A remark on the structure of the Busemann representative of a polyconvex function

J. J. Bevan †

September 23, 2008

1 Introduction

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

\[
W(\xi) = \varphi(R(\xi))
\]

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

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].

The Busemann representative \( \varphi_W \) of a given polyconvex function \( W : \mathbb{R}^{N \times n} \to \mathbb{R} \cup \{ \infty \} \) can be expressed as

\[
\varphi_W(X) = \inf \left\{ \sum_{i=1}^{d+1} \lambda_j W(\xi_j) : \lambda_j \geq 0, \sum_{j=1}^{d+1} \lambda_j = 1 \text{ and } \sum_{j=1}^{d+1} \lambda_j R(\xi_j) = X \right\}.
\] (1.1)

Here, \( d \) is the least integer such that \( R(\xi) \in \mathbb{R}^d \) for all \( \xi \in \mathbb{R}^{N \times n} \) and \( X \) lies in \( \mathbb{R}^d \). Busemann et al. proved that

\[
\varphi_W(X) = \sup \{ a(X) : a \in \mathcal{L} \},
\] (1.2)
where
\[ \mathcal{L} = \{ \phi \text{ affine} : \phi(R(\xi)) \leq W(\xi) \ \forall \ \xi \in \mathbb{R}^{N \times n} \}. \]

The graph of any \( \phi \in \mathcal{L} \) is a hyperplane, so (1.2) states that \( \phi_w \) is built from hyperplanes which lie below the set \( G_w := \{(R(\xi), W(\xi)) : \ \xi \in \mathbb{R}^{N \times n}\} \).

The main result in this short note is that there is no redundancy in the expression (1.2) in the case \( N = n = 2 \). To be precise, one cannot replace \( \mathcal{L} \) in (1.2) by the smaller class \( \mathcal{T} \), where
\[ \mathcal{T} = \{ \phi \in \mathcal{L} : \exists \ \xi \in \mathbb{R}^{2 \times 2} \text{ s.t. } W(\xi) = \phi(R(\xi)) \}. \]

Thus \( \mathcal{T} \) represents the collection of supporting hyperplanes to \( G_w \) which meet \( G_w \) in at least one point. We define
\[ \varphi_\tau(X) = \sup\{a(X) : a \in \mathcal{T}\}. \]

Note that \( \varphi_w \geq \varphi_\tau \) in view of the inclusion \( \mathcal{T} \subset \mathcal{L} \). It is proved in the next section that for a certain choice of \( W \) it is the case that \( \varphi_w > \varphi_\tau \) on a large set. This result is surprising since the set \( \{R(\xi) : \ \xi \in \mathbb{R}^{2 \times 2}\} \) is large: its convex hull is the whole of \( \mathbb{R}^5 \). (For a proof of this fact see [1].) Certainly one might expect \( \varphi_w = \varphi_\tau \) 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 \( \varphi_w \) 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 \( \mathcal{S} = \{R(\xi) : \ \xi \in \mathbb{R}^{N \times n}\} \) and suppose \( W : \mathbb{R}^{N \times n} \to \mathbb{R} \) is polyconvex. Define \( \varphi_w \) by (1.1). Then for each \( X \in \mathbb{R}^d \) 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 \( e \in \mathbb{R}^d \) such that for all \( Y \in \mathbb{R}^d \) and all \( t \in \mathbb{R} \) 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 \( \varphi_w \) on \( \mathcal{S} \) implies that \( \varphi_w \) 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 \( \mathbb{R}^5 \), using \( \cdot \) for both. Here, \( \mathbb{R}^5 \) is shorthand for \( \mathbb{R}^{2 \times 2} \times \mathbb{R} \), and in this case the inner product of \((\xi, s)\) with \((\eta, t)\) is given by
\[ (\xi, s) \cdot (\eta, t) = \xi \cdot \eta + st, \]
where $\xi, \eta$ are two matrices in $\mathbb{R}^{2 \times 2}$, $s, t \in \mathbb{R}$ and
\[ \xi \cdot \eta = \text{tr}(\xi^T \eta). \]
Finally, if $a, b \in \mathbb{R}^2$ then the $2 \times 2$ 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 $\mathbb{R}^{2 \times 2}$, so that $R(\xi) = (\xi, \det \xi)$ for each $2 \times 2$ matrix $\xi$. To begin with we recall some basic facts about the subgradients of $\varphi_W$ (for the definition of the subgradient of a convex function see [7]). When $W : \mathbb{R}^{2 \times 2} \to \mathbb{R}$ is polyconvex and differentiable on an open set $U \subset \mathbb{R}^{2 \times 2}$ it can be shown that for each $\xi \in U$
\[ \partial \varphi_W(R(\xi)) = \{(DW(\xi) - \rho \text{cof} \xi, \rho) : \rho_{\max}(\xi) \leq \rho \leq \rho_{\max}(\xi)\}, \quad (2.1) \]
where the functions $\rho_{\max}, \rho_{\min} : \mathbb{R}^{2 \times 2} \to \mathbb{R}$ are defined by
\[ \rho_{\max}(\xi) = \min \left\{ \frac{W(\eta + \xi) - W(\xi) - DW(\xi) \cdot \eta}{\det \eta} : \det \eta > 0 \right\}, \quad (2.2) \]
\[ \rho_{\min}(\xi) = \max \left\{ \frac{W(\eta + \xi) - W(\xi) - DW(\xi) \cdot \eta}{\det \eta} : \det \eta < 0 \right\}. \quad (2.3) \]

The proof of these assertions can be found in [3, Section 2]. Thus when $\xi \in U$, a sufficient condition for the differentiability of $\varphi_W$, and hence of $\varphi_\tau$ (because $\varphi_W \geq \varphi_\tau$ on $\mathbb{R}^5$, and because $\varphi_W$ and $\varphi_\tau$ agree on $\mathcal{S}$—see [2, Corollary 2.5]), at $R(\xi)$ is that there exists a number $\rho(\xi)$ such that
\[ W(\xi + \eta) \geq W(\xi) + DW(\xi) \cdot \eta + \rho(\xi) \det \eta \]
for all $2 \times 2$ matrices $\eta$.

