Convexity and Coercivity in nonlinear, anisotropic elasticity and some useful relations.
CISM-NOTES.

Patrizio Neff

Department of Mathematics, Technische Universität Darmstadt

February 27, 2008

Contents

1 Introduction

These notes present relations which the author has found useful in treating convexity questions in nonlinear elasticity. They cannot be exhaustive and remain most of the time on a formal level. The notes are mainly meant for engineering students wishing to complement their mathematical knowledge. One of the goals was to provide a lot of detailed calculations which show how the notions of convexity work.

After working through these notes the student should understand the meaning of convexity, polyconvexity and ellipticity. He should know how to prove or disprove the convexity and ellipticity in simple circumstances, avoid the pitfalls of convexity conditions, understand the meaning of coercivity and prove coercivity in different situations. Since we are in a vectorvalued setting, the student needs a good deal of knowledge in matrix analysis and in manipulating tensor differentiation. The presented material on matrix analysis can be complemented by studying the books of Horn/Johnson [?, ?].

As far as tensor differentiation is concerned the author has found it useful to calculate the differentials directly based on the Taylor expansion without calculating cumbersome partial derivatives. The advantages are that the Hesse-matrix of second derivatives of a scalarvalued function of a matrix-valued argument $W : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ is immediately available in terms of the generated bilinear form. This is, in my opinion, a decisive advantage.

When it comes to checking polyconvexity of a given function, the second derivative approach has its limitations: in the relevant applications one has to check the convexity of a function $P : \mathbb{R}^{19} \mapsto \mathbb{R}$. In the isotropic case the investigations are facilitated by going over to the eigenvalue representation. However, in the anisotropic setting this reduction step is not possible. Instead of checking the convexity of P by direct means, it will be easier to provide constructive rules of how to obtain polyconvex functions from simpler ones.

Good sources for covering the mathematical background are Ciarlet [?], Marsden/Hughes [?] and Dacorogna [?, ?]. For an introduction to the theory of function spaces and partial differential equations the work of Evans [?, ?] is a valuable source. More about nonlinear elasticity can be found in Ogden [?], Silhavy [?], Gurtin [?], Holzapfel [?] and the forthcoming book of Schröder [?].

Of the many important contributions of Sir John Ball the articles [?, ?, ?] are also in reach of a non-expert audience.

An unsolved problem

Given a class of anisotropic, frame-indifferent materials with positive definite elasticity tensor $\mathcal{C}$ in the stress-free reference configuration $1\!\!1$: is it always (constructive?) possible to find a polyconvex (quasiconvex, elliptic) function W such that

$$D^2W(1\!\!1).(H,H) = \langle \mathcal{C}.H, H\rangle \,.$$

In the isotropic case, the positive answer for polyconvexity (and thence quasiconvexity, ellipticity) has been given by Ciarlet [?].

2 Mathematical background

2.1 Some linear algebra and bilinear forms

2.2 Trace, scalarproduct and norm

We let $\mathbb{M}^{3\times 3}$ denote the set of three by three matrices and GL(3) be the set of invertible three by three matrices. Let $1\!\!1$ be the identity tensor on $\mathbb{M}^{3\times 3}$. We recall the property of the trace:

Lemma 2.1
Let $A, B \in \mathbb{M}^{3\times 3}$ *then* $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Proof.

$$\begin{aligned}\operatorname{tr}(AB) &= \operatorname{tr}(\sum_j A_{ij}B_{jk}) = \sum_k \left(\sum_j A_{kj}B_{jk}\right)\\ &= \sum_k \sum_j B_{kj}A_{jk} = \operatorname{tr}(\sum_j B_{kj}A_{ji}) = \operatorname{tr}(BA)\,.\end{aligned} \qquad \blacksquare$$

We define a scalarproduct on $\mathbb{M}^{3\times 3}$ by setting:

Definition 2.2 (Scalarproduct on $\mathbb{M}^{3\times 3}$)

$$\langle A, B\rangle := \operatorname{tr}(A\,B^T)\,.$$

Lemma 2.3 ($\langle A, B\rangle$ is a scalarproduct on $\mathbb{M}^{3\times 3}$)

Proof. The bilinearity of $\langle A, B\rangle$ is clear. The symmetry follows from invariance of the trace under transposition:

$$\langle A, B\rangle = \operatorname{tr}(AB^T) = \operatorname{tr}((AB^T)^T) = \operatorname{tr}(BA^T) = \langle B, A\rangle \tag{2.1}$$

and the positivity of $\langle A, A\rangle$ can be seen in the following way

$$\langle A, A\rangle = \operatorname{tr}(AA^T) = \sum_{i,j=1}^{3} A_{ij}\cdot A_{ij} = \sum_{i,j=1}^{3} A_{ij}^2 > 0\,,$$

for $A \neq 0$. $\blacksquare$

Remark 2.4 (Trace and scalarproduct)
With the introduced notation we have $\operatorname{tr}(A) = \langle A, 1\!\!1\rangle$ *where* $1\!\!1$ *is the second order identity tensor.*

We may define a corresponding norm:

Definition 2.5 (Euclidean norm on $\mathbb{M}^{3\times 3}$)
Let $A \in \mathbb{M}^{3\times 3}$ then

$$\|A\| := \sqrt{\langle A, A\rangle}$$

defines a norm.

Proof. Obvious. ■
For $Q \in O(3)$ it holds then $\|Q\| = \sqrt{\|Q\|^2} = \sqrt{\langle Q, Q\rangle} = \sqrt{\operatorname{tr}(QQ^T)} = \sqrt{\operatorname{tr}(1\!\!1)} = \sqrt{3}$.

Lemma 2.6 (Compatible matrixnorm)
Let $A \in \mathbb{M}^{3\times 3}$. Then $\|A.\eta\|_{\mathbb{R}^3} \leq \|A\| \cdot \|\eta\|_{\mathbb{R}^3}$.

Proof. Well known. Subsequently we will skip the index $\mathbb{R}^3$ if there is no danger of confusion. ■
The introduced scalarproduct on matrices exhibits some useful commutation properties. It holds

Lemma 2.7 (Commutativity)
Let $A, B, C \in \mathbb{M}^{3\times 3}$ then

$$\langle A, B\,C\rangle = \langle A\,C^T, B\rangle = \langle B^T\,A, C\rangle = \langle C, B^T\,A\rangle\,.$$

Proof.

$$\begin{aligned}\langle A, BC\rangle &= \operatorname{tr}(A(BC)^T) = \operatorname{tr}(AC^TB^T) = \operatorname{tr}((AC^T)B^T)\\ &= \langle AC^T, B\rangle = \operatorname{tr}(AC^TB^T) \quad \text{with (\textbf{??})}\\ &= \operatorname{tr}((AC^T)B^T) = \operatorname{tr}(B^T(AC^T)) = \operatorname{tr}((B^TA)C^T)\\ &= \langle B^TA, C\rangle \quad \text{and with (\textbf{??})}\\ &= \langle C, B^TA\rangle\,.\end{aligned}$$ ■

This implies immediately

Corollary 2.8

$$\langle A, B\rangle = \langle A^T, B^T\rangle\,.$$

Proof.

$$\begin{aligned}\langle A, B\rangle &= \langle A, B\cdot 1\!\!1\rangle \quad \text{with (\textbf{??})}\\ &= \langle AB^T, 1\!\!1\rangle = \langle B^T, A^T\cdot 1\!\!1\rangle = \langle B^T, A^T\rangle\,.\end{aligned}$$ ■

Lemma 2.9 (Symmetric matrices and scalarproduct)
Let $S \in \operatorname{Sym}(3)$, where $\operatorname{Sym}(3)$ denotes the symmetric three by three matrices and let $T \in \mathbb{M}^{3\times 3}$. Then

$$\begin{aligned}&\langle S, T\rangle = \langle S^T, T\rangle = \langle S, T^T\rangle = \langle S, \tfrac{1}{2}(T^T + T)\rangle\,,\\ &\forall\; S \in \operatorname{Sym}(3): \quad \langle S, T\rangle = 0 \Rightarrow T = -T^T\,.\end{aligned}$$

Proof. The first part is obvious. For the second statement assume the contrary, i.e. T is not antisymetric, but $\forall\ S \in \mathrm{Sym}(3): \quad \langle S, T\rangle = 0$. Now T has a nonzero symmetric part $\mathrm{sym}(T) = \frac{1}{2}T^T + T$. Choose $S = \mathrm{sym}(T)$ in the first statement. This yields $\langle S, T\rangle = \langle \mathrm{sym}(T), \mathrm{sym}(T)\rangle = \|\mathrm{sym}(T)\|^2 = 0$. A contradiction. ■

Lemma 2.10 (Invariance of the trace)
Let $A \in \mathbb{M}^{3\times 3}$ *and* $S \in \mathrm{GL}(3,\mathbb{R})$. *Then* $\mathrm{tr}(S^{-1}AS) = \mathrm{tr}(A)$.

Proof.

$$\mathrm{tr}(S^{-1}AS) = \langle S^{-1}AS, 1\!\!1\rangle = \langle AS, S^{-T}1\!\!1\rangle = \langle A, S^{-T}S^T\rangle = \langle A, 1\!\!1\rangle = \mathrm{tr}(A)\,. \qquad ■$$

Lemma 2.11 (Invariance of the norm under orthogonal mappings)
Let $A \in \mathbb{M}^{3\times 3}$ *and* $Q \in O(3)$. *Then* $\|QA\| = \|A\|$ *and* $\|AQ\| = \|A\|$.

Proof.

$$\|QA\|^2 = \langle QA, QA\rangle = \langle A, Q^TQA\rangle = \langle A, A\rangle = \|A\|^2\,. \qquad ■$$

Corollary 2.12
For $A \in \mathbb{M}^{3\times 3}$ *and* $Q \in O(3)$ *it holds* $\|Q^TAQ\| = \|A\|$. ■

2.3 Positive definite bilinear forms

Consider the set $\mathrm{PSym}(3)$ of positive definite three by three matrices. A useful observation is

Theorem 2.13
Let $A \in \mathrm{PSym}(3)$ *and* B *be positiv semi-definite. Then*

$$\mathrm{tr}(AB) \geq 0\,,$$

while in general $A\,B \notin \mathrm{PSym}(3)$.

Proof. Since $A \in \mathrm{PSym}(3)$, there exists an orthogonal matrix $Q \in O(3)$ which diagonalizes A, i.e., $Q^TAQ = D_A$. In D_A appear the positive Eigenvalues of A since $A \in \mathrm{PSym}(3)$. The matrix $\hat{B} := Q^TBQ$ is still positive semi-definite since for arbitrary $\xi \in \mathbb{R}^3$ it holds

$$\langle \xi, \hat{B}.\xi\rangle = \langle \xi, Q^TBQ.\xi\rangle = \langle Q.\xi, BQ.\xi\rangle \geq 0\,.$$

Now

$$\mathrm{tr}(AB) = \mathrm{tr}(Q^TABQ) = \mathrm{tr}(Q^TA1\!\!1BQ) = \mathrm{tr}(Q^TAQQ^TBQ) = \mathrm{tr}(D_A\hat{B})\,.$$

But

$$\begin{aligned}\mathrm{tr}(D_A\hat{B}) &= \mathrm{tr}\left(\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{pmatrix}\cdot\begin{pmatrix}\hat{b}_{11} & \hat{b}_{12} & \hat{b}_{13}\\ \hat{b}_{21} & \hat{b}_{22} & \hat{b}_{23}\\ \hat{b}_{31} & \hat{b}_{32} & \hat{b}_{33}\end{pmatrix}\right)\\ &= \mathrm{tr}\begin{pmatrix}\lambda_1\hat{b}_{11} & \lambda_1\hat{b}_{12} & \lambda_1\hat{b}_{13}\\ \lambda_2\hat{b}_{21} & \lambda_2\hat{b}_{22} & \lambda_2\hat{b}_{23}\\ \lambda_3\hat{b}_{31} & \lambda_3\hat{b}_{32} & \lambda_3\hat{b}_{33}\end{pmatrix} = \sum_{i=1}^{3}\lambda_i\,\hat{b}_{ii} \geq 0\,,\end{aligned}$$

because $\hat{b}_{ii} \geq 0$ since $\hat{B}$ is positive semi-definite and therefore has only non-negative diagonal entries (test with e_i). ■

Theorem 2.14
Let D be a positive definite diagonal matrix and let $B \in \mathbb{M}^{3\times 3}$. Then

$$\langle DB, BD\rangle \geq \lambda^2_{min}(D) \cdot \|B\|^2 .$$

Proof. $\langle DB, BD\rangle = \langle DBD, B\rangle$ and

$$\begin{aligned} DBD &= \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 b_{11} & \lambda_1 b_{12} & \lambda_1 b_{13} \\ \lambda_2 b_{21} & \lambda_2 b_{22} & \lambda_2 b_{23} \\ \lambda_3 b_{31} & \lambda_3 b_{32} & \lambda_3 b_{33} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1\lambda_1 b_{11} & \lambda_1\lambda_2 b_{12} & \lambda_1\lambda_3 b_{13} \\ \lambda_2\lambda_1 b_{21} & \lambda_2\lambda_2 b_{22} & \lambda_2\lambda_3 b_{23} \\ \lambda_3\lambda_1 b_{31} & \lambda_3\lambda_2 b_{32} & \lambda_3\lambda_3 b_{33} \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \langle DBD, B\rangle &= \langle \begin{pmatrix} \lambda_1\lambda_1 b_{11} & \lambda_1\lambda_2 b_{12} & \lambda_1\lambda_3 b_{13} \\ \lambda_2\lambda_1 b_{21} & \lambda_2\lambda_2 b_{22} & \lambda_2\lambda_3 b_{23} \\ \lambda_3\lambda_1 b_{31} & \lambda_3\lambda_2 b_{32} & \lambda_3\lambda_3 b_{33} \end{pmatrix}, \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \rangle \\ &= \sum_{i,j=1}^{3} \lambda_i \lambda_j b_{ij}^2 \geq \lambda^2_{\min}(D) \sum_{i,j=1}^{3} b_{ij}^2 = \lambda^2_{\min}(D)\|B\|^2 , \end{aligned}$$

since $\lambda_i\, \lambda_j \geq 0$ by assumption. ■

Corollary 2.15
Let D be a positive definite diagonal matrix and $B \in \mathbb{M}^{3\times 3}$. Then

$$\langle BD, B\rangle \geq \lambda_{\min}(D)\|B\|^2 .$$

Proof.

$$\langle BD, B\rangle = \langle B\sqrt{D}\sqrt{D}, B\rangle = \langle B\sqrt{D}, B\sqrt{D}\rangle \geq \sqrt{\lambda^2_{\min}(D)}\|B\|^2 = \lambda_{\min}(D)\|B\|^2 .$$ ■

Corollary 2.16
Let D be a positive diagonal matrix and $B \in \mathbb{M}^{3\times 3}$. Then

$$\langle DB, B\rangle \geq \lambda_{\min}(D)\|B\|^2 .$$

Proof.

$$\begin{aligned} \langle DB, B\rangle &= \langle \sqrt{D}\sqrt{D}B, B\rangle = \langle \sqrt{D}B, \sqrt{D}B\rangle = \langle (\sqrt{D}B)^T, (\sqrt{D}B)^T\rangle \\ &= \langle B^T\sqrt{D}, B^T\sqrt{D}\rangle \geq \lambda_{\min}(D)\|B^T\|^2 = \lambda_{\min}(D)\ \|B\|^2 \end{aligned}$$ ■

Corollary 2.17
Let $A \in \mathbb{M}^{3\times 3}$ *and* $B \in \operatorname{PSym}(3)$*. Then* $\langle AB, BA\rangle \geq \lambda_{\min}^2(B)\|A\|^2$*.*

Proof. Set $B = Q^T D_B Q$ where D_B is the diagonalization of B. Then

$$\begin{aligned}\langle AB, BA\rangle = \langle AQ^T D_B Q, Q^T D_B QA\rangle &= \langle QAQ^T D_B, D_B QAQ^T\rangle \\ &\text{with Theorem ?? it follows} \\ &\geq \lambda_{\min}^2(D_B)\|QAQ^T\|^2 = \lambda_{\min}^2(D_B)\|A\|^2 = \lambda_{\min}^2(B)\cdot\|A\|^2 .\end{aligned}$$ ∎

Corollary 2.18
Let $A, B \in \operatorname{PSym}$*. Then* $\langle AB, BA\rangle \geq \lambda_{\min}^2(A)\|B\|^2$ *and* $\langle AB, BA\rangle \geq \lambda_{\min}^2(B)\,\|A\|^2$*.*

Proof.

$$\begin{aligned}\langle AB, BA\rangle &= \langle Q^T D_A QB, BQ^T D_A Q\rangle = \langle D_A QB, QBQ^T D_A Q\rangle = \langle D_A QBQ^T, QBQ^T D_A\rangle \\ &\geq \lambda_{\min}^2(D_A)\|QBQ^T\|^2 = \lambda_{\min}^2(D_A)\|B\|^2 = \lambda_{\min}^2(A)\|B\|^2 .\end{aligned}$$ ∎

Corollary 2.19
Let $F \in \operatorname{GL}(3,\mathbb{R})$ *and* $H \in \mathbb{M}^{3\times 3}$*. Then*

$$\|F^T H F\|^2 \geq \lambda_{\min}^2(FF^T)\cdot\|H\|^2 .$$

Proof.

$$\|F^T HF\|^2 = \langle F^T HF, F^T HF\rangle = \langle FF^T H, HFF^T\rangle \geq \lambda_{\min}^2(FF^T)\|H\|^2 .$$ ∎

Corollary 2.20
Let $F \in \operatorname{GL}(3,\mathbb{R})$ *and* $H \in \mathbb{M}^{3\times 3}$*. Then*

$$\|F^T H + H^T F\|^2 \geq \lambda_{\min}^2(FF^T)\cdot\|HF^{-1} + (HF^{-1})^T\|^2 .$$

Proof.

$$\begin{aligned}\|F^T H + H^T F\|^2 &= \|F^T(HF^{-1} + F^{-T}H^T)F\|^2 \\ &\geq \lambda_{\min}^2(FF^T)\cdot\|HF^{-1} + (HF^{-1})^T\|^2 .\end{aligned}$$ ∎

Corollary 2.21
Let $F, H \in \mathbb{M}^{3\times 3}$*. Then* $\|FH\|^2 \geq \lambda_{\min}(F^T F)\|H\|^2$*.*

Proof.

$$\begin{aligned}\|FH\|^2 &= \langle FH, FH\rangle = \langle H, F^T FH\rangle = \langle HH^T, F^T F\rangle \\ &= \langle HH^T, Q^T D_{F^TF} Q\rangle = \langle QHH^T Q^T, D\rangle \\ &= \langle (QH)(QH)^T, D\rangle = \langle (QH), D(QH)\rangle \\ &\geq \lambda_{\min}(D)\|QH\|^2 = \lambda_{\min}(D)\langle QH, QH\rangle \\ &= \lambda_{\min}(D)\langle H, Q^T QH\rangle = \lambda_{\min}(D)\langle H, H\rangle \\ &= \lambda_{\min}(D)\|H\|^2 = \lambda_{\min}(F^T F)\|H\|^2 .\end{aligned}$$ ∎

Corollary 2.22
Let $F, H \in \mathbb{M}^{3\times 3}$ *and* $\det[F^TF] = 1$. *Then* $\|FH\|^2 \geq \frac{1}{\|F^TF\|^2} \cdot \|H\|^2$.

Proof. With the previous Lemma we only need to estimate $\lambda_{\min}(F^TF)$. We have

$$\begin{aligned}
1 &= \lambda_{\min}(F^TF) \cdot \lambda_2(F^TF) \cdot \lambda_{\max}(F^TF) \\
\lambda_{\min}(F^TF) &= \frac{1}{\lambda_2(F^TF) \cdot \lambda_{\max}(F^TF)} \\
&\geq \frac{1}{\lambda_{\max}(F^TF) \cdot \lambda_{\max}(F^TF)} = \frac{1}{\lambda_{\max}(F^TF)^2} \\
&\text{and let } \eta \text{ be an eigenvector corresponding to } \lambda_{\max} \text{ then} \\
\lambda_{\max} &= \frac{\langle F^TF.\eta, \eta\rangle}{\|\eta\|^2} \leq \|F^TF\|
\end{aligned}$$

hence for $\lambda_{\min}$:

$$\lambda_{\min}(F^TF) \geq \frac{1}{\|F^TF\|^2}\,. \qquad \blacksquare$$

Corollary 2.23
If $A, H \in \mathbb{M}^{3\times 3}$ *then* $\langle A^TAH^TH, H^THA^TA\rangle \geq 0$.

Proof.

$$\begin{aligned}
\langle A^TAH^TH, H^THA^TA\rangle &= \langle Q^TD_{A^TA}QH^TH, H^THQ^TDQ\rangle \\
&= \langle DQH^THQ^T, QH^THQ^TD\rangle \geq \lambda^2_{\min}(D)\|QH^THQ^T\|^2 \\
&= \lambda^2_{\min}(D)\|H^TH\|^2 = \lambda^2_{\min}(A^TA)\|H^TH\|^2\,. \qquad \blacksquare
\end{aligned}$$

Corollary 2.24
Let $F, H \in \mathbb{M}^{3\times 3}$. *Then* $\langle F^TF, H^TH\rangle \geq \lambda_{\min}(F^TF) \cdot \|H\|^2$.

Proof.

$$\begin{aligned}
\langle F^TF, H^TH\rangle &= \langle Q^TD_{F^TF}Q, H^TH\rangle = \langle D, QH^THQ^T\rangle = \langle D, (HQ^T)^T(HQ^T)\rangle \\
&= \langle (HQ^T)D, HQ^T\rangle \geq \lambda_{\min}(D)\|HQ^T\|^2 \\
&= \lambda_{\min}(D)\|H\|^2 = \lambda_{\min}(F^TF) \cdot \|H\|^2\,. \qquad \blacksquare
\end{aligned}$$

Corollary 2.25
Let $P \in \mathrm{PSym}(3)$ *and* $F \in \mathbb{M}^{3\times 3}$. *Then* $\langle FP, P^{-1}F\rangle \geq \frac{\lambda_{\min}(P)}{\lambda_{\max}(P)} \cdot \|F\|^2$.

Proof. Without loss of generality we may assume that $P = D = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$. Then

$$\begin{aligned}
\langle FD, D^{-1}F\rangle &= \langle D^{-1}FD, F\rangle\,, \\
D^{-1}FD &= \begin{pmatrix} \frac{1}{\lambda_1} & 0 & 0 \\ 0 & \frac{1}{\lambda_2} & 0 \\ 0 & 0 & \frac{1}{\lambda_3} \end{pmatrix} \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}
\end{aligned}$$

$$= \begin{pmatrix} \frac{\lambda_1}{\lambda_1}F_{11} & \frac{\lambda_2}{\lambda_1}F_{12} & \frac{\lambda_3}{\lambda_1}F_{13} \\ \frac{\lambda_1}{\lambda_2}F_{21} & \frac{\lambda_2}{\lambda_2}F_{22} & \frac{\lambda_3}{\lambda_2}F_{23} \\ \frac{\lambda_1}{\lambda_3}F_{31} & \frac{\lambda_2}{\lambda_3}F_{32} & \frac{\lambda_3}{\lambda_3}F_{33} \end{pmatrix},$$

$$\langle D^{-1}FD, F\rangle = \sum_{i,j=1}^{3} \frac{\lambda_j}{\lambda_i} F_{ij}^2 \geq \frac{\lambda_{\min}}{\lambda_{\max}} \sum_{i,j=1}^{3} F_{ij}^2 = \frac{\lambda_{\min}}{\lambda_{\max}} \|F\|^2 . \qquad \blacksquare$$

A good source for any kind of matrix analysis is Horn/Johnson [**?**, **?**].

2.4 Properties of the tensor- or dyadic product

Let $\eta, \xi \in \mathbb{R}^3$, then $\eta \otimes \xi \in \mathbb{M}^{3\times 3}$ and $(\eta \otimes \xi)_{ij} = \eta_i\, \xi_j$. This yields the following

Lemma 2.26 (Basic properties of the tensor product)
Let $A \in \mathbb{M}^{3\times 3}$, $v \in \mathbb{R}^3$ *and* $\eta \otimes \xi \in \mathbb{M}^{3\times 3}$ *then*

1. $(\eta \otimes \xi).v = \eta\, \langle \xi, v\rangle$.
2. $(\eta \otimes \xi)^T = \xi \otimes \eta$.
3. $\operatorname{tr}[\eta \otimes \xi] = \langle \eta, \xi\rangle$.
4. $\operatorname{tr}[\eta \otimes \eta] = \|\eta\|^2$.
5. $\|\eta \otimes \xi\|^2 = \|\eta\|^2 \|\xi\|^2$.
6. $\langle \eta \otimes \xi, (\eta \otimes \xi)^T\rangle = \langle \eta \otimes \xi, (\xi \otimes \eta)\rangle = \langle \eta, \xi\rangle^2 \geq 0$.
7. $\operatorname{tr}[(\eta \otimes \xi)^2)] = (\operatorname{tr}[\eta \otimes \xi])^2$.
8. $\|(\eta \otimes \xi)^T + (\eta \otimes \xi)\|^2 \geq 2\|\eta\|^2\|\xi\|^2$.
9. $A \cdot (\eta \otimes \xi) = A.\eta \otimes \xi$.
10. $(\eta \otimes \xi) \cdot A = \eta \otimes A^T.\xi$.
11. $A \cdot (\eta \otimes \xi) \cdot A^T = A.\eta \otimes A.\xi$.
12. *rank*$(\eta \otimes \xi) = 1$.
13. *for every matrix* $A \in \mathbb{M}^{3\times 3}$ *with rank*$(A) = 1$ *there exists vectors* $\eta, \xi \in \mathbb{R}^3$ *such that* $A = \eta \otimes \xi$.
14. $B = 1\!\!1 + \eta \otimes \xi \Longrightarrow B^{-1} = 1\!\!1 - \frac{1}{1+\langle \eta, \xi\rangle} \cdot \eta \otimes \xi$ *if* $\langle \eta, \xi\rangle \neq 1$.

Proof.

1. trivial.
2. trivial.

3. $\operatorname{tr}[\eta \otimes \xi] = \langle \eta \otimes \xi, \mathbb{1} \rangle = \sum_{i,j=1}^{3} \eta_i \cdot \xi_j \cdot \delta_{ij} = \sum_i \eta_i \cdot \xi_i = \langle \eta, \xi \rangle$.
4. with the previous line clear.
5. $\|\eta \otimes \xi\|^2 = \sum_i \sum_j (\eta_i\, \xi_j)^2 = \sum_i \sum_j \eta_i^2 \xi_j^2 = (\sum_i \eta_i^2) \cdot \left(\sum_j \xi_j^2 \right) = \|\eta\|^2 \|\xi\|^2$.
6. $\langle \eta \otimes \xi, (\xi \otimes \eta) \rangle = \sum_{i,j} (\eta \otimes \xi)_{ij} (\xi \otimes \eta)_{ij} = \sum_{i,j} \eta_i\, \xi_j \quad \xi_i\, \eta_j =$
   $\sum_{i,j} (\eta_i\, \xi_i) \quad (\xi_j\, \eta_j) = \sum_i (\eta_i\, \xi_i) \quad \sum_j (\xi_j\, \eta_j) = \left(\sum_i (\eta_i\, \xi_i)\right)^2 = \langle \eta, \xi \rangle^2 \geq 0$.
7. $\operatorname{tr}[(\eta \otimes \xi)^2] = \langle (\eta \otimes \xi)^2, \mathbb{1} \rangle = \langle \eta, \xi \rangle^2 = (\operatorname{tr}[\eta \otimes \xi])^2$.
8. $\|(\eta \otimes \xi)^T + (\eta \otimes \xi)\|^2 = 2\|\eta \otimes \xi\|^2 + 2\langle \eta \otimes \xi, (\xi \otimes \eta) \rangle \geq 2\|\eta\|^2 \|\xi\|^2$.
9. $A \cdot (\eta \otimes \xi).v = A.\eta \ \langle \xi, v \rangle = (A.\eta \otimes \xi).v$.
10. $(\eta \otimes \xi) \cdot A.v = \eta \ \langle \xi, A.v \rangle = \eta \ \langle A^T.\xi, v \rangle = (\eta \otimes A^T.\xi).v$.
11. then o.k.
12. $\eta \in \mathbb{R}^3$ is a basis of the image of $\eta \otimes \xi$.
13. Let $rank(A) = 1$, then there exists $\xi \in \mathbb{R}^3 : \ A.\xi \neq 0$ ($\xi \in image(A^T)$). We set $\eta = \frac{A.\xi}{\|\xi\|^2}$. Then $A = \eta \otimes \xi$ since $(\eta \otimes \xi).\xi = A.\xi$ and let $\zeta \in kern(A)$, therefore $(\eta \otimes \xi).\zeta = \eta \cdot \langle \xi, \zeta \rangle = \eta \cdot \langle A^T.\mu, \zeta \rangle = \eta \cdot \langle \mu, A.\zeta \rangle = 0$.
14. direct computation. ■

Observe that $\langle F^T F H^T H, \xi \otimes \xi \rangle$ is not necessarily positive, since taking e.g.,

$$F^T F = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}, \quad H^T H = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad \xi = \begin{pmatrix} x \\ 1 \end{pmatrix} \ \Rightarrow$$
$$\langle F^T F H^T H, \xi \otimes \xi \rangle = \langle F^T F.\xi, H^T H.\xi \rangle = x^2 + (1 + \lambda)\, x + 2\lambda < 0$$

for some x and some positive λ. Note, however, that if $F^T F$ and $H^T H$ can be diagonalized simultaneously, then the positivity of $\langle F^T F H^T H, \xi \otimes \xi \rangle$ holds true.

2.5 The adjugate and determinant

Definition 2.27 (Projector $P : \mathbb{M}^{3\times 3} \mapsto \mathbb{M}^{2\times 2}$)
Let $A \in \mathbb{M}^{3\times 3}$. *Then*

$$P_{ij}.A := \text{the matrix which remains if row } j \text{ and column } i \text{ are cancelled.} \tag{2.2}$$

Remark 2.28
P is linear and $(\operatorname{Cof} A)_{ij} = (-1)^{i+j} \det[P_{ji}.A]$. ■

Definition 2.29 (Adjugate of a matrix)
Let $A \in \mathbb{M}^{3\times 3}$. We define the adjugate $\operatorname{Adj} A$ *in the following way:*

$$\begin{aligned}(\operatorname{Adj} A)_{ij} &:= (-1)^{i+j}\cdot \det[(P_{ij}.A)]\\ &\textit{if } A \textit{ is invertible, then}\\ \operatorname{Adj} A &= \det[A]\cdot A^{-1}.\end{aligned} \tag{2.3}$$

The expression $\det[(P_{ij}.A)]$ *is called the minor.*

For ease of reference we supply also the component form:

$$\begin{aligned}\operatorname{Adj} H &= \operatorname{Adj}\begin{pmatrix} H_{11} & H_{12} & H_{13}\\ H_{21} & H_{22} & H_{23}\\ H_{31} & H_{32} & H_{33}\end{pmatrix}\\ &= \begin{pmatrix} \begin{vmatrix} H_{22} & H_{23}\\ H_{32} & H_{33}\end{vmatrix} & -\begin{vmatrix} H_{21} & H_{23}\\ H_{31} & H_{33}\end{vmatrix} & \begin{vmatrix} H_{21} & H_{22}\\ H_{31} & H_{32}\end{vmatrix}\\ -\begin{vmatrix} H_{12} & H_{13}\\ H_{32} & H_{33}\end{vmatrix} & \begin{vmatrix} H_{11} & H_{13}\\ H_{31} & H_{33}\end{vmatrix} & -\begin{vmatrix} H_{11} & H_{12}\\ H_{31} & H_{32}\end{vmatrix}\\ \begin{vmatrix} H_{12} & H_{13}\\ H_{22} & H_{23}\end{vmatrix} & -\begin{vmatrix} H_{11} & H_{13}\\ H_{21} & H_{23}\end{vmatrix} & \begin{vmatrix} H_{11} & H_{12}\\ H_{21} & H_{22}\end{vmatrix}\end{pmatrix}\\ &= \begin{pmatrix} H_{22}H_{33}-H_{32}H_{23} & H_{31}H_{23}-H_{21}H_{33} & H_{21}H_{32}-H_{31}H_{22}\\ H_{32}H_{13}-H_{12}H_{33} & H_{11}H_{33}-H_{31}H_{13} & H_{31}H_{12}-H_{11}H_{32}\\ H_{12}H_{23}-H_{22}H_{13} & H_{21}H_{13}-H_{11}H_{23} & H_{11}H_{22}-H_{21}H_{12}\end{pmatrix}.\end{aligned}$$

For $H \in \mathbb{M}^{2\times 2}$ we have

$$\begin{aligned}\det[(1\!\!1 + H)] &= \begin{vmatrix} 1+H_{11} & H_{12}\\ H_{21} & 1+H_{22}\end{vmatrix}\\ &= 1 + H_{11} + H_{22} + H_{11}H_{22} - H_{21}H_{12} = 1 + \operatorname{tr}(H) + \det[H].\end{aligned}$$

Hence for invertible $F \in \mathbb{M}^{2\times 2}$ we get

$$\begin{aligned}\det[(F+H)] &= \det[((1\!\!1 + HF^{-1})F)] = \det[(1\!\!1 + HF^{-1})]\,\det[F]\\ &= \det[F]\,\left(1 + \operatorname{tr}(HF^{-1}) + \det[(HF^{-1})]\right)\\ &= \det[F] + \det[F]\,\operatorname{tr}(HF^{-1}) + \det[H]\\ &= \det[F] + \operatorname{tr}(H\,\operatorname{Adj} F) + \det[H].\end{aligned}$$

This representation remains true for non invertible $F \in \mathbb{M}^{2\times 2}$ by density. Thus for $F, H \in \mathbb{M}^{3\times 3}$ we have in components

$$\begin{aligned}(\operatorname{Adj}(F+H))_{ij} &= (-1)^{i+j}\,\det[(P_{ij}.(F+H))]\\ &= (-1)^{i+j}\,\det[(P_{ij}.F + P_{ij}.H)]\\ &= (-1)^{i+j}\,\big(\det[(P_{ij}.F)] + \det[(P_{ij}.F)]\,\operatorname{tr}(P_{ij}.H\,(P_{ij}.F)^{-1}) + \det[(P_{ij}.H)]\big)\\ &= (-1)^{i+j}\det[(P_{ij}.F)] + (-1)^{i+j}\,\det[(P_{ij}.F)]\,\operatorname{tr}(P_{ij}.H\,(P_{ij}.F)^{-1}) +\\ &\quad (-1)^{i+j}\det[(P_{ij}.H)].\end{aligned}$$

This means

$$\operatorname{Adj}(F+H) = \operatorname{Adj} F + D(\operatorname{Adj} F).H + \operatorname{Adj} H.$$

where

$$\begin{aligned} D(\operatorname{Adj} F).H &= \operatorname{Adj} F \cdot \left\{\langle F^{-T}, H\rangle \mathbb{1} - HF^{-1}\right\} \\ D^2(\operatorname{Adj} F).(H,H) &= 2 \operatorname{Adj} H\,. \end{aligned}$$

2.6 Adjugate and determinant

Definition 2.30 (Adjugate $\operatorname{Adj} F = \operatorname{Cof}(F)^T$**)**

$$\begin{aligned} D(\operatorname{Adj} F).H &= \operatorname{Adj} F \cdot \left\{\langle F^{-T}, H\rangle \mathbb{1} - HF^{-1}\right\} \\ D^2(\operatorname{Adj} F).(H,H) &= 2 \operatorname{Adj} H\,, \end{aligned} \tag{2.4}$$

since $\operatorname{Adj} F$ is a quadratic expression in the entries of F.

Lemma 2.31 (More properties of the Adjugate)
Let $A, B, P \in \mathrm{GL}(3,\mathbb{R})$ and $Q \in O(3)$. Then we have:

1. $\operatorname{Adj} A = \operatorname{Cof}(A)^T$.
2. $\operatorname{Adj}(\xi \otimes \eta) = 0$.
3. $\operatorname{Adj}(\mathbb{1} + \xi \otimes \eta) = \mathbb{1} + \langle \xi, \eta\rangle \mathbb{1} - \xi \otimes \eta$.
4. $\operatorname{Adj}(AB) = \operatorname{Adj} B\ \operatorname{Adj} A$.
5. $\operatorname{Adj} A^T = (\operatorname{Adj} A)^T$.
6. $\operatorname{Adj} A \cdot A = \det[A] \cdot \mathbb{1}$.
7. $\operatorname{Adj}(P^{-1}AP) = P^{-1} \operatorname{Adj} AP$, *hence* Adj *is an isotropic tensor function.*
8. $\operatorname{Adj}(A^{-1}) = (\operatorname{Adj} A)^{-1}$.
9. *Let D be a diagonal matrix, then* $\operatorname{Adj} D = \begin{pmatrix} \lambda_2\lambda_3 & 0 & 0 \\ 0 & \lambda_1\lambda_3 & 0 \\ 0 & 0 & \lambda_1\lambda_2 \end{pmatrix}$.
10. $\operatorname{tr}[\operatorname{Adj} D] = \lambda_2\lambda_3 + \lambda_1\lambda_3 + \lambda_1\lambda_2$.
11. $\operatorname{tr}[\operatorname{Adj}(P^{-1}AP)] = \operatorname{tr}[\operatorname{Adj} A]$.
12. $\|\operatorname{Adj}(Q^{-1}AQ)\|^2 = \|\operatorname{Adj} A\|^2$.
13. $\langle \operatorname{Adj} F^T F, \mathbb{1}\rangle = \|\operatorname{Adj} F\|^2$.

14. *For* $Q \in O(3): \ \operatorname{Adj}(QF) = (\operatorname{Adj} F)\, Q^T$ *and* $\|\operatorname{Adj}(QF)\| = \|\operatorname{Adj} F\|$.

Remark 2.32
Every noninvertible matrix can be approximated by invertible matrices. Therefore, by density, the above properties carry over to non-invertible matrices as well. ■

Proof.

1. obvious: definition.
2. obviously, $\xi \otimes \eta$ has rank one, hence there are no nonvanishing 2 by 2 minors.
3. we apply the formula for the derivative of the adjugate (**??**) to obtain $\operatorname{Adj}(\mathbb{1} + \xi \otimes \eta) = \operatorname{Adj} \mathbb{1} + D \operatorname{Adj} \mathbb{1}.\xi \otimes \eta + \operatorname{Adj} \xi \otimes \eta = \mathbb{1} + D \operatorname{Adj} \mathbb{1}.\xi \otimes \eta + 0 = \mathbb{1} + \mathbb{1}\,(\langle \mathbb{1}, \xi \otimes \eta \rangle \cdot \mathbb{1} - \xi \otimes \eta) = \mathbb{1} + \langle \xi, \eta \rangle \mathbb{1} - \xi \otimes \eta$.
4. obvious: definition.
5. obvious: definition.
6. obvious: definition.
7. $\operatorname{Adj}(P^{-1}AP) = \det[P^{-1}AP] \cdot (P^{-1}AP)^{-1} = \det[A] \cdot P^{-1}A^{-1}P$
   $= P^{-1}\det[A] \cdot A^{-1}P = P^{-1} \operatorname{Adj} AP$.
8. obvious: definition.
9. obvious: definition.
10. to be read off immediately.
11. $\|\operatorname{Adj}(Q^{-1}AQ)\|^2 = \|Q^{-1} \operatorname{Adj} AQ\|^2 = \|\operatorname{Adj} A\|^2$.
12. $\langle \operatorname{Adj} F^T F, \mathbb{1} \rangle = \langle \operatorname{Adj} F \ \operatorname{Adj} F^T, \mathbb{1} \rangle = \langle \operatorname{Adj} F\ (\operatorname{Adj} F)^T, \mathbb{1} \rangle = \langle \operatorname{Adj} F, \operatorname{Adj} F \rangle = \|\operatorname{Adj} F\|^2$.
13. obvious: definition and isotropy. ■

Theorem 2.33 (Cayley-Hamilton)
Let $A \in \mathbb{M}^{3\times3}$. *Then* A *is solution of its characteristic polynomial* $\det[(\lambda \cdot \mathbb{1} - A)] = 0$, *i.e.*

$$\begin{aligned} 0 &= \lambda^3 - \operatorname{tr}[A] \cdot \lambda^2 + \operatorname{tr}[\operatorname{Adj} A] \cdot \lambda - \det[A] \cdot \lambda^0 \\ &\textit{which means} \\ 0 &= A^3 - \operatorname{tr}[A] \cdot A^2 + \operatorname{tr}[\operatorname{Adj} A] \cdot A - \det[A] \cdot A^0 \\ 0 &= A^3 - \operatorname{tr}[A] \cdot A^2 + \operatorname{tr}[\operatorname{Adj} A] \cdot A - \det[A] \cdot \mathbb{1}. \end{aligned} \tag{2.5}$$

Proof. Standard exercise in classical linear algebra. ■

Lemma 2.34 (Invariants)
For all real diagonalizable $A \in \mathbb{M}^{3\times 3}$ *we set*

$$\begin{aligned} I_1(A) &:= \mathrm{tr}(A) = \lambda_1 + \lambda_2 + \lambda_3\,, \\ I_2(A) &:= \mathrm{tr}(\mathrm{Adj}\,A) = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3\,, \\ I_3(A) &:= \det[A] = \lambda_1\lambda_2\lambda_3\,. \end{aligned}$$

Because of Theorem **??** *this implies*

$$\begin{aligned} \mathrm{tr}(F)^2 &= \mathrm{tr}(F^2) + 2\ \mathrm{tr}(\mathrm{Adj}\,F) \\ (\lambda_1+\lambda_2+\lambda_3)^2 &= \lambda_1^2+\lambda_2^2+\lambda_3^2 + 2\,(\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_1\lambda_3)\,. \end{aligned}$$

■

Lemma 2.35 (Coefficients of the characteristic polynomial)
Let A *be real diagonalizable and assume that* $\det[A] \geq 0$*. Then we have*

$$\begin{aligned} I_1^2(A) &\geq 3 \cdot I_2(A) \\ I_2^2(A) &\geq 3 \cdot I_1(A)\ I_3(A). \end{aligned}$$

Proof. Young's inequality shows that $\lambda_i \cdot \lambda_j \leq \frac{1}{2}\lambda_i^2 + \frac{1}{2}\lambda_j^2$ (consider $(\lambda_i - \lambda_j)^2 \geq 0$). Therefore $\lambda_1^2+\lambda_2^2+\lambda_3^2 \geq \lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_1\lambda_3$. Hence

$$\begin{aligned} (\lambda_1+\lambda_2+\lambda_3)^2 &= \left(\lambda_1^2+\lambda_2^2+\lambda_3^2\right) + 2\,(\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_1\lambda_3) \\ &\geq 3\,(\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_1\lambda_3)\ , \end{aligned}$$

which proves $I_1(A)^2 \geq 3 \cdot I_2(A)$. In order to prove the second statement note that we may assume $\lambda_i(A) \neq 0$ without loss of generality since otherwise the statement is true anyway. Let therefore $\det[A] \neq 0$. Then the inverse $A^{-1} \in \mathbb{M}^{3\times 3}$ exists and with the first statement we know $I_1(A^{-1})^2 \geq 3 \cdot I_2(A^{-1})$. Moreover $\hat{\lambda}(A^{-1})_i = \frac{1}{\lambda(A)_i}$. Therefore

$$\begin{aligned} \left(\frac{1}{\lambda_1}+\frac{1}{\lambda_2}+\frac{1}{\lambda_3}\right)^2 &\geq 3\left(\frac{1}{\lambda_1\lambda_2}+\frac{1}{\lambda_2\lambda_3}+\frac{1}{\lambda_3\lambda_1}\right) \\ \left(\frac{\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_1\lambda_3}{\lambda_1\lambda_2\lambda_3}\right)^2 &\geq 3\left(\frac{\lambda_1+\lambda_2+\lambda_3}{\lambda_1\lambda_2\lambda_3}\right) \\ (\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_1\lambda_3)^2 &\geq 3(\lambda_1+\lambda_2+\lambda_3)\cdot(\lambda_1\lambda_2\lambda_3)\,, \end{aligned}$$

which is just $I_2(A)^2 \geq 3\,I_1(A)\cdot I_3(A)$. ■

Corollary 2.36 (Estimates between $\|F\|$, $\|\mathrm{Adj}\,F\|$ and $\det[F]$)
Let $F \in \mathbb{M}^{3\times 3}$*. Then we have*

$$\begin{aligned} \|F\|^3 &\geq 3\sqrt{3}\,\det[F]\,, \\ \|F\|^2 &\geq \sqrt{3}\,\|\mathrm{Adj}\,F\|\,, \\ \|\mathrm{Adj}\,F\|^3 &\geq 3\sqrt{3}\,(\det[F])^2\,, \\ \|F\|^2 = \langle F^TF, 1\!\!1\rangle &\leq \sqrt{3}\ \|F^TF\|\,. \end{aligned}$$

Proof. Set $A = F^TF$. The symmetry of A ensures the applicability of the foregoing Lemma **??**. Thus we have

$$\begin{aligned}
I_1(A) &= I_1(F^TF) = \operatorname{tr}(F^TF) = \|F\|^2\,,\\
I_2(A) &= I_2(F^TF) = \operatorname{tr}(\operatorname{Adj}(F^TF)) = \operatorname{tr}(\operatorname{Adj}F \cdot \operatorname{Adj}F^T) = \|\operatorname{Adj}F\|^2\,,\\
I_3(A) &= I_3(F^TF) = \det[F^TF] = (\det[F])^2\,,
\end{aligned}$$

and

$$\begin{aligned}
I_1^2(A) \geq& 3\,I_2(A) \Leftrightarrow \|F\|^2 \geq \sqrt{3}\|\operatorname{Adj}F\|\\
I_2^2(A) \geq& 3\,I_1(A)\cdot I_3(A) \Leftrightarrow \|\operatorname{Adj}F\|^2 \geq \sqrt{3}\|F\|\,\det[F].
\end{aligned}$$

The last two lines lead immediately to the second statement and third statement. The last statement is only a simple algebraic estimate. ■

Theorem 2.37 (Representation for linear mappings of $\mathbb{M}^{3\times3} \mapsto \mathbb{R}$.)
For every linear mapping $L : \mathbb{M}^{3\times3} \mapsto \mathbb{R}$ there exists a uniquely determined matrix $M_L \in \mathbb{M}^{3\times3}$ such that

$$\forall X \in \mathbb{M}^{3\times3}: \quad L.X := L(X) = \langle M_L, X\rangle\,.$$

Proof. The matrix M can be constructed easily. ■

Remark 2.38
This statement allows us to identify $\mathbb{M}^{3\times3}$ and $L(\mathbb{M}^{3\times3}, \mathbb{R})$.

2.7 Formal first derivatives

By direct expansion of tensorvalued functions we are able to identify termwise the appearing derivatives having in mind the corresponding Taylor-series representation. The simple yet efficient method is illustrated by examples. Let us begin with recalling the Neumann-series for matrices:

$$(1\!\!1 - F)^{-1} = 1\!\!1 + F + F^2 + F^3 + \dots$$

This series converges whenever $\|F\| < 1$.

$$(F+H)^T = F^T + H^T \Rightarrow D(F^T).H = H^T\,.$$

$$\begin{aligned}
(F+H)^2 = F^2 + FH + HF + H^2 &\Rightarrow D(F^2).H = FH + HF\\
&\Rightarrow D^2(F^2).(H,H) = 2\cdot H^2\,.
\end{aligned}$$

If we differentiate the first derivative again with fixed argument, i.e.,

$$D(D(F^2).H).\hat{H} = D(FH + HF).\hat{H} = \hat{H}H + H\hat{H}\,,$$

then one determines not only the second derivative in a fixed direction H (the quadratic form) but the complete bilinearform. This has to be noted in the subsequent calculations.

$$\begin{aligned}(F+H)^3 &= F^3 + F^2H + FHF + HF^2 + FH^2 + HFH + H^2F + H^3\\ &\Rightarrow D(F^3).H = F^2H + FHF + HF^2\\ &\Rightarrow D^2(F^3).(H,H) = 2\{FH^2 + HFH + H^2F\}\\ &\Rightarrow D^3(F^3).(H,H,H) = 6\cdot H^3\end{aligned}$$

$$\begin{aligned}(F+H)^T(F+H) &= F^TF + F^TH + H^TF + H^TH\\ &\Rightarrow D(F^TF).H = F^TH + H^TF\\ &\Rightarrow D^2(F^TF).(H,H) = 2\cdot H^TH\,.\end{aligned}$$

$$\begin{aligned}(F+H)^{-1} &= (F(\mathbb{1}+F^{-1}H))^{-1} = (\mathbb{1}+F^{-1}H)^{-1}F^{-1}\\ &= (\mathbb{1}-(-F^{-1}H))^{-1}F^{-1}\\ &\text{using Neumanns series}\\ &= (\mathbb{1}-F^{-1}H+(F^{-1}H)^2+\ldots)F^{-1}\\ &= F^{-1}-F^{-1}HF^{-1}+(F^{-1}H)^2F^{-1}+\ldots\\ &\Rightarrow D(F^{-1}).H = -F^{-1}HF^{-1}\\ &\Rightarrow D^2(F^{-1}).(H,H) = (F^{-1}H)^2F^{-1} = F^{-1}HF^{-1}HF^{-1}\,.\end{aligned}$$

$$\begin{aligned}((F+H)^T(F+H))^{-1} &= (F^TF+F^TH+H^TF+H^TH)^{-1}\\ &= \big((F^TF)(\mathbb{1}+(F^TF)^{-1}(F^TH+H^TF+H^TH))\big)^{-1}\\ &= \big((\mathbb{1}+(F^TF)^{-1}(F^TH+H^TF+H^TH))\big)^{-1}\cdot(F^TF)^{-1}\\ &= \big(\mathbb{1}-(-(F^TF)^{-1}(F^TH+H^TF+H^TH))\big)^{-1}\cdot(F^TF)^{-1}\\ &\text{expanding with Neumanns series}\\ &=(\mathbb{1}-(F^TF)^{-1}(F^TH+H^TF+H^TH)+\\ &\qquad\big[(F^TF)^{-1}(F^TH+H^TF+H^TH)\big]^2+\ldots)\cdot(F^TF)^{-1}\\ &= (F^TF)^{-1}-(F^TF)^{-1}(F^TH+H^TF+H^TH))(F^TF)^{-1}+\\ &\qquad\big[(F^TF)^{-1}(F^TH+H^TF+H^TH)\big]\cdot\\ &\qquad\quad\big[(F^TF)^{-1}(F^TH+H^TF+H^TH)\big]\cdot(F^TF)^{-1}+\ldots\\ &\text{keeping terms up to second order}\\ &=(F^TF)^{-1}-(F^TF)^{-1}(F^TH+H^TF+H^TH))(F^TF)^{-1}+\\ &(F^TF)^{-1}(F^TH+H^TF)\cdot(F^TF)^{-1}(F^TH+H^TF)\cdot(F^TF)^{-1}+\ldots\\ &\Rightarrow D((F^TF)^{-1}).H = -(F^TF)^{-1}(F^TH+H^TF))(F^TF)^{-1}\\ &\Rightarrow D^2((F^TF)^{-1}).(H,H) = -(F^TF)^{-1}(H^TH))(F^TF)^{-1}+\\ &(F^TF)^{-1}(F^TH+H^TF)\cdot(F^TF)^{-1}(F^TH+H^TF)\cdot(F^TF)^{-1}\,.\end{aligned}$$

$$\begin{aligned}\|F+H\|^2 &= \langle F+H, F+H\rangle\\ &= \langle F,F\rangle + \langle F,H\rangle + \langle H,F\rangle + \langle H,H\rangle\\ &\Rightarrow D(\|F\|^2).H = 2\langle F,H\rangle\,.\end{aligned}$$

$$\begin{aligned}\|F+H\| &= \sqrt{\langle F+H, F+H\rangle} = \sqrt{\|F\|^2 + 2\langle F,H\rangle + \|H\|^2}\\ &= \|F\|\{\sqrt{1+\frac{2}{\|F\|^2}\langle F,H\rangle + \frac{\|H\|^2}{\|F\|^2}}\}\\ &= \|F\|\{1+\frac{1}{\|F\|^2}\langle F,H\rangle + \frac{1}{2}\frac{\|H\|^2}{\|F\|^2} + \ldots\}\\ &= \|F\| + \frac{1}{\|F\|}\langle F,H\rangle + \ldots \quad , F \neq 0\,,\end{aligned}$$

where we have used the expansion $\sqrt{1+x} = 1+\frac{1}{2}x+\ldots$ for $x \in \mathbb{R}$. Therefore for matrices as well as for real numbers we have:

$$D(\|F\|).H = \frac{1}{\|F\|}\langle F,H\rangle, \quad F \neq 0\,.$$

Using the chain-rule we obtain thus for $p \geq 1$:

$$D.(\|F\|^p).H = p\|F\|^{p-1}\frac{1}{\|F\|}\langle F,H\rangle = p\|F\|^{p-2}\langle F,H\rangle\,.$$

The **product-rule:** obtains in the usual format

$$D(G(F)\cdot J(F)).H = (DG(F).H)\cdot J(F) + G(F)\cdot(DJ(F).H)\,.$$

Differentiation with the help of the inverse function and the **chain-rule**:

$$G(J(F)) = F \Rightarrow DG(J(F)).DJ(F).H = \mathrm{I}d.H\,,$$

and if this can be inverted it follows:

$$DJ(F).H = (DG(J(F)))^{-1}.H\,.$$

If G is continuously differentiable, then also J is differentiable.
Let us demonstrate this with an example. Let $A, B \in \mathrm{PSym}(3)$ and $G(A) = A^2$ and $J(B) = B^{\frac{1}{2}}$. Then we have

$$G(J(B)) = B \Rightarrow DG(J(B)) \bullet DJ(B).H = H$$

and with $DG(A).H = AH + HA$ it follows

$$\begin{aligned}&J(B)\ DJ(B).H + DJ(B).H\ J(B) = H\\ &\qquad\Leftrightarrow B^{\frac{1}{2}}\ DJ(B).H + DJ(B).H\ B^{\frac{1}{2}} = H\,.\end{aligned}$$

This is a matrix equation for $DJ(B).H$ (**Lyapunov matrixsystem**), which, for $H \in$ PSym admits a unique solution. The precise expression is to difficult for our purposes and we do not consider here functions of the type $F \mapsto \Psi((F^TF)^{\frac{1}{2}})$. Such functions, however, appear naturally for scalarvalued functions defined on the eigenvalues of $(F^TF)^{\frac{1}{2}}$.

$$D(\|F^TF\|^2).H = 2\|F^TF\|\langle \frac{F^TF}{\|F^TF\|}, F^TH + H^TF\rangle = 2\langle F^TF, F^TH + H^TF\rangle\,.$$

For $p \geq 2$ it holds:

$$\begin{aligned} D(\frac{1}{p}\|F^TF\|^p).H &= \|F^TF\|^{p-1}\langle \frac{F^TF}{\|F^TF\|}, F^TH + H^TF\rangle \\ &= \|F^TF\|^{p-2}\langle F^TF, F^TH + H^TF\rangle\,. \end{aligned}$$

$$\begin{aligned} \operatorname{tr}(F+H) = \langle F+H, 1\!\!1\rangle &= \langle F, 1\!\!1\rangle + \langle H, 1\!\!1\rangle = \operatorname{tr}(F) + \operatorname{tr}(H) \\ &\Rightarrow D(\operatorname{tr}(F)).H = \operatorname{tr}(H)\,. \end{aligned}$$

The formal expansion of the determinant shows for dimension n:

$$\det[(1\!\!1 + H)] = \begin{cases} 1 + \det[H] & \text{if } n = 1 \\ 1 + \operatorname{tr}(H) + \det[H] & \text{if } n = 2 \\ 1 + \operatorname{tr}(H) + \operatorname{tr}(\operatorname{Adj} H) + \det[H] & \text{if } n = 3 \\ 1 + \operatorname{tr}(H) + \cdots + \operatorname{tr}(\operatorname{Adj} H) + \det[H] & \text{if } n > 3\,. \end{cases}$$

From now onwards we restrict attention to dimension $n = 3$:

$$\det[F] = \det[(1\!\!1 + (F - 1\!\!1))] \qquad = 1 + \operatorname{tr}(F - 1\!\!1) + \operatorname{tr}(\operatorname{Adj} F - 1\!\!1) + \det[(F - 1\!\!1)]$$

This shows that for $F - 1\!\!1 = \nabla u$

$$\det[(]1\!\!1 + \nabla u) = 1 + \operatorname{tr}(\nabla u) + \ldots = 1 + \operatorname{Div} u + \ldots$$

Lemma 2.39 (Expansion of the determinant for $n = 3$)
Let $\det[F] > 0$, *then*

$$\det[F + H] = \det[F] + \langle H, \operatorname{Adj} F^T\rangle + \langle \operatorname{Adj} H, F^T\rangle + \det[H]\,.$$

Proof.

$$\begin{aligned} \det[F+H] &= \det((1\!\!1 + HF^{-1})F) = \det[(1\!\!1 + HF^{-1})]\det[F] \\ &= \det[F](\det[1\!\!1] + \operatorname{tr}(HF^{-1}) + \operatorname{tr}(\operatorname{Adj}(HF^{-1})) + \det[HF^{-1}]) \\ &= \det[F] + \det[F]\langle HF^{-1}, 1\!\!1\rangle + \det[F]\langle \det[(HF^{-1})](HF^{-1})^{-1}, 1\!\!1\rangle \\ &\qquad + \det[F]\det[H]\det[F^{-1}] \\ &= \det[F] + \det[F]\langle HF^{-1}, 1\!\!1\rangle + \det[F]\langle \det[(HF^{-1})]FH^{-1}, 1\!\!1\rangle + \\ &\qquad \det[F]\det[H]\det[F^{-1}] \\ &= \det[F] + \det[F]\langle HF^{-1}, 1\!\!1\rangle + \det[F]\langle \det[(HF^{-1})]FH^{-1}, 1\!\!1\rangle + \\ &\qquad \det[H] \\ &= \det[F] + \langle H, \operatorname{Adj} F^T\rangle + \langle \operatorname{Adj} H, F^T\rangle + \det[H]\,. \end{aligned}$$

By continuity this result holds as well for non-invertible matrices $F, H \in \mathbb{M}^{3\times 3}$. ■

On $\mathbb{M}^{3\times 3}$ the operator $\operatorname{Adj} : \mathbb{M}^{3\times 3} \mapsto \mathbb{M}^{3\times 3}, \quad F \mapsto \operatorname{Adj} F$ is a polynom of second grade in the entries of F, therefore, the Taylor expansion must end with the second term.

Lemma 2.40 (Expansion of the adjugate for $n = 3$)
It holds

$$\operatorname{Adj}(F+H) = \operatorname{Adj} F + D(\operatorname{Adj} F).H + \operatorname{Adj} H\,.$$

Proof.

$$\begin{aligned} \operatorname{Adj}(F+H) &= \operatorname{Adj} F + D(\operatorname{Adj} F).H + \operatorname{Adj} H \\ &\Rightarrow D^2(\operatorname{Adj} F).(H,H) = 2 \cdot \operatorname{Adj} H\,. \end{aligned}$$ ■

Let $G(F) = F^{-1}$, then

$$D(F^{-1}).H = -F^{-1}HF^{-1}\,.$$

$$D(\det[F]).H = \det[F] \cdot \operatorname{tr}(HF^{-1}) = \det[F] \cdot \langle F^{-T}, H\rangle = \langle (\operatorname{Adj} F)^T, H\rangle\,.$$

$$D(\ln(\det[F])).H = \frac{1}{\det[F]} D(\det[F]).H = \frac{1}{\det[F]} \det[F]\langle F^{-T}, H\rangle = \langle F^{-T}, H\rangle\,.$$

$$\begin{aligned} D((\ln\det[F])^2)).H &= 2\ln\det[F] \cdot \frac{1}{\det[F]} D(\det[F]).H \\ &= 2\ln\det[F] \cdot \frac{1}{\det[F]} \det[F]\langle F^{-T}, H\rangle = 2\ln\det[F] \cdot \langle F^{-T}, H\rangle\,. \end{aligned}$$

$$\begin{aligned} D(\operatorname{Adj} F).H &= D(\det[F]F^{-1}).H \\ &= (D(\det[F]).H)F^{-1} + (\det[F])D(F^{-1}).H \\ &= \det[F]\langle F^{-T}, H\rangle \cdot F^{-1} + \det[F](-F^{-1}HF^{-1}) \\ &= \operatorname{Adj} F\langle F^{-T}, H\rangle - (\operatorname{Adj} F)HF^{-1} \\ &= \operatorname{Adj} F\{\langle F^{-T}, H\rangle \cdot 1\!\!1 - HF^{-1}\}\,. \end{aligned}$$

For this identity we might have as well differentiated the identity $\operatorname{Adj} F \cdot F = \det[F] \cdot 1\!\!1$ with the same result.

$$\begin{aligned} D(\operatorname{Adj}(F^TF)).H &= D(\operatorname{Adj} F \operatorname{Adj} F^T).H \\ &= (D(\operatorname{Adj} F).H)\operatorname{Adj} F^T + (\operatorname{Adj} F)D(\operatorname{Adj} F^T).H \\ &= \operatorname{Adj} F\{\langle F^{-T}, H\rangle 1\!\!1 - HF^{-1}\}\operatorname{Adj} F^T + \\ &\qquad (\operatorname{Adj} F)\{\operatorname{Adj} F\{\langle F^{-T}, H\rangle 1\!\!1 - HF^{-1}\}\}^T\,. \end{aligned}$$

$$
\begin{aligned}
D(\frac{1}{\det[F]}).H &= D(\det[F^{-1}]).H = D((\det[F])^{-1}).H\\
&= (-1)(\det[F])^{-2}D(\det[F]).H\\
&= (-1)(\det[F])^{-2}\det[F]\langle F^{-T},H\rangle = (-1)(\det[F])^{-1}\langle F^{-T},H\rangle\,.
\end{aligned}
$$

$$
\begin{aligned}
D(\operatorname{Adj}(F^{-1})).H &= D(\det[F^{-1}]F).H\\
&= D((\det[F])^{-1}F).H\\
&= \{(-1)(\det[F])^{-2}\det[F]\langle F^{-T},H\rangle\}F + (\det[F])^{-1}H\\
&= (-1)(\det[F])^{-1}F\langle F^{-T},H\rangle + (\det[F])^{-1}H\\
&= -\operatorname{Adj}F^{-1}\langle F^{-T},H\rangle + (\det[F])^{-1}H\,.
\end{aligned}
$$

$$
\begin{aligned}
D(F^{-T}).H &= D((F^T)^{-1}).H = -F^{-T}(D(F^T).H)F^{-T}\\
&= -F^{-T}H^TF^{-T} = \{D(F^{-1}).H\}^T\,.
\end{aligned}
$$

$$
\begin{aligned}
\{G(F+H)\}^T &= \{G(F)+DG(F).H+\ldots\}^T = G(F)^T+\{DG(F).H\}^T+\ldots\\
&\Rightarrow D(G(F)^T).H = \{DG(F).H\}^T\,.
\end{aligned}
$$

$$
\begin{aligned}
D(\operatorname{tr}(\operatorname{Adj}F)).H &= \operatorname{tr}(D(\operatorname{Adj}F).H)\\
&= \operatorname{tr}(\operatorname{Adj}F\{\langle F^{-T},H\rangle 1\!\!1 - HF^{-1}\})\\
&= \operatorname{tr}(\operatorname{Adj}F\{\langle 1\!\!1,F^{-1}H\rangle 1\!\!1 - HF^{-1}\})\\
&= \operatorname{tr}(\operatorname{Adj}F)\operatorname{tr}(F^{-1}H) - \operatorname{tr}((\operatorname{Adj}F)HF^{-1})\,.
\end{aligned}
$$

$$
D(\|F\|^2).H = D(\langle F,F\rangle).H = 2\langle F,H\rangle\,.
$$

$$
\begin{aligned}
D(\|F\|).H &= D(\langle F,F\rangle^{\frac{1}{2}}).H\\
&= \frac{1}{2}\langle F,F\rangle^{-\frac{1}{2}}D(\langle F,F\rangle).H = \langle\frac{F}{\|F\|},H\rangle\,.
\end{aligned}
$$

$$
D(\frac{F}{\|F\|}).H = \frac{H}{\|F\|} - \frac{1}{\|F\|^3}\langle F,H\rangle\cdot F\,.
$$

$$
D(F^TF).H = F^TH + H^TF\,.
$$

$$
\begin{aligned}
D(\Psi(F^TF)).H &= \partial_{F^TF}\Psi(F^TF).(F^TH+H^TF)\\
&= \langle\partial_{F^TF}\Psi(F^TF),(F^TH+H^TF)\rangle\\
&= \langle F\cdot\partial_{F^TF}\Psi(F^TF),H\rangle + \langle\partial_{F^TF}\Psi(F^TF)\cdot F^T,H^T\rangle\\
&= \langle F\cdot\left(\partial_{F^TF}\Psi(F^TF)+\partial_{F^TF}\Psi(F^TF)^T\right),H\rangle\,.
\end{aligned}
$$

$$D(\|F-1\!\!1\|^2).H = 2\langle F-1\!\!1, H\rangle\,.$$

$$\begin{aligned} D(\|F^TF-1\!\!1\|^2).H &= 2\|F^TF-1\!\!1\|\langle\frac{F^TF-1\!\!1}{\|F^TF-1\!\!1\|}, F^TH+H^TF\rangle \\ &= 2\langle F^TF-1\!\!1, F^TH+H^TF\rangle\,. \end{aligned}$$

$$\begin{aligned} D((\mathrm{tr}(F^TF-1\!\!1))^2).H &= D(\langle F^TF-1\!\!1, 1\!\!1\rangle^2).H \\ &= 2\langle F^TF-1\!\!1, 1\!\!1\rangle\langle F^TH+H^TF, 1\!\!1\rangle\,. \end{aligned}$$

$$\begin{aligned} D(\mathrm{tr}(\mathrm{Adj}\,(F^TF)).H &= D(\langle \mathrm{Adj}\,F^TF, 1\!\!1\rangle).H = D(\langle \mathrm{Adj}\,F\,\mathrm{Adj}\,F^T, 1\!\!1\rangle).H \\ &= D(\langle \mathrm{Adj}\,F, \mathrm{Adj}\,F\rangle).H\,. \end{aligned}$$

$$\begin{aligned} D(\|\,\mathrm{Adj}\,F\|^2).H &= 2\langle \mathrm{Adj}\,F, D(\mathrm{Adj}\,F).H\rangle \\ &= 2\langle \mathrm{Adj}\,F, \mathrm{Adj}\,F\{\langle F^{-T}, H\rangle 1\!\!1 - HF^{-1}\}\rangle \\ &= 2\langle F^{-T}, H\rangle\|\,\mathrm{Adj}\,F\|^2 - 2\langle \mathrm{Adj}\,F, (\mathrm{Adj}\,F)HF^{-1}\rangle\,. \end{aligned}$$

$$\begin{aligned} D(\det[F^TF]).H &= D(\det[F^T]\det[F]).H = D((\det[F])^2).H \\ &= 2\{D(\det[F]).H\}\det[F] \\ &= 2(\det[F])^2\langle F^{-T}, H\rangle = 2\det[F]\langle(\mathrm{Adj}\,F)^T, H\rangle\,. \end{aligned}$$

$$\begin{aligned} D(\frac{1}{\det[F^TF]}).H &= D(\frac{1}{\det[F^T]\det[F]}).H = D(\frac{1}{(\det[F])^2}).H \\ &= -2\,(\det[F])^{-3}\langle\det[F]F^{-T}, H\rangle = -2\,(\det[F])^{-2}\langle F^{-T}, H\rangle\,. \end{aligned}$$

$$\begin{aligned} D(\frac{F}{\det[F]^{\frac{1}{3}}}).H &= D(F\,\det[F]^{-\frac{1}{3}}).H \\ &= H\cdot\det[F]^{-\frac{1}{3}} - F\cdot\frac{1}{3}(\det[F])^{-\frac{4}{3}}\cdot\det[F]\;\langle F^{-T}, H\rangle \\ &= H\cdot\det[F]^{-\frac{1}{3}} - \frac{1}{3}\frac{F}{\det[F]^{\frac{1}{3}}}\langle F^{-T}, H\rangle \\ &\text{for } F = 1\!\!1 \text{ it follows} \\ &= H - \frac{1}{3}\langle 1\!\!1, H\rangle = \mathrm{dev}(H)\,. \end{aligned}$$

2.8 Derivative for symmetric argument

Lemma 2.41 (Derivative for symmetric argument)
Let $W : \mathbb{M}^{3\times3} \mapsto \mathbb{R}$ *be differentiable and let* $\hat{W}$ *be the restriction of* W *on symmetric matrices, i.e.,*

$$\forall\, C \in \mathrm{Sym}(3): \quad \hat{W}(C) = W(C)\,.$$

Then

$$D_C\hat{W}(C) = \frac{1}{2}\left(DW(C) + DW(C)^T\right)\,.$$

It follows that the derivative of a function defined on a symmetric argument can always be assumed to be symmetric.

Proof. Let $H \in \mathrm{Sym}(3)$

$$\begin{aligned}\hat{W}(C+H) &= W(C+H) = W(C) + \langle DW(C), H\rangle + \dots\\ &\text{and since } H \in \mathrm{Sym}(3) \text{ it follows}\\ &= \hat{W}(C) + \langle \frac{1}{2}\left(DW(C) + DW(C)^T\right), H\rangle + \dots\\ &= \hat{W}(C) + \langle D_C\hat{W}(C), H\rangle + \dots\end{aligned}$$

Compare this result with a discussion in the literature [?], p.213. ■

2.9 Formal second derivatives

The second derivatives are needed for the investigation of convexity properties and ellipticity.

$$D^2(\|F\|^2).(H,H) = D(D(\|F\|^2).H).H = D(2\langle F,H\rangle).H = 2\langle H,H\rangle = 2\|H\|^2\,.$$

$$\begin{aligned}D^2(\|F\|^p).(H,H) &= D(D(\|F\|^p).H).H = D(p\|F\|^{p-2}\cdot\langle F,H\rangle).H\\ &= p\cdot(p-2)\|F\|^{p-4}\langle F,H\rangle^2 + p\|F\|^{p-2}\cdot\langle H,H\rangle\,.\end{aligned}$$

$$\begin{aligned}D^2(\|F\|).(H,H) &= D(D(\|F\|).H).H = D(\langle\frac{F}{\|F\|},H\rangle).H\\ &= \frac{\langle H,H\rangle}{\|F\|} - \frac{\langle F,H\rangle\langle F,H\rangle}{\|F\|^3} = \frac{\|H\|^2}{\|F\|} - \frac{\langle F,H\rangle^2}{\|F\|^3}\,.\end{aligned}$$

$$\begin{aligned}D^2(\|F^TF\|^2).(H,H) &= D(D(\|F^TF\|^2).H).H\\ &= D(2\langle F^TF, F^TH + H^TF\rangle).H\\ &= 2\langle F^TH + H^TF, F^TH + H^TF\rangle +\\ &\quad\; 2\langle F^TF, H^TH + H^TH\rangle\\ &= 2\langle F^TH + H^TF, F^TH + H^TF\rangle +\\ &\quad\; 4\langle F^TF, H^TH\rangle\\ &= 2\|F^TH + H^TF\|^2 + 4\langle F^TF, H^TH\rangle\,.\end{aligned}$$

$$\begin{aligned}
D^2(\frac{1}{p}\|F^TF\|^p).(H,H) &= D(D(\frac{1}{p}\|F^TF\|^p).H).H \\
&= D(\|F^TF\|^{p-2}\langle F^TF, F^TH + H^TF\rangle).H \\
&= (p-2)\|F^TF\|^{p-4}\langle F^TF, F^TH + H^TF\rangle^2 + \\
&\quad \|F^TF\|^{p-2}[\langle F^TF, 2H^TH\rangle + \\
&\qquad \langle F^TH + H^TF, F^TH + H^TF\rangle] \\
&= (p-2)\|F^TF\|^{p-4}\langle F^TF, F^TH + H^TF\rangle^2 + \\
&\quad \|F^TF\|^{p-2}[2\langle F^TF, H^TH\rangle + \|F^TH + H^TF\|^2]\,.
\end{aligned}$$

$$\begin{aligned}
D^2(\|F^TF - \mathbb{1}\|^2).(H,H) &= D(D(\|F^TF - \mathbb{1}\|^2).H).H \\
&= D(2\langle F^TF - \mathbb{1}, F^TH + H^TF\rangle).H \\
&= 2\{\langle F^TH + H^TF, F^TH + H^TF\rangle + \\
&\qquad \langle F^TF - \mathbb{1}, H^TH + H^TH\rangle\} \\
&= 2\{\langle F^TH + H^TF, F^TH + H^TF\rangle + \\
&\qquad 2\langle F^TF - \mathbb{1}, H^TH\rangle\} \\
&= 2\|F^TH + H^TF\|^2 + 4\langle F^TF - \mathbb{1}, H^TH\rangle\,.
\end{aligned}$$

$$\begin{aligned}
D^2(\|F - \mathbb{1}\|^2).(H,H) &= D(D(\|F - \mathbb{1}\|^2).H).H \\
&= D(2\langle F - \mathbb{1}, H\rangle).H = 2\langle H, H\rangle = 2\|H\|^2\,.
\end{aligned}$$

$$\begin{aligned}
D^2((\mathrm{tr}(F^TF - \mathbb{1})^2).(H,H) &= D^2(\langle F^TF - \mathbb{1}, \mathbb{1}\rangle^2).(H,H) \\
&= D(2\langle F^TF - \mathbb{1}, \mathbb{1}\rangle\langle F^TH + H^TF, \mathbb{1}\rangle).H \\
&= 2\{\langle F^TH + H^TF, \mathbb{1}\rangle\langle F^TH + H^TF, \mathbb{1}\rangle + \\
&\qquad 2\langle F^TF - \mathbb{1}, \mathbb{1}\rangle\langle H^TH, \mathbb{1}\rangle\} \\
&= 2\{\langle F^TH + H^TF, \mathbb{1}\rangle\langle F^TH + H^TF, \mathbb{1}\rangle + \\
&\qquad 2\langle F^TF - \mathbb{1}, \mathbb{1}\rangle\|H\|^2\} \\
&= 2\{\langle F^TH + H^TF, \mathbb{1}\rangle^2 + 2\langle F^TF - \mathbb{1}, \mathbb{1}\rangle\|H\|^2\} \\
&= 8\langle F, H\rangle^2 + 4(\|F\|^2 - 3)\|H\|^2\,.
\end{aligned}$$

$$\begin{aligned}
D^2(\frac{1}{\det[F]}).(H,H) &= D((-1)(\det[F])^{-1}\langle F^{-T}, H\rangle).H \\
&= (\det[F])^{-1}\langle F^{-T}, H\rangle^2 + (\det[F])^{-1}\langle HF^{-1}, (HF^{-1})^T\rangle\,.
\end{aligned}$$

$$\begin{aligned}
D^2(\det[F]).(H,H) &= D(\det[F]\langle H^T, F^{-1}\rangle).H \\
&= D(\det[F]).H\langle H^T, F^{-1}\rangle + \det[F]D(\langle H^T, F^{-1}\rangle).H \\
&= \det[F]\langle H^T, F^{-1}\rangle\langle H^T, F^{-1}\rangle + \det[F]\langle H^T, -F^{-1}HF^{-1}\rangle \\
&= \det[F]\{\langle F^{-T}, H\rangle^2 - \langle F^{-T}H^T, HF^{-1}\rangle\} \\
&= \det[F]\{\langle F^{-T}, H\rangle^2 - \langle (HF^{-1})^T, HF^{-1}\rangle\}\,.
\end{aligned}$$

But with Lemma **??** it follows

$$D^2(\det[F]).(H,H) = 2\langle \operatorname{Adj} H, F^T\rangle\,,$$

which shows the unexpected formula:

$$\det[F]\{\langle F^{-T}, H\rangle^2 - \langle (HF^{-1})^T, HF^{-1}\rangle\} = 2\langle \operatorname{Adj} H, F^T\rangle\,, \tag{2.6}$$

and for $F = 1\!\!1$ we obtain (cf. Satz **??**)

$$\begin{aligned} 1\{\langle 1\!\!1, H\rangle^2 - \langle H^T, H\rangle\} &= 2\langle \operatorname{Adj} H, 1\!\!1\rangle \\ \{\operatorname{tr}(H)^2 - \operatorname{tr}(H^2)\} &= 2\operatorname{tr}(\operatorname{Adj} H)\,. \end{aligned} \tag{2.7}$$

$$\begin{aligned} D^2(\frac{1}{\det[F^TF]}).(H,H) &= D(-2\cdot(\det[F])^{-2}\langle F^{-T}, H\rangle).H \\ &= 4\cdot(\det[F])^{-2}\langle F^{-T}, H\rangle^2 + \\ &\quad 2\cdot(\det[F])^{-2}\langle HF^{-1}, (HF^{-1})^T\rangle\,. \end{aligned}$$

$$\begin{aligned} D^2(\ln\det[F]).(H,H) &= D(\langle F^{-T}, H\rangle).H \\ &= \langle -F^{-T}H^TF^{-T}, H\rangle \\ &= -\langle F^{-T}H^T, HF^{-1}\rangle = -\langle (HF^{-1})^T, HF^{-1}\rangle\,. \end{aligned}$$

$$\begin{aligned} D^2((\ln\det[F])^2)).(H,H) &\underbrace{=}_{\det[F]>0} D(2\ln\det[F]\cdot\langle F^{-T}, H\rangle).H \\ &= 2\ln\det[F]\ \langle -F^{-T}H^TF^{-T}, H\rangle + 2\langle F^{-T}, H\rangle\langle F^{-T}, H\rangle \\ &= -2\ln\det[F]\langle F^{-T}H^T, HF^{-1}\rangle + 2\langle F^{-T}, H\rangle^2 \\ &= -2\ln\det[F]\langle (HF^{-1})^T, HF^{-1}\rangle + 2\langle F^{-T}, H\rangle^2\,. \end{aligned}$$

$$\begin{aligned} D^2(\ln(\det[F^TF])).(H,H) &\underbrace{=}_{\det[F]>0} 2D^2(\ln(\det[F])).(H,H) \\ &= -2\langle (HF^{-1})^T, HF^{-1}\rangle\,. \end{aligned} \tag{2.8}$$

Lemma 2.42 (Formal second derivative of $\Psi(F) := W(F^T + F)$)
Let $W : \mathrm{Sym}(3) \mapsto \mathbb{R}$. Then the second derivative of $\Psi(F) := W(F^T + F)$ is given by:

$$D^2\Psi(F).(H,H) = D^2W(F^T+F).(H^T+H, H^T+H)\,.$$

If $D^2W(S).(H,H) \geq c^+\cdot\|H\|^2$, then

$$D^2\Psi(F).(H,H) \geq c^+\cdot\|H^T+H\|^2\,,$$

and one only looses strict convexity of Ψ (choose $H^T = -H$) . ■

Lemma 2.43 (Formal second derivative of $\Psi(F) := W(F^TF)$)
Let $W : \mathrm{PSym}(3) \mapsto \mathbb{R}$. Then the second derivative of $\Psi(F) := W(F^TF)$ is given by:

$$D^2\Psi(F).(H,H) = 2\langle \partial_C W(F^TF), H^TH\rangle + \partial_C^2 W(F^TF).(F^TH + H^TF, F^TH + H^TF)\,.$$

Proof. We formally derive the expression of the second derivative of $\Psi(F) := W(F^TF)$.

$$\begin{aligned} W((F+H)^T(F+H)) &= W(F^TF + F^TH + H^TF + H^TH) \\ &= W(F^TF) + \langle \partial_C W(F^TF), F^TH + H^TF\rangle + \cdots \end{aligned}$$

thus $D\Psi(F).H = \langle \partial_C W(F^TF), F^TH + H^TF\rangle$ and expanding this term itself it follows

$$\begin{aligned} &\langle \partial_C W((F+H)^T(F+H)), (F+H)^TH + H^T(F+H)\rangle = \\ &\quad= \langle \partial_C W((F^TF + F^TH + H^TF + H^TH)), (F^TH + H^TF + H^TH + H^TH)\rangle \\ &\quad= \langle \partial_C W(F^TF) + \partial_C^2 W(F^TF).(F^TH + H^TF + \ldots), (F^TH + H^TF + 2H^TH)\rangle \\ &\quad= \langle \partial_C W(F^TF), F^TH + H^TF\rangle + 2\langle \partial_C W(F^TF), H^TH\rangle \\ &\qquad + \partial_C^2 W(F^TF).(F^TH + H^TF, F^TH + H^TF) + \ldots \end{aligned}$$

Thus for $\Psi(F) := W(F^TF)$

$$D^2\Psi(F).(H,H) = 2\langle \partial_C W(F^TF), H^TH\rangle + \partial_C^2 W(F^TF).(F^TH + H^TF, F^TH + H^TF)\,. \blacksquare$$

3 Convexity

Let us now come to the important notion of convexity.

Definition 3.1 (Convex sets)
A set K is called convex, whenever

$$\lambda\, F_1 + (1-\lambda)\, F_2 \in K \tag{3.9}$$

for $F_1, F_2 \in K$ and $\lambda \in (0,1)$.

In geometric terms this definition says that the line, joining F_1 and F_2 is included in K. Obviously, every vectorspace is also a convex set, but consider for example the set of all positive definite matrices PSym. This is not a vectorspace, since multiplication with negative real numbers will make the positive definite matrix negative definite. However, PSym is a convex set, since if $P_1\,, P_2$ are both positive definite, then the matrix $\lambda P_1 + (1-\lambda)P_2$ is still symmetric and positive definite.

Definition 3.2 (Convexity of functions)
Let K be a convex set and let $W : K \mapsto \mathbb{R}$. We say that W is convex if

$$W(\lambda F_1 + (1-\lambda)F_2) \leq \lambda W(F_1) + (1-\lambda)W(F_2)$$

for all $F_1, F_2 \in K$ and $\lambda \in (0,1)$.

Remark 3.3
Observe that in this definition it is necessary that the function W is **defined on a convex set** K.

Lemma 3.4 (Second derivative and convexity)
Let K be a convex set and let $W : K \mapsto \mathbb{R}$ be two-times continuously differentiable. Then it is equivalent:

1. *W ist convex.*
2. *$D^2W(F).(H,H) \geq 0 \quad \forall F \in K, \ \forall H \in Lin(K)$.*

Proof. [?], page 27. ■

Remark 3.5
Here it should be noted that the possible increments H have to be considered in the linear hull of the convex set K and that certainly not every convex set is a linear vecorspace as seen above for the set PSym.

Lemma 3.6 (Convexity of the square)
Let $P : \mathbb{R}^n \mapsto \mathbb{R}$ be convex and $P(Z) \geq 0$. Then the function

$$Z \in \mathbb{R}^n \mapsto [P(Z)] \cdot [P(Z)]$$

is convex.

Proof. Assume first that P is a smooth function. The second differential of $E(Z) = P(Z) \cdot P(Z)$ can be easily calculated. We get

$$\begin{aligned} D_Z E(Z).H &= P(Z) \cdot D_Z P(Z).H + D_Z P(Z).H \cdot P(Z) \\ D_Z^2 E(Z).(H,H) &= 2\big(P(Z) \cdot D_Z^2 P(Z).(H,H) + D_Z P(Z).H \cdot D_Z P(Z).H\big) \geq 0. \end{aligned}$$

Hence $E(Z)$ is convex. In the non-smooth case we proceed as follows:

$$E(\lambda Z_1 + (1-\lambda)Z_2) = [P(\lambda Z_1 + (1-\lambda)Z_2)] \cdot [P(\lambda Z_1 + (1-\lambda)Z_2)].$$

The assumed convexity of P shows that

$$[P(\lambda Z_1 + (1-\lambda)Z_2)] \leq [\lambda P(Z_1) + (1-\lambda)P(Z_2)].$$

Since the square function is a monotone increasing function for positive values and assuming that $[\lambda P(Z_1) + (1-\lambda)P(Z_2)]$ is positive we get the estimate

$$E(\lambda Z_1 + (1-\lambda)Z_2) \leq [\lambda P(Z_1) + (1-\lambda)P(Z_2)]^2 .$$

However, since the square function is itself convex we may proceed to write

$$\begin{aligned} E(\lambda Z_1 + (1-\lambda)Z_2) &\leq \lambda P(Z_1)^2 + (1-\lambda)P(Z_2)^2 \\ &= \lambda E(Z_1) + (1-\lambda)E(Z_2). \end{aligned}$$

The proof is complete. ■

Corollary 3.7
Let $P : \mathbb{R}^n \mapsto \mathbb{R}$ *be convex and assume that* $P(Z) \geq 0$. *Then the function*

$$Z \in \mathbb{R}^n \mapsto [P(Z)]^p, \ p \geq 1$$

is convex.

Proof. The same ideas as before carry over to this situation. ■

Lemma 3.8 (Convexity and monotone composition)
Let $P : \mathbb{R}^n \mapsto \mathbb{R}$ *be convex and let* $m : \mathbb{R} \mapsto \mathbb{R}$ *be convex and monotone increasing. Then the function* $\mathbb{R}^n \mapsto \mathbb{R}, \ X \mapsto m(P(X))$ *is convex.*

Proof. A direct check of the convexity condition. ■

Remark 3.9 (Nonconvexity of mixed products)
Let $P_i : \mathbb{R}^n \mapsto \mathbb{R}, \ i = 1, 2$ *be convex and assume* $P_i \geq 0$. *Then the functions*

$$\begin{aligned} &Z \in \mathbb{R}^n \mapsto [P_1(Z)] \cdot [P_2(Z)] \,, \\ &Z \in \mathbb{R}^n \mapsto [P_1(Z)]^q \cdot [P_2(Z)]^p, \quad p, q \geq 1 \,, \end{aligned}$$

are in general non-convex. As simple examples may serve $x \mapsto x^2(x-1)^2, \ x \mapsto e^x \cdot x^2$. *The function* $(x, y) \mapsto x^2 \cdot y^2$ *may serve as an example where functions in different variables are convex and positive, but their product is not convex.* ■

Remark 3.10
In order that $W : K \mapsto \mathbb{R}$ *be convex it is not sufficient to assume only*

$$D^2W(C).(H, H) \geq 0$$

for all $C \in K, \ \forall H \in K$. *Since for example with* $W : \mathrm{PSym} \mapsto \mathbb{R}, \quad W(C) = \det[C]$ *we have that* $K = \mathrm{PSym}$ *is a convex set (cone) and*

$$D^2W(C).(H, H) = 2 \, \langle C, \mathrm{Adj}\, H \rangle \geq 0$$

for $C, H \in \mathrm{PSym}$ *since* $\mathrm{Adj}\, H$ *is positive definite for positive definite* H[1] *, but* $W(C) = \det[C]$ *is not convex as a function of* C. ■

We deduce

Lemma 3.11 (Convexity on $\mathbb{M}^{3\times 3}$ and $\mathrm{PSym}(3)$)
Let $C \in \mathrm{PSym}(3)$ *and* $W : \mathrm{PSym}(3) \mapsto \mathbb{R}$. *Assume that* $\forall H \in \mathrm{Sym}(3) : \ \partial_C^2 W(C).(H, H) \geq 0$ *and* $\partial_C W(C) \in \mathrm{PSym}_0(3)$. *Then the function*

$$\Psi : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}, \quad F \mapsto W(F^T F)$$

is convex.

---

[1] Every positive definite H can be written as $H = X^T X$. Then $\mathrm{Adj}\, H = \mathrm{Adj}\, X^T X = \mathrm{Adj}\, X \ \mathrm{Adj}\, X^T$ is positive definite.

Proof. Use Lemma **??** for the second derivative of W and observe that

$$Lin(\mathrm{PSym}) = \mathrm{Sym}\ .$$

Apply then basic properties of the scalar product. ■

Let us again remark that $\mathbb{M}^{3\times 3}$ and $\mathrm{Sym}(3)$ are vectorspaces while PSym is a positive, convex cone. Thus there are no problems involved with the domain of definition of the convex functions. In all cases the domain of definition is convex.

Example 3.12

Let $W(C) = \|C\|^2$, *then* $\Psi(F) = \|F^TF\|^2$ *is convex in* F, *as has already been seen. Here, using Lemma* **??** *shows* $\partial_C W(C) = 2C \in \mathrm{PSym}(3)$ *and* $\partial_C^2 W(C).(H,H) = 2\|H\|^2 \geq 0$. *In fact,* $\Psi(F) = \|F^TF\|^2$ *is strictly convex, which can be shown by considering the higher derivatives.*

Example 3.13

Let $A, C \in \mathrm{PSym}(3)$; $W(C) = \langle AC, CA\rangle$, *then* $\Psi(F) = \langle AF^TF, F^TFA\rangle$ *is convex in* F *since* $\partial_C W(C) = A^TCA + ACA^T \in \mathrm{PSym}(3)$ *and* $\partial_C^2 W(C).(H,H) = 2\langle AH, HA\rangle \geq \lambda^2_{\min}(A)\cdot\|H\|^2$ *and the conclusion follows with Lemma* **??**.

Example 3.14

Let $A, C \in \mathrm{PSym}(3)$ *and* $W(C) = \langle C, A\rangle$ *(*W *is linear in* C*). Then* $\Psi(F) = \langle F^TF, A\rangle$ *is convex since* $\partial_C W(C) = A \in \mathrm{PSym}(3)$ *and* $\partial_C^2 W(C).(H,H) = 0$. *If, on the other side* $W(C) = \|AC\|^2$ *then one cannot use Lemma* **??** *to conclude the convexity of* $\Psi(F) = \|F^TFA\|^2$ *since in general* $\partial_C W(C) \notin \mathrm{PSym}(3)$ *because* $\partial_C W(C).H = 2\langle A^TAC, H\rangle$ *and* $A^TAC \notin \mathrm{PSym}(3)$.

Remark 3.15

For the conclusion in Lemma **??** *it is not sufficient that* $\partial_C^2 W(C).(H,H) \geq c^+\cdot\|H\|^2$ *for* $H \in \mathrm{Sym}(3)$ *and* $\partial_C W(C) \in \mathrm{Sym}(3)$. *We show this with a counterexample.*

Let $W(C) = -\ln\det[C]$. Then

$$\forall H \in \mathrm{PSym}(3) : \partial_C W(C).H = -\langle C^{-T}, H\rangle = \langle -C^{-1}, H\rangle\ .$$

Moreover $\partial_C W(C) \in \mathrm{Sym}(3)$ since $-C^{-1} \in \mathrm{Sym}(3)$ if $C \in \mathrm{PSym}(3)$ and

$$\begin{aligned}\forall H \in \mathrm{Sym}(3): \quad \partial_C^2 W(C).(H,H) &= \langle (C^{-1}H)^T, C^{-1}H\rangle = \langle HC^{-1}, C^{-1}H\rangle \\ &\geq \lambda^2_{\min}(C^{-1})\ \|H\|^2\ .\end{aligned}$$

Therefore, the function $W(C) = -\ln\det[C]$ is strictly convex on $\mathrm{PSym}(3)$ [**?**], but the function $\Psi(F) = -\ln\det[F^TF] = -\ln(\det[F]^2)$ is not convex on $M^{3\times 3}$ which can be easily seen: choose $F_1, F_2 \in \mathbb{M}^{3\times 3}$

$$F_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\ , \quad F_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{pmatrix}\ .$$

For $0 < \lambda < 1$ we have $\det[F_1] = 1$ and $\det[F_2] = 4$. Thus

$$-\lambda \ln(\det[F_1]^2) - (1-\lambda)\ln(\det[F_2]^2) = -2(1-\lambda)\ln 4 < 0$$

and

$$\begin{aligned}\det[\lambda F_1 + (1-\lambda)F_2] &= \det\begin{pmatrix}\lambda + (1-\lambda) & 0 & 0\\ 0 & \lambda - (1-\lambda) & 0\\ 0 & 0 & \lambda - 4(1-\lambda)\end{pmatrix}\\ &= (2\lambda - 1)\,(2\lambda - 4)\,.\end{aligned}$$

This implies for $\hat\lambda = \frac{3}{8}$

$$(2\hat\lambda - 1)\,(2\hat\lambda - 4) = \frac{52}{64} < 1\,.$$

Therefore

$$-\ln(\det[\lambda F_1 + (1-\lambda)F_2]^2) = -2\ln[2\lambda - 1)\,(2\lambda - 4)] = -2\ln\frac{52}{64} > 0\,,$$

contradicting the convexity condition for $\Psi(F) = -\ln(\det[F]^2)$. ■

Remark 3.16
It is also not sufficient for the conclusion in Lemma **??** *that* $\partial_C^2 W(C).(H,H) \geq 0$ *for all* $H \in \mathrm{PSym}(3)$ *and* $\partial_C W(C) \in \mathrm{PSym}_0(3)$. *Again, we provide a counterexample.*

Example 3.17
Let $W(C) = \det[C]$. *Then* $\partial_C W(C).H = \langle H, \mathrm{Adj}\, C\rangle$ *and for* $C \in \mathrm{PSym}$ *it follows that* $\mathrm{Adj}\, C \in \mathrm{PSym}_0$. *Moreover* $\partial_C^2 W(C).(H,H) = 2\langle \mathrm{Adj}\, H, C\rangle \geq 0$ *for* $H \in \mathrm{PSym}_0$. *But neither* $\Psi(F) = \det[F^T F]$ *is convex nor* $W(C) = \det[C]$ *is convex on* PSym.

3.1 Singularity and convexity

Remark 3.18 (Singularity for $\det[F] \to 0$ and convexity)
Let $\Psi(F)$ *such that* $\Psi(F) \to \infty$ *for* $\det[F] \to 0$. *Choose* $F_1, F_2 \in \mathbb{M}^{3\times 3}$

$$F_1 = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\,, \quad F_2 = \begin{pmatrix}-1 & 1 & 1\\ 0 & -1 & 1\\ 0 & 0 & 1\end{pmatrix}\,,$$

this implies $\det[F_1] = \det[F_2] = 1$ *and for* $\lambda = \frac{1}{2} + \varepsilon$ *it holds*

$$\begin{aligned}\det[\lambda F_1 + (1-\lambda)F_2] &= \det\begin{pmatrix}\lambda - (1-\lambda) & 1-\lambda & 1-\lambda\\ 0 & \lambda - (1-\lambda) & 1-\lambda\\ 0 & 0 & \lambda + (1-\lambda)\end{pmatrix}\\ &= (2\lambda - 1)^2 = 4\varepsilon^2\,.\end{aligned}$$

The convexity condition for Ψ *is violated if* $\varepsilon \to 0$. *Thus, a potential* Ψ *with such a singularity can never be convex on* $\mathbb{M}^{3\times 3}$.

Lemma 3.19 (Convexity and frame-indifference)
Assume that $W : \mathrm{GL}^+(3) \mapsto \mathbb{R}$ *is frame-indifferent and*

$$\min_{F \in \mathrm{GL}^+(3)} W(F) = W(1\!\!1)\,, \quad W(F) > W(1\!\!1) \quad \forall\, F \notin \mathrm{SO}(3)\,. \tag{3.10}$$

Then W *is not convex.*

Proof. Frame-indifference implies that $W(Q\,1\!\!1) = W(1\!\!1)$. Thus W is minimal on SO(3), the group of proper orthogonal matrices. However, $SO(3)$ is not a convex set. We may construct $F = \lambda\, Q_1 + (1-\lambda)\, Q_2 \notin \mathrm{SO}(3)$. Assume that W is convex. Then

$$W(F) = W(\lambda\, Q_1 + (1-\lambda)\, Q_2) \le \lambda\, W(Q_1) + (1-\lambda)\, W(Q_2) = W(1\!\!1)\,, \tag{3.11}$$

contradicting $W(F) > W(1\!\!1)$ forall $F \notin \mathrm{SO}(3)$. ■

Lemma 3.20 ($\exp : \mathrm{Sym}(3) \mapsto \mathrm{PSym}(3)$ **is a homeomorphism)**
Let $A \in \mathrm{Sym}(3)$, *then the matrix exponential mapping*

$$\exp(X) := \sum_{j=1}^{\infty} \frac{1}{j!} X^j$$

is such that $\exp(A) \in \mathrm{PSym}$ *and is continuous and the inverse function* $\ln : \mathrm{PSym}(3) \mapsto \mathrm{Sym}(3)$ *exists and is continuous.*

Proof. See [**?**], page 19. ■

Lemma 3.21 ($D(\exp(X)) : \mathrm{Sym}(3) \mapsto L(\mathrm{Sym}(3), \mathrm{PSym}(3))$ **is bijective)**
The derivative of the matrix-exponential $\exp(X)$ *is bijective on* $L(\mathrm{Sym}(3), Lin(\mathrm{PSym}(3)))$.

Proof. See [**?**, p.208]. ■

Corollary 3.22 ($\exp : \mathrm{Sym}(3) \mapsto \mathrm{PSym}(3)$ **is a diffeomorphism)**

Proof. The bijectivity of $D(\exp(C))$ and the inverse function theorem imply the differentiability of the inverse function $\ln C$. ■

Lemma 3.23 (Derivative of scalarvalued isotropic tensorfunctions)
Let $\Psi : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ *be an isotropic scalarvalued tensorfunction. Then the derivative* $D\Psi : \mathbb{M}^{3\times 3} \mapsto \mathbb{M}^{3\times 3}$ *is an isotropic tensorfunction.*

Proof.

$$\begin{aligned}
\Psi(F+H) &= \Psi(F) + D\Psi(F).H + \ldots = \Psi(F) + \langle D\Psi(F), H\rangle + \ldots \\
&\qquad \text{because of isotropy of } \Psi \text{ it holds for all } Q \in O(3) \\
\Psi(F+H) &= \Psi(Q^T(F+H)Q) = \Psi(Q^TFQ + Q^THQ) \\
&= \Psi(Q^TFQ) + D\Psi(Q^TFQ).(Q^THQ) + \ldots \\
&= \Psi(F) + \langle Q D\Psi(Q^TFQ)Q^T, H\rangle + \ldots
\end{aligned}$$

also for all $H \in \mathbb{M}^{3\times 3}$

$$\begin{aligned}\langle D\Psi(F), H\rangle &= \langle QD\Psi(Q^TFQ)Q^T, H\rangle \\ Q^TD\Psi(F)Q &= D\Psi(Q^TFQ)\,.\end{aligned} \qquad \blacksquare$$

3.2 Chain rule and Sansours formula

Lemma 3.24 (Sansours formula)
Let $A \in \mathrm{Sym}(3)$ and let $\Psi : \mathrm{Sym}(3) \mapsto \mathbb{R}$ be a differentiable isotropic scalarvalued tensorfunction. Set $W(A) = \Psi(\exp(A))$. Then the chain rule obtains in the following form:

1.

$$\begin{aligned} D_A\left[\Psi(\exp(A))\right] &= \exp(A)\cdot D\Psi(\exp(A))\,, \\ D_A W(A) &= D\Psi(\exp(A))\cdot\exp(A)\,. \end{aligned}$$

2. *Setting $A = \ln C$ it holds as well*

$$D\Psi(C) = D_A W(\ln C)\cdot C^{-1}\,.$$

Proof. We follow [?]. We expand first the matrix-exponential function. It holds

$$\begin{aligned}\exp(X+H) &= 1\!\!1 + (X+H) + \frac{1}{2}(X+H)^2 + \frac{1}{6}(X+H)^3 + \ldots \\ &= 1\!\!1 + (X+H) + \frac{1}{2}(X^2 + XH + HX + H^2) + \\ &\frac{1}{6}(X^3 + XHX + HXX + H^2X + X^2H + XH^2 + HXH + H^3) + \ldots \\ &= 1\!\!1 + X + \frac{1}{2}X^2 + \frac{1}{6}X^3 + \ldots + \\ &\quad H + \frac{1}{2}(XH + HX) + \frac{1}{6}(XXH + XHX + HXX) \\ &= \exp(X) + \\ &\quad \underbrace{H + \frac{1}{2}(XH + HX) + \frac{1}{6}(XXH + XHX + HXX) + \ldots}_{D(\exp(X)).H}\end{aligned}$$

Now we consider the expansion of $\Psi(\exp(A))$

$$\begin{aligned}\Psi(\exp(A+H)) &= \Psi(\exp(A) + D(\exp(A)).H + \ldots) \\ &= \Psi(\exp(A)) + \langle D\Psi(\exp(A)), D(\exp(A)).H\rangle + \ldots \\ &= \Psi(\exp(A)) + \langle D\Psi(\exp(A)), H + \frac{1}{2}(AH + HA)\rangle + \\ &\quad \langle D\Psi(\exp(A)), \frac{1}{6}(AAH + AHA + HAA) + \ldots\rangle + \ldots\end{aligned}$$

$$
\begin{aligned}
&= \Psi(\exp(A)) + \langle D\Psi(\exp(A)), H\rangle + \frac{1}{2}\langle D\Psi(\exp(A)), AH + HA\rangle + \\
&\frac{1}{6}\langle D\Psi(\exp(A)), AAH + AHA + HAA\rangle + \dots \\
&= \Psi(\exp(A)) + \langle D\Psi(\exp(A)), H\rangle + \\
&\frac{1}{2}\left[\langle A^T D\Psi(\exp(A)), H\rangle + \langle D\Psi(\exp(A))A^T, H\rangle\right] + \\
&\quad \frac{1}{6}\left[\langle A^T A^T D\Psi(\exp(A)), H\rangle + \langle A^T D\Psi(\exp(A))A^T, H\rangle\right] + \\
&\qquad \frac{1}{6}\langle D\Psi(\exp(A))A^T A^T, H\rangle + \dots \\
&\text{since here } A = A^T\text{, it follows} \\
&= \Psi(\exp(A)) + \langle D\Psi(\exp(A)), H\rangle + \\
&\frac{1}{2}\left[\langle AD\Psi(\exp(A)), H\rangle + \langle D\Psi(\exp(A))A, H\rangle\right] + \\
&\frac{1}{6}\left[\langle AAD\Psi(\exp(A)), H\rangle + \langle AD\Psi(\exp(A))A, H\rangle\right] + \\
&\qquad \frac{1}{6}\langle D\Psi(\exp(A))AA, H\rangle + \dots\,,
\end{aligned}
$$

and since $D\Psi$ is an isotropic tensorfunction and $\exp(A)$ is isotropic as well it follows that $D\Psi(\exp(A))$ is an isotropic tensorfunction and therefore

$$
D\Psi(\exp(A)) \cdot A = A \cdot D\Psi(\exp(A))\,.
$$

This implies

$$
\begin{aligned}
\Psi(\exp(A+H)) &= \Psi(\exp(A)) + \langle D\Psi(\exp(A)), H\rangle + \langle D\Psi(\exp(A))A, H\rangle + \\
&\quad \frac{1}{2}\left[\langle D\Psi(\exp(A))A^2, H\rangle + \dots\right] \\
&= \Psi(\exp(A)) + \langle D\Psi(\exp(A)) \cdot [\mathbb{1} + A + \frac{1}{2}A^2 + \dots], H\rangle \\
&= \Psi(\exp(A)) + \langle D\Psi(\exp(A)) \cdot \exp(A), H\rangle \\
&\text{and since } \Psi \text{ is isotropic, } D\Psi(\exp(A)) \text{ and } \exp(A) \text{ commute} \\
&= \Psi(\exp(A)) + \langle \exp(A) \cdot D\Psi(\exp(A)), H\rangle \\
&\text{but at the same time it holds} \\
\Psi(\exp(A+H)) &= \Psi(\exp(A)) + \langle D_A\Psi(\exp(A)), H\rangle + \dots
\end{aligned}
$$

implying $\forall H \in \operatorname{Sym}(3)$

$$
\begin{aligned}
\langle D_A\Psi(\exp(A)), H\rangle &= \langle \exp(A) \cdot D\Psi(\exp(A)), H\rangle \\
\langle D_A W(A), H\rangle &= \langle \exp(A) \cdot D\Psi(\exp(A)), H\rangle\,.
\end{aligned}
$$

Setting $A = \ln C$ shows the other part. ■

3.3 Convexity and ellipticity

Let $W \in C^2(\mathbb{M}^{3\times 3}, \mathbb{R})$ be a scalarvalued function. In the following let $F, H \in \mathbb{M}^{3\times 3}$ and $\eta, \xi \in \mathbb{R}^3$. Subsequently we classify the possibilities for the second derivative of W. The matrix representation of D^2W is also called elasticity tensor or stiffness matrix. We call W in a given $F \in \mathbb{M}^{3\times 3}$

1. uniformly positive or stable, if $\quad D^2W(F).(H,H) \geq c^+ \cdot \|H\|^2$.
2. strictly Legendre elliptic, if $\quad D^2W(F).(H,H) > 0 \quad \forall H \neq 0$.
3. strictly convex, if
   $W(\lambda F_1 + (1-\lambda)F_2) < \lambda \cdot W(F_1) + (1-\lambda) \cdot W(F_2) \quad \lambda \in (0,1)$.
4. Legendre elliptic, if $D^2W(F).(H,H) \geq 0$.
5. convex, if $W(\lambda F_1 + (1-\lambda)F_2) \leq \lambda \cdot W(F_1) + (1-\lambda) \cdot W(F_2) \quad \lambda \in (0,1)$.
6. strictly uniformly Korn-elliptic, if $\quad D^2W(F).(H,H) \geq c^+ \cdot \|H^T + H\|^2$.
7. strictly Korn-elliptic, if $D^2W(F).(H,H) > 0 \quad \forall H : \quad H^T + H \neq 0$.
8. uniformly Korn elliptic, if $\quad D^2W(F).(H,H) \geq c^+ \cdot \|F^T H + H^T F\|^2$.
9. polyconvex, if there is a convex function $P : \mathbb{M}^{3\times 3} \times \mathbb{M}^{3\times 3} \times \mathbb{R} \mapsto \mathbb{R}$ with $W(F) = P(F, \operatorname{Adj} F, \det[F])$.
10. strictly polyconvex, if W is polyconvex and P strictly convex.
11. quasiconvex, if for all $\Omega \subset \mathbb{R}^3$ and all $F \in \mathbb{M}^{3\times 3}$ and all $v \in C_0^\infty(\Omega)$ it holds
    $$W(F) \cdot |\Omega| = \int_\Omega W(F) \ \mathrm{dx} \leq \int_\Omega W(F + Dv(x)) \ \mathrm{dx}\,,$$
    i.e., $u(x) = F \cdot x + c$ is minimizer to his own boundary conditions.
12. uniformly strictly quasiconvex, if for all $\Omega \subset \mathbb{R}^3$ and all $F \in \mathbb{M}^{3\times 3}$ and all $v \in C_0^\infty(\Omega)$ it holds
    $$\int_\Omega W(F) + c^+ \cdot \|Dv\|^2 \ \mathrm{dx} \leq \int_\Omega W(F + Dv(x)) \ \mathrm{dx}\,.$$
13. strictly rank-one convex, if the function $f(t) := W(F + t \cdot (\eta \otimes \xi))$ is strictly convex in t.
14. rank-one convex, if the function $f(t) := W(F + t \cdot (\eta \otimes \xi))$ is convex in t for all $F \in \mathbb{M}^{3\times 3}$.
15. uniformly Legendre-Hadamard-elliptic (LH-elliptic), if $D^2W(F).(\eta \otimes \xi, \eta \otimes \xi) \geq c^+ \cdot \|\eta \otimes \xi\|^2$.

16. uniformly Legendre-Hadamard elliptic, if
$D^2W(F).(\eta \otimes \xi, \eta \otimes \xi) \geq c^+ \cdot \|\eta\|^2 \cdot \|\xi\|^2$.

17. strictly LH-elliptic, if $D^2W(F).(\eta \otimes \xi, \eta \otimes \xi) > 0 \quad \forall \quad \eta, \xi \neq 0$.

18. strongly elliptic, if
$\langle \xi, M(\eta).\xi \rangle > 0 \quad \forall \quad \eta, \xi \neq 0$, with

$$\langle \xi, M(\eta).\xi \rangle := D^2W(F).(\eta \otimes \xi, \eta \otimes \xi)$$

M is also called **acoustic tensor** and with this definition M is symmetric.

19. elliptic, if $\quad \det[M(\eta)] \neq 0 \quad \forall \eta \neq 0$.

20. Legendre-Hadamard elliptic, if $\quad D^2W(F).(\eta \otimes \xi, \eta \otimes \xi) \geq 0$.

The following relations between these definitions obtain:

1. (??)⇔(??)
Proof. (??)⇒(??) is obvious. Assume (??). Then

$$\forall H \neq 0 \quad D^2W(F).(\frac{H}{\|H\|}, \frac{H}{\|H\|}) > 0\,,$$

but the unit sphere is compact on $\mathbb{M}^{3\times3}$ and since $D^2W(F).(H,H)$ is continuous, the minimium is achieved, let us say in $\hat{H}$, thus

$$D^2W(F).(\frac{H}{\|H\|}, \frac{H}{\|H\|}) \geq D^2W(F).(\frac{\hat{H}}{\|\hat{H}\|}, \frac{\hat{H}}{\|\hat{H}\|})$$

$$\frac{1}{\|H\|^2} \cdot D^2W(F).(H,H) \geq D^2W(F).(\frac{\hat{H}}{\|\hat{H}\|}, \frac{\hat{H}}{\|\hat{H}\|})$$

set $c^+ = D^2W(F).(\frac{\hat{H}}{\|\hat{H}\|}, \frac{\hat{H}}{\|\hat{H}\|})$, therefore

$$D^2W(F).(H,H) \geq c^+ \cdot \|H\|^2\,.$$

Alternatively one might consider a contradiction argument. The benefit of this argument being the possibility to extend it to the infinite-dimensional case with appropriate modifications: assume $\|H_n\| = 1$ such that $D^2W(F).(H_n, H_n) \to 0$. Since the unit sphere is compact there is a subsequence H_{n_j} which converges to $\hat{H}$. Since $D^2W(F).(H,H)$ is continuous in H we have $D^2W(F).(\hat{H}, \hat{H}) = 0$. However, $\|\hat{H}\| = 1$, contrary to the assumption that $D^2W(F).(H,H) > 0 \; \forall \; H \neq 0$. ■

2. (??)⇒ (??)
Proof. Since W is smooth. ■

3. (??)$\not\Rightarrow$ (??)
   Proof. $W(F) = \|F\|^4$ is strictly convex, but
   $D^2W(F).(H,H) = 8\langle F,H\rangle^2 + 4\|F\|^2 \cdot \|H\|^2$ is not uniformly positive in $F = 0$. ■

4. (??)$\Rightarrow$ (??)
   Proof. $\|H^T + H\|^2 \leq 4 \cdot \|H\|^2$. ■

5. (??)$\Rightarrow$ (??)
   Proof. Obvious. ■

6. (??) $\Leftrightarrow$ (??)
   Proof. Again the differentiability of W. ■

7. (??)$\Leftrightarrow$ (??)
   Proof. Compactness argument. ■

8. (??)$\Rightarrow$ (??)
   Proof. Obvious. ■

9. (??)$\not\Rightarrow$ (??)
   Proof. Consider again $W(F) = \|F\|^4$. ■

10. (??)$\Rightarrow$(??)
    Proof. Use the properties of the tensorproduct. ■

11. (??)$\not\Rightarrow$(??)
    Proof. There are $F, H \neq 0$ such that $F^T H = 0$. ■

12. (??)$\Rightarrow$(??), if $\lambda_{\min}(F^T F) > 0$. More precisely

    $$D^2W(F).(\eta \otimes \xi, \eta \otimes \xi) \geq 2c^+ \cdot \lambda_{\min}(F^T F) \cdot \|\eta\|^2 \cdot \|\xi\|^2$$

    Proof. Use the properties of the tensorproduct and the scalarproduct. ■

13. (??)$\Rightarrow$(??)
    Proof. See [?]. ■

14. (**??**)⇒(**??**)
    Proof. Assume that W is polyconvex. We define $f(t) = W(F + t(\eta \otimes \xi))$ and check the convexity of f directly. Let $\lambda \in (0, 1)$.

$$\begin{aligned}
f(\lambda t_1 + (1-\lambda)t_2) &= W(F + [\lambda t_1 + (1-\lambda)t_2]\ (\eta \otimes \xi)) \\
&= P(F + [\lambda t_1 + (1-\lambda)t_2]\ (\eta \otimes \xi), \\
&\quad\ \operatorname{Adj}(F + [\lambda t_1 + (1-\lambda)t_2]\ (\eta \otimes \xi)), \\
&\quad \det[(F + [\lambda t_1 + (1-\lambda)t_2]\ (\eta \otimes \xi))])
\end{aligned}$$

Using the expansion formula for the determinant and the adjugate we obtain therefore

$$\begin{aligned}
f(\lambda t_1 + (1-\lambda)t_2) = P(&\lambda[F + t_1\ (\eta \otimes \xi)] + (1-\lambda)[F + t_2\ (\eta \otimes \xi)], \\
&\lambda[\operatorname{Adj}(F + t_1\ (\eta \otimes \xi))] + (1-\lambda)[\operatorname{Adj}(F + t_2\ (\eta \otimes \xi))], \\
&\lambda[\det[(F + t_1\ (\eta \otimes \xi))]] + (1-\lambda)[\det[(F + t_2\ (\eta \otimes \xi))]])
\end{aligned}$$

$$\begin{aligned}
f(\lambda t_1 + (1-\lambda)t_2) \le\ &\lambda\ P(F + t_1\ (\eta \otimes \xi), \\
&\qquad \operatorname{Adj}(F + t_1\ (\eta \otimes \xi)), \\
&\qquad \det[(F + t_1\ (\eta \otimes \xi))]) \\
&+ (1-\lambda)P(F + t_2\ (\eta \otimes \xi), \\
&\qquad \operatorname{Adj}(F + t_2\ (\eta \otimes \xi)), \\
&\qquad \det[(F + t_2\ (\eta \otimes \xi))]) \\
&= \lambda\ f(t_1) + (1-\lambda)\ f(t_2)
\end{aligned}$$

which proves the statement. ■

15. (**??**)$\not\Rightarrow$(**??**)
    Proof. [?]. ■

16. (**??**)⇒(**??**)
    Proof. [?]. ■

17. (**??**)$\not\Rightarrow$(**??**)
    Proof. [?]. ■

18. (**??**)⇒(**??**)
    Proof. Differentiate f twice w.r.t. t. ■

19. (??)$\Rightarrow$(??)
    Proof. Obvious. ■

20. (??)$\Leftrightarrow$(??)$\Leftrightarrow$(??) $\Leftrightarrow$(??)
    Proof. Use properties of the tensorproduct and compactness.

21. (??)$\not\Rightarrow$(??)
    Proof. Set $W(F) = \varepsilon\cdot\|F\|^2 + \det[F]$, then $D^2W(F).(H,H) = 2\varepsilon\|H\|^2 + 2\langle \operatorname{Adj} H, F^T\rangle$ and $D^2W(F).(H,H)$ is strictly Legendre-Hadamard elliptic. Choose $F = -\mu\mathbb{1}$ and $H = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$ with $\frac{1}{2} \le \lambda_i \le 1$. This implies

    $$\begin{aligned} D^2W(F).(H,H) &= 2\varepsilon(\lambda_1^2 + \lambda_2^2 + \lambda_3^2) - 2\mu(\lambda_3\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_1) \\ &\le 2\varepsilon \cdot 3 - 2\mu \cdot \frac{3}{4}\,. \end{aligned}$$

    If $\mu > 4\varepsilon$ then $D^2W(F).(H,H) < 0$. ■

22. (??)$\not\Rightarrow$(??)(??)(??)(??)(??)(??)(??) . ■

23. (??)$\not\Rightarrow$(??)
    Proof. Let $A \in \mathbb{M}^{3\times 3}$ with $\det[A] \neq 0$. Define $W(F) = \langle F, A\cdot F\rangle$, then $D^2W(F).(H,H) = 2\langle H, A\cdot H\rangle$ and $\langle \xi, M(\eta).\xi\rangle = \langle \eta\otimes\xi, A.(\eta\otimes\xi)\rangle = \langle \eta\otimes\xi, A.\eta\otimes\xi\rangle$, now choose $A.\eta = -\eta$. ■

24. (??)$\Leftrightarrow$(??)
    Proof. Differentiate f twice. ■

25. (??)$\not\Rightarrow$(??)
    Proof. Consider $f(t) = \|F + t\cdot(\eta\otimes\xi)\|^4$. ■

Apart for quasiconvexity all mentioned properties are pointwise conditions.

Example 3.25 (LH-ellipticity for $\|\operatorname{dev}(F + F^T - 2\,\mathbb{1})\|^2$)
The second differential of $\|\operatorname{dev}(F + F^T - 2\,\mathbb{1})\|^2$ *estimates* $\|\operatorname{dev}(H^T + H)\|^2$. *We write*

$$\begin{aligned} \|\operatorname{dev}(H^T + H)\|^2 &= \|H^T + H\|^2 - \frac{1}{n}\operatorname{tr}\left[H^T + H\right]^2 \\ &= \|H^T + H\|^2 - \frac{4}{n}\operatorname{tr}[H]^2 \quad \text{for}\, n = 3 \;\Rightarrow \\ &= 2\|H\|^2 + 2\operatorname{tr}[H]^2 - 4\langle \operatorname{Cof} H, \mathbb{1}\rangle - \frac{4}{n}\operatorname{tr}[H]^2 \end{aligned}$$

$$
\begin{aligned}
&= 2\|H\|^2 + \frac{2n-4}{n}\mathrm{tr}\,[H]^2 - 4\langle \mathrm{Cof}\, H, \mathbb{1}\rangle\,, \quad H = \xi\otimes\eta \;\Rightarrow \\
\|\,\mathrm{dev}(\xi\otimes\eta+\eta\otimes\xi)\|^2 &\geq 2\,\|\xi\|^2\,\|\eta\|^2\,.
\end{aligned}
$$

For $n=2$ it holds

$$
\begin{aligned}
\|\,\mathrm{dev}(H^T+H)\|^2 &= \|H^T+H\|^2 - \frac{1}{n}\,\mathrm{tr}\left[H^T+H\right]^2 \\
&= \|H^T+H\|^2 - \frac{4}{n}\,\mathrm{tr}\,[H]^2 \quad \textit{for } n=2 \;\Rightarrow \\
&= 2\|H\|^2 + 2\mathrm{tr}\,[H]^2 - 4\mathrm{det}[H] - \frac{4}{n}\,\mathrm{tr}\,[H]^2 \\
&= 2\|H\|^2 - 4\mathrm{det}[H]\,, \quad H = \xi\otimes\eta \;\Rightarrow \\
\|\,\mathrm{dev}(\xi\otimes\eta+\eta\otimes\xi)\|^2 &\geq 2\,\|\xi\|^2\,\|\eta\|^2\,.
\end{aligned}
$$

This shows uniform LH-ellipticity. However, $\|\,\mathrm{dev}(F+F^T-2\,\mathbb{1})\|^2$ is not uniformly convex. Take e.g., $F=\lambda^+(\mathbb{1}+A)$ with $A\in\mathfrak{so}(3)$.

Remark 3.26 (Ellipticity for differential-operators in Div-form)
Let $F=\nabla u(x)\in\mathbb{M}^{3\times3}$ and $A:\mathbb{M}^{3\times3}\mapsto\mathbb{M}^{3\times3}$ be a continuously differentiable function. We say that the PDE-system

$$
\mathrm{Div}\,A(\nabla u) = f\,, \quad u_{\partial\Omega} = g\,,
$$

is elliptic, if $B(F).(H,H) := \langle H, DA(F).H\rangle$ instead of $D^2W(F).(H,H)$ has the corresponding property. If the PDE-system is the Euler-Lagrange equation of a corresponding energetic formulation, then the ellipticity conditions can be read off in terms of the energy density W.

Example 3.27 (Ellipticity of $\mathrm{Div}\,\nabla u^T\nabla u=0$)
In this case $A(F)=F^TF$ and therefore $DA(F).H=F^TH+H^TF$, thus $B(F).(H,H):=\langle H, DA(F).H\rangle = \langle H, F^TH+H^TF\rangle$. We have

$$
\begin{aligned}
B(F).(H,H) &= \langle H, F^TH+H^TF\rangle \\
&= \langle (F^T-\mathbb{1})H, H\rangle + \langle H,H\rangle + \langle H^T(F-\mathbb{1}), H\rangle + \langle H^T, H\rangle \\
&\geq -2\|F-\mathbb{1}\|\cdot\|H\|^2 + \|H\|^2 + \langle H^T, H\rangle \\
&\quad \textit{for } H=\eta\otimes\xi \textit{ it follows} \\
B(F).(\eta\otimes\xi,\eta\otimes\xi) &\geq (1-2\|F-\mathbb{1}\|)\cdot\|\eta\|^2\|\xi\|^2\,.
\end{aligned}
$$

Hence, the operator A is elliptic as long as the pointwise condition $(1-2\|\nabla u(x)-\mathbb{1}\|)>0$ is satisfied.

How to know, whether a given operator is the differential of some energy? The answer follows in the next statement.

Lemma 3.28 (Pseudopotential and integrability conditions)
Let $A\in C^1(\mathbb{M}^{3\times3},\mathbb{M}^{3\times3})$. The following is equivalent:

1. *$\exists\, W \in C^2(\mathbb{M}^{3\times 3}, \mathbb{R}) : \quad DW(F) = A(F)$*
2. *$\forall\, F, H, \hat{H} \in \mathbb{M}^{3\times 3} : \quad \langle DA(F).H, \hat{H}\rangle = \langle DA(F).\hat{H}, H\rangle$ integrability conditions*

Proof. One direction is just the theorem of Schwarz on commuting second derivatives. For the reverse direction consider the so called **pseudopotential**

$$W(F) := \int_0^1 \langle A(tF), F\rangle \,\mathrm{dt}\,,$$

which is well defined on $\mathbb{M}^{3\times 3}$. Let $H \in \mathbb{M}^{3\times 3}$. Then

$$\begin{aligned} DW(F).H &= \int_0^1 \langle F, DA(tF).(tH)\rangle + \langle H, A(tF)\rangle \,\mathrm{dt} \\ &= \int_0^1 \langle \frac{\mathrm{d}}{\mathrm{dt}}\,[t\, A(tF)], H\rangle - t\,\langle DA(tF).F, H\rangle + \langle F, DA(tF).(tH)\rangle \,\mathrm{dt} \\ &= \int_0^1 \langle \frac{\mathrm{d}}{\mathrm{dt}}\,[t\, A(tF)], H\rangle \,\mathrm{dt} \\ &= \langle [1\, A(F)], H\rangle - \langle [0\, A(0F)], H\rangle = \langle A(F), H\rangle\,, \end{aligned}$$

and since H is arbitrary it follows $DW(F) = A(F)$. ■

Example 3.29
The tensorfunction $A(F) = F^T F$ is not the derivative of some potential, since $\langle \hat{H}, DA(F).H\rangle = \langle \hat{H}, F^T H + H^T F\rangle \neq \langle H, F^T \hat{H} + \hat{H}^T F\rangle$.

3.4 Invariance properties of convexity

Lemma 3.30 (Invariance of convexity)
Let $\hat{W} : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ be a convex mapping and let $A \in \mathbb{M}^{3\times 3}$. Then $W(F) := \hat{W}(A \cdot F)$ and $W(F) := \hat{W}(F \cdot A)$ are also convex.

Proof. Check directly the convexity condition for W by using the linearity of $F \mapsto A \cdot F$ and $F \mapsto F \cdot A$. ■

Lemma 3.31 (Invariance of strict convexity)
Let $\hat{W} : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ be a strictly convex function and let $A \in \mathrm{GL}(3, \mathbb{R})$. Then $W(F) := \hat{W}(A \cdot F)$ or $W(F) := \hat{W}(F \cdot A)$ are also strictly convex.

Proof. Since $A \in \mathrm{GL}(3, \mathbb{R})$ we have $F_1 A \neq F_2 A$ whenever $F_1 \neq F_2$. For $0 < \lambda < 1$ it hold therefore

$$\begin{aligned} W(\lambda F_1 + (1-\lambda)F_2) &= \hat{W}((\lambda F_1 + (1-\lambda)F_2)A) = \hat{W}(\lambda F_1 A + (1-\lambda)F_2 A) \\ &< \lambda \hat{W}(\lambda F_1 A) + (1-\lambda)\hat{W}(F_2 A) \\ &= \lambda W(F_1) + (1-\lambda)W(F_2)\,. \end{aligned}$$ ■

Lemma 3.32 (Invariance of strict ellipticity)
Let $\hat{W} : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ *be a smooth uniformly Legendre-Hadamard elliptic energy expression and let* $A \in \mathrm{GL}(3,\mathbb{R})$*. Then* $W(F) := \hat{W}(A \cdot F)$ *viz.* $W(F) := \hat{W}(F \cdot A)$ *is also uniformly Legendre-Hadamard elliptic.*

Proof. We have $D^2W(F).(H,H) = D^2\hat{W}(AF).(AH,AH)$. Taking $H = \xi \otimes \eta$ yields

$$\begin{aligned}
D^2W(F).(\xi\otimes\eta,\xi\otimes\eta) &= D^2\hat{W}(AF).(A(\xi\otimes\eta),A(\xi\otimes\eta)) \\
&= D^2\hat{W}(AF).(A.\xi\otimes\eta,A.\xi\otimes\eta) \\
&\geq c^+ \, \|A.\xi\|^2 \, \|\eta\|^2 \\
&\geq c^+ \, \lambda_{\min}(A^TA) \, \|\xi\|^2 \, \|\eta\|^2 \\
&= \hat{c}^+ \, \|\xi\|^2 \, \|\eta\|^2 \, . \qquad \blacksquare
\end{aligned}$$

Lemma 3.33 (Invariance of stability)
Let $\hat{W} : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ *be smooth and stable and* $A \in \mathrm{GL}(3,\mathbb{R})$*. Then* $W(F) := \hat{W}(A \cdot F)$ *viz.* $W(F) := \hat{W}(F \cdot A)$ *is also stable.*

Proof. It holds $D^2W(F).(H,H) = D^2\hat{W}(AF).(AH,AH) \geq c^+ \cdot \|AH\|^2$. Therefore $D^2W(F).(H,H) \geq c^+ \cdot \lambda_{\min}(A^TA)\|H\|^2$. $\blacksquare$

3.5 Invariance of polyconvexity

Let us define polyconvexity.

Definition 3.34 (Polyconvexity)
$F \mapsto W(x,F)$ *is polyconvex if and only if there exists a function* $P : \mathbb{R}^3 \times \mathbb{M}^{3\times 3} \times \mathbb{M}^{3\times 3} \times \mathbb{R} \mapsto \mathbb{R}$ *(in general non unique) such that*

$$W(x,F) = P(x,F,\operatorname{Adj} F,\det[F])$$

and the function $\mathbb{R}^{19} \mapsto \mathbb{R}$, $(X,Y,Z) \mapsto P(x,X,Y,Z)$ *is convex forall* $x \in \mathbb{R}^3$. $\blacksquare$

Lemma 3.35 (Invariance of polyconvexity)
Let $\hat{W} : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ *be a polyconvex function and let* $A \in \mathbb{M}^{3\times 3}$*. Then* $W(F) := \hat{W}(A \cdot F)$ *viz.* $W(F) := \hat{W}(F \cdot A)$ *is also polyconvex.*

Proof. Let $\hat{P} : \mathbb{M}^{3\times 3}\times\mathbb{M}^{3\times 3}\times\mathbb{R}$ be convex, i.e., for $\lambda \in (0,1)$ and arbitrary $(X_1,Y_1,z_1),(X_2,Y_2,z_2) \in \mathbb{M}^{3\times 3} \times \mathbb{M}^{3\times 3} \times \mathbb{R}$ it holds:

$$\hat{P}\left(\lambda \cdot \begin{pmatrix} X_1 \\ Y_1 \\ z_1 \end{pmatrix} + (1-\lambda) \cdot \begin{pmatrix} X_2 \\ Y_2 \\ z_2 \end{pmatrix}\right) \leq \lambda \cdot \hat{P}\left(\begin{pmatrix} X_1 \\ Y_1 \\ z_1 \end{pmatrix}\right) + (1-\lambda) \cdot \hat{P}\left(\begin{pmatrix} X_2 \\ Y_2 \\ z_2 \end{pmatrix}\right) . \quad (3.12)$$

Define $T : \mathbb{M}^{3\times 3} \mapsto \mathbb{M}^{3\times 3} \times \mathbb{M}^{3\times 3} \times \mathbb{R}$

$$T(F) := (F, \operatorname{Adj} F, \det[F]) \, .$$

This implies

$$\begin{aligned}T(A\cdot F) &= (A\cdot F, \operatorname{Adj}(A\cdot F), \det[(A\cdot F)])\\ &= (A\cdot F, \operatorname{Adj} F\cdot \operatorname{Adj} A, \det[A]\cdot\det[F])\,.\end{aligned}$$

$$\begin{aligned}\hat{W}(F) &= \hat{P}(T(F))\\ W(F) &= \hat{W}(A\cdot F) = \hat{P}(T(A\cdot F))\\ &= \hat{P}(A\cdot F, \operatorname{Adj} F\cdot\operatorname{Adj} A, \det[A]\cdot\det[F])\,.\end{aligned}$$

We define

$$P\left(\begin{pmatrix} X_1\\ Y_1\\ z_1\end{pmatrix}\right) = \hat{P}\left(\begin{pmatrix} A\cdot X_1\\ Y_1\cdot\operatorname{Adj} A\\ \det[A]\cdot z_1\end{pmatrix}\right)\,.$$

It remains to show that this P is itself again convex. For this we write

$$\begin{aligned}&P\left(\lambda\cdot\begin{pmatrix} X_1\\ Y_1\\ z_1\end{pmatrix} + (1-\lambda)\cdot\begin{pmatrix} X_2\\ Y_2\\ z_2\end{pmatrix}\right) =\\ &\qquad\hat{P}\left(\lambda\cdot\begin{pmatrix} A\cdot X_1\\ \operatorname{Adj} A\cdot Y_1\\ \det[A]\cdot z_1\end{pmatrix} + (1-\lambda)\cdot\begin{pmatrix} A\cdot X_2\\ \operatorname{Adj} A\cdot Y_2\\ \det[A]\cdot z_2\end{pmatrix}\right)\,.\end{aligned}$$

The convexity of $\hat{P}$ with the arguments $(A\cdot X_1, \operatorname{Adj} A\cdot Y_1, \det[A]\cdot z_1), (A\cdot X_2, \operatorname{Adj} A\cdot Y_2, \det[A]\cdot z_2)$ yields, however,

$$\begin{aligned}&\le \lambda\cdot\hat{P}\left(\begin{pmatrix} A\cdot X_1\\ \operatorname{Adj} A\cdot Y_1\\ \det[A]\cdot z_1\end{pmatrix}\right) + (1-\lambda)\cdot\hat{P}\left(\begin{pmatrix} A\cdot X_2\\ \operatorname{Adj} A\cdot Y_2\\ \det[A]\cdot z_2\end{pmatrix}\right)\\ &\qquad = \lambda P\left(\begin{pmatrix} X_1\\ Y_1\\ z_1\end{pmatrix}\right) + (1-\lambda)\cdot P\left(\begin{pmatrix} X_2\\ Y_2\\ z_2\end{pmatrix}\right)\,.\end{aligned}$$

Thus $W(F)$ is polyconvex since $W(F) = P(T(F))$ and P is convex. Similarly one may show the polyconvexity of $W(F) = \hat{W}(F\cdot A)$. ■

Corollary 3.36 (Invariance of polyconvexity under $F \mapsto A\cdot F\cdot B$)
Let $A, B \in \mathbb{M}^{3\times 3}$. Assume that $\hat{W}(F)$ is polyconvex. Then $W(F) := \hat{W}(A\cdot F\cdot B)$ is polyconvex in F.

Proof. Use the preceeding Lemma first from left and then from the right. ■

Corollary 3.37 (Inhomogeneous invariance of polyconvexity)
Let $\hat{W} : \Omega\times\mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ be a polyconvexe function, this means for fixed $x_0 \in \Omega$ the function $\hat{W}(x_0, \cdot)$ is polyconvex. Assume that $A(x) \in \mathbb{M}^{3\times 3}$. Then $W(x, F) := \hat{W}(x, A(x)\cdot F)$ viz. $W(x, F) := \hat{W}(x, F\cdot A(x))$ is also polyconvex.

Proof. For fixed $x_0 \in \Omega$ repeat the preceeding argument for $W(x_0, F)$. The matrix $A(x)$ only changes the inhomogeneity. ■

Corollary 3.38 (Invariance of strict polyconvexity)
Let $\hat{W} : \Omega \times \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ *be a strictly polyconvex function and assume that* $A(x) \in \mathrm{GL}(3, \mathbb{R})$. *Then* $W(x, F) := \hat{W}(x, A(x) \cdot F)$ *viz.* $W(x, F) := \hat{W}(x, F \cdot A(x))$ *is also strictly polyconvex.*

Corollary 3.39
Let $F = F_e \cdot F_p$ *with* $\det[F]_p = 1$. *Then the function*

$$W(x, F) := \hat{W}(F_e)$$

is strictly polyconvex if $\hat{W} : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ *is itself strictly polyconvex. Therefore, it suffices to define a strictly polyconvex function in terms of* F_e.

Remark 3.40
The same results carry over to weak lower semicontinuity and quasiconvexity.

Lemma 3.41 (Additive polyconvex functions)
Let $P_1, P_2 : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ *and* $P_3 : \mathbb{R} \mapsto \mathbb{R}$ *be convex, respectively. Then*

$$\begin{aligned} &P : \mathbb{M}^{3\times 3} \times \mathbb{M}^{3\times 3} \times \mathbb{R} \mapsto \mathbb{R}\,, \\ &P(X, Y, z) := P_1(X) + P_2(Y) + P_3(z) \end{aligned}$$

is convex.

Proof. Check the convexity conditions directly. ■

Corollary 3.42
Functions of the type $W(F) = P_1(F) + P_2(\mathrm{Adj}\, F) + P_3(\det[F])$ *with* P_i *convex, are polyconvex.*

There are many characterisations of polyconvexity. Let us mention just these ones for special functions.

Lemma 3.43 (Polyconvexity for special functions)
Let $h : \mathbb{R}_+ \mapsto \mathbb{R}$ *be smooth. Then*

$$\begin{aligned} &W_1(F) = h(\|F\|)\,, \\ &W_2(F) = \|F\|^\alpha + h(\det[F]), \quad 1 \le \alpha \le 6\,, \end{aligned}$$

are polyconvex if and only if h *is convex.*

Proof. In case of W_1 see [**?**], in case of W_2 see [**?**]. ■

Example 3.44 (Polyconvex functions I)
Let $A, F \in \mathbb{M}^{3\times 3}$ *und* $\det[F] \geq c^+ > 0$. *Then the following list presents examples of simple polyconvex functions* $H \mapsto W(H)$.

$$\begin{aligned}
W(H) &= \langle A, H\rangle^2 \,, \\
W(H) &= \langle \operatorname{Adj} H, A\rangle \,, \\
W(H) &= \langle \operatorname{Adj} H, A\rangle^2 \,, \\
W(H) &= \| \operatorname{Adj} H\|^2 \,, \\
W(H) &= \|D(\operatorname{Adj} F).H\|^2 \,, \\
W(H) &= -\langle \operatorname{Adj} H, A\rangle \,, \\
W(H) &= \langle (HF^{-1})^T, HF^{-1}\rangle \,, \\
&= \langle F^{-T}, H\rangle^2 - \frac{2}{\det[F]} \cdot \langle \operatorname{Adj} H, F^T\rangle \,, \\
W(H) &= \|H^T H\|^2 + \| \operatorname{Adj}(H^T H)\|^2 + \det[H] - \ln \det[H] \,, \\
W(H) &= \frac{\|H\|^2}{\det[H]^{\frac{2}{3}}} \,, \\
W(H) &= \|H\|^2 + \det[H] \cdot \ln \det[H] \,.
\end{aligned}$$

Let us also recall the following result in convex analysis.

Lemma 3.45
A differentiable function $W : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ *is convex if and only if*

$$\begin{aligned}
\forall F \in \mathbb{M}^{3\times 3}\ \forall H \in \mathbb{M}^{3\times 3} : &\quad W(F+H) \geq W(F) + DW(F).H \,, \\
\forall F \in \mathbb{M}^{3\times 3}\ \forall E \in \mathbb{M}^{3\times 3} : &\quad W(E) \geq W(F) + DW(F).(E-F) \,.
\end{aligned}$$

This statement can be used to check convexity properties if the function W *is not twice differentiable.*

Proof. Well known characterization of convexity. ■
For polyconvex functions we may conclude thus

Lemma 3.46
Let $W : M^{3\times 3} \mapsto \mathbb{R}$ *be a smooth polyconvex function and let* $P : \mathbb{M}^{3\times 3} \times \mathbb{M}^{3\times 3} \times \mathbb{R} \mapsto \mathbb{R}$ *be the corresponding convex function. Then*

$$\begin{aligned}
&\forall E, F \in \mathbb{M}^{3\times 3} \\
W(E) &\geq W(F) + D_1 P(F, \operatorname{Adj} F, \det[F]).\,[E-F] \\
&+ D_2 P(F, \operatorname{Adj} F, \det[F]).\,[D(\operatorname{Adj} F).(E-F) + \operatorname{Adj}(E-F)] \\
&+ D_3 P(F, \operatorname{Adj} F, \det[F]).\,\left[\langle E-F, \operatorname{Adj} F^T\rangle + \langle \operatorname{Adj}(E-F), F^T\rangle + \det[(E-F)]\right]
\end{aligned}$$

Proof.

$$\begin{aligned}
W(E) &= P(E, \operatorname{Adj} E, \det[E]) \\
&\geq P(F, \operatorname{Adj} F, \det[F]) + DP(F, \operatorname{Adj} F, \det[F]).\left[\begin{pmatrix} E \\ \operatorname{Adj} E \\ \det[E] \end{pmatrix} - \begin{pmatrix} F \\ \operatorname{Adj} F \\ \det[F] \end{pmatrix}\right] \\
&= W(F) + D_1P(F, \operatorname{Adj} F, \det[F]).\,[E - F] \\
&+ D_2P(F, \operatorname{Adj} F, \det[F]).\,[\operatorname{Adj} E - \operatorname{Adj} F] \\
&+ D_3P(F, \operatorname{Adj} F, \det[F]).\,[\det[E] - \det[F]]\ .
\end{aligned}$$

Insert now the expansion for the adjugate and the determinant. Recall that, e.g., $\operatorname{Adj} E = \operatorname{Adj} F + D(\operatorname{Adj} F).(E - F) + \operatorname{Adj}\,(E - F)$. Thus

$$\begin{aligned}
W(E) \geq{}& W(F) + D_1P(F, \operatorname{Adj} F, \det[F]).\,[E - F] \\
&+ D_2P(F, \operatorname{Adj} F, \det[F]).\,[D(\operatorname{Adj} F).(E - F) + \operatorname{Adj}\,(E - F)] \\
&+ D_3P(F, \operatorname{Adj} F, \det[F]).\,\left[\langle E - F, \operatorname{Adj} F^T\rangle + \langle \operatorname{Adj}\,(E - F), F^T\rangle + \det[(E - F)]\right]\ .
\end{aligned}$$

■

Corollary 3.47
For additive polyconvex functions we obtain therefore

$$\begin{aligned}
&\forall E, F \in \mathbb{M}^{3\times 3} \\
W(E) \geq{}& W(F) + DP_1(F).\,[E - F] \\
&+ DP_2(\operatorname{Adj} F).\,[D(\operatorname{Adj} F).(E - F) + \operatorname{Adj}\,(E - F)] \\
&+ DP_3(\det[F]).\,\left[\langle E - F, \operatorname{Adj} F^T\rangle + \langle \operatorname{Adj}\,(E - F), F^T\rangle + \det[(E - F)]\right]\ .
\end{aligned}$$

■

For strictly polyconvex functions the statements carry over with $>$ instead of $\geq$.

Lemma 3.48 (Polyconvexