_\epsilon$.*

Proof. Recall that $\varphi_\mathcal{T}(X) = \sup\{a(X) : a \in \mathcal{T}\}$, where $\mathcal{T}$ consists of all those affine functions a satisfying $a(\xi, \det\xi) \leq W(\xi)$ for all $\xi \in \mathbf{R}^{2\times 2}$, and $a(\xi_0, \det\xi_0) = W(\xi_0)$ for at least one ξ_0. Suppose a_{ξ_0} is such that $a_{\xi_0}(\xi_0, \det\xi_0) = W(\xi_0)$. Standard arguments from convex analysis together with the differentiability of φ_W (Proposition 2.1 above) at all $(\xi_0, \det\xi_0)$ show that the gradient of the affine function a_{ξ_0} at $(\xi_0, \det\xi_0)$ must be $D\varphi_W(\xi_0, \det\xi_0)$. Since a_{ξ_0} is affine, and in view of (3), it follows that for all X in $\mathbf{R}^5$

$$\begin{aligned} a_{\xi_0}(X) &= W(\xi_0) + D\varphi_W(\xi_0, \det\xi_0)\cdot(X - (\xi_0, \det\xi_0)) \\ &= ([\hat{X}], X' - y))\cdot\frac{([\xi_0], \det\xi_0 - y)}{W(\xi_0)}. \end{aligned}$$

Here we have used the notation $X = (\hat{X}, X') \in \mathbf{R}^{2\times 2}\times\mathbf{R}$. Thus

$$\varphi_\mathcal{T}(X) = \sup\left\{([\hat{X}], X'-y))\cdot\frac{([\xi_0], \det\xi_0 - y)}{W(\xi_0)} :\ \xi_0 \in \mathbf{R}^{2\times 2}\right\}. \tag{6}$$

Provided we can find ξ_0 such that $([\hat{X}], X'-y))$ and $([\xi_0], \det\xi_0 - y)$ are parallel, or asymptotically parallel (which will be made clear below), then it will follow essentially from the Cauchy-Schwarz inequality that $\varphi_\mathcal{T}(X) = |([\hat{X}], X'-y))|$. There are three cases to consider, and in doing so we shall refer to the unit vector $\frac{([\xi_0], \det\xi_0 - y)}{W(\xi_0)}$ by $u(\xi_0)$.

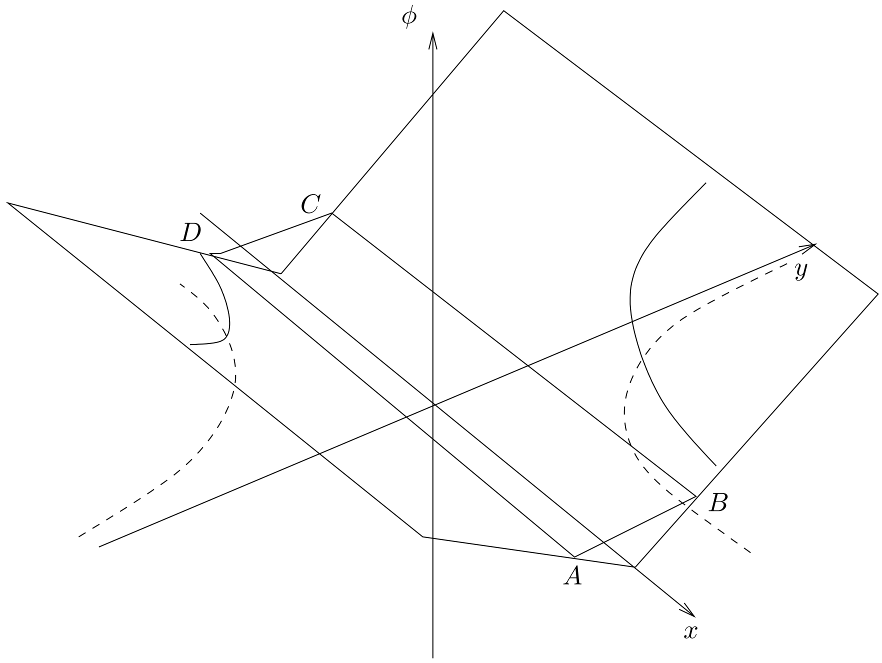

Figure 2.1: A graphical interpretation of the constructions of W, φ_W and φ_τ. $\mathcal{S}$ can be thought of as the union of the two curves in the $x-y$ plane, the graph of φ_W as the union of the plane $ABCD$ together with the two sloping planes it meets at AD and BC, and the graph of φ_τ as the union of the two sloping planes. The function W is represented by the restriction of φ_W to $\mathcal{S}$. Clearly, $\varphi_W > \varphi_\tau$ in the projection of $ABCD$ in the $x-y$ plane.

(i) $[\hat{X}] = 0$. Note that $u(0) = (0, -1)$, which gives $\varphi_\tau(X) = |X' - y|$ provided $y > X'$. Otherwise note that $u(kQ) \to (0, 1)$ as $k \to \infty$ whenever Q is a rotation matrix (i.e. $Q \in SO(2)$), which implies $u(kQ) \cdot (0, X' - y) \to |X' - y|$ whenever $X' > y$. If $y = X'$ then $\varphi_\tau(X) = |X' - y|$ trivially.

(ii) $\hat{X}_{22} \neq 0$. Set $\xi_0 = [\hat{X}]$ and consider $\xi_\mu = \xi_0 + \mu e_1 \otimes e_1$. We require $\det \xi_\mu = X'$. But this can easily be satisfied by an appropriate choice of μ, and on using $\hat{X}_{22} \neq 0$ in $\det \xi_\mu = \det \xi_0 + \mu \hat{X}_{22}$.

(iii) $[\hat{X}] \neq 0$, $\hat{X}_{22} = 0$. As before, choose ξ_0 to satisfy $\xi_0 = [\hat{X}]$ and let $\xi_{\mu,\nu} = \xi_0 + \mu e_1 \otimes e_1 + \nu e_2 \otimes e_2$, where μ and ν are parameters. Now we seek μ and ν such that $\det \xi_{\mu,\nu} = X'$, that is,

$$\mu\nu = X' + \hat{X}_{12}\hat{X}_{21}. \tag{7}$$

But $[\xi_{\mu,\nu}] = [\hat{X}] + \nu e_2 \otimes e_2$, and hence

$$u(\xi_{\mu,\nu}) \to \frac{([\hat{X}], X' - y)}{|([\hat{X}], X' - y)|}$$

provided $\mu \to \infty$ and $\nu \to 0$ consistent with (7).

Thus in each case we have $\varphi_\tau(X) = |([\hat{X}], X' - y)|$. To conclude the proof note that $W(\xi)$ can be interpreted as the distance of the point $(\xi, \det \xi)$ to the centre of the tube T_ϵ. The construction of T_ϵ above therefore implies that $W(\xi) \geq \epsilon$ for all 2×2 matrices ξ. Hence $\varphi_W(X) \geq \epsilon$ for all X, while $\varphi_\tau(X) < \epsilon$ for all X inside the tube T_ϵ. □

With reference to the statement of [3, Lemma 2.4] given in the introduction, we remark that because alternative (a) of [3, Lemma 2.4] fails for points X in the tube T_ϵ it must be that (b) holds for such points. It was shown in [3, Proposition 3.5] that if alternative (a) held at all X and if φ_W was differentiable on $\mathcal{S}$ then φ_W was the unique convex representative of W. This result is clearly false when alternative (b) holds at some X, even when, as we have seen above, φ_W is differentiable on $\mathcal{S}$.

Acknowledgements. I thank Prof. B. Kirchheim for reading a draft version of the paper and for the idea leading to Figure 2.1.

References

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

[2] J. M. Ball, B. Kirchheim, J. Kristensen: Regularity of quasiconvex envelopes, Calc. Var. Partial Differ. Equ. 11 (2000) 333–359.

[3] J. Bevan: On convex representatives of polyconvex functions, Proc. R. Soc. Edinb., Sect. A, Math. 136(1) (2006) 23–51.

[4] H. Busemann, G. Ewald, G. C. Shepherd: Convex bodies and convexity on Grassmann cones. I–IV, Math. Ann. 151 (1963) 1–41.

[5] B. Kirchheim: Private communication.

[6] C. B. Morrey, Jr.: Quasiconvexity and the lower semicontinuity of multiple integrals, Pacific J. Math. 2 (1952) 25–53.

[7] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton (1970).