Now let $[\xi] = \xi - \xi_1 e_1 \otimes e_1$, where $e_1$ is the first canonical basis vector in $\mathbb{R}^2$, and define $W(\xi) = |[\xi], \det (\xi - y)|$, where $|z|$ is the usual Euclidean norm in $\mathbb{R}^5$ 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 $y \neq 0$, is empty. With the above remarks in mind the following proposition shows that $\varphi_W$ is differentiable at all points $R(\xi)$ in $\mathcal{S}$.

**Proposition 2.1.** Let $\xi \in \mathbb{R}^{2 \times 2}$ and let $W$ be as above. Then for all $\eta$
\[ W(\xi + \eta) - W(\xi) - DW(\xi) \cdot \eta \geq \rho(\xi) \det \eta, \]
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. \( \square \)
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 $\delta$, evaluating at $R(\xi)$, where $\xi \neq 0$, and referring to (2.1) gives a candidate $\rho(\xi) = \frac{\det \xi}{f(\xi)}$.

Now consider the line $L := \text{Span}\{e_1 \otimes e_1\}$. Clearly det $l = 0$ for all $l \in L$. Since $D^2 \det(\xi)[\eta, \eta] = 2\det \eta$ for all $2 \times 2$ matrices $\xi$ and $\eta$, we can assume that the curvature of the graph of the determinant (i.e., the curvature of $S$) 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 $\mathbb{R}^5$, satisfies dist $(T_\epsilon, S) > 0$. With $W$ as above it is claimed that $\varphi_w(\xi) > \varphi_\tau(\xi)$ on the tube $T_\epsilon$. Figure 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 $\epsilon$ has been chosen so that the tube $T_\epsilon$ does not meet $S$. Then $\varphi_w(\xi) > \varphi_\tau(\xi)$ for all $X \in T_\epsilon$.

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

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

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

$$\varphi_\tau(X) = \sup \left\{([\hat{X}], X' - y) \cdot \frac{([\xi_0], \det \xi_0 - y)}{W(\xi_0)}: \xi_0 \in \mathbb{R}^{2 \times 2}\right\}. \quad (2.4)$$

Provided we can find $\xi_0$ 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 $\varphi_\tau(X) = ||([\hat{X}], X' - y)||$. 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 $u(\xi_0)$.

(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.
Figure 1: A graphical interpretation of the contructions of $W$, $\varphi_w$ and $\varphi$. $S$ can be thought of as the union of the two curves in the $x - y$ plane, the graph of $\varphi_w$ as the union of the plane $ABCD$ together with the two sloping planes it meets at $AD$ and $BC$, and the graph of $\varphi$ as the union of the two sloping planes. The function $W$ is represented by the restriction of $\varphi_w$ to $S$; its graph is shown with dotted lines. Clearly, $\varphi_w > \varphi$ in the projection of $ABCD$ in the $x - y$ plane.

(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 $\mu$, 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 $\xi_0$ 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 $\mu$ and $\nu$ are parameters. Now we seek $\mu$ and $\nu$ such that $\det \xi_{\mu, \nu} = X'$, that is,

$$\mu \nu = X' + \hat{X}_{12} \hat{X}_{21}. \tag{2.5}$$

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

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

provided $\mu \rightarrow \infty$ and $\nu \rightarrow 0$ consistent with (2.5).
Thus in each case we have \( \varphi_r(X) = |([\hat{X}], X' - y)| \). 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 \( T_\epsilon \). The construction of \( T_\epsilon \) above therefore implies that \( W(\xi) \geq \epsilon \) for all \( 2 \times 2 \) matrices \( \xi \). Hence \( \varphi_w(X) \geq \epsilon \) for all \( X \), while \( \varphi_r(X) < \epsilon \) for all \( X \) inside the tube \( T_\epsilon \).

\[\square\]

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 \( T_\epsilon \) 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 \( \varphi_w \) was differentiable on \( S \) then \( \varphi_w \) 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, \( \varphi_w \) is differentiable on \( S \).

Acknowledgement 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. I thank Prof. B. Kirchheim for reading a draft version of the paper and for the idea leading to Figure 1.

References

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

[2] J.M. Ball, B. Kirchheim, J. Kristensen. Regularity of quasiconvex envelopes. Calc. Var., 11:333-359, 2000.

[3] J. Bevan. On convex representatives of polyconvex functions. Proc. Roy. Soc. Edinburgh Sect. A 136 (2006), no. 1, 23–51.

[4] H. Busemann, G. Ewald and G.C. Shepherd. Convex bodies and convexity on Grassman cones, parts I-IV. Math. Ann., 151:1-41, 1963.

[5] B. Kirchheim. Private Communication.

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

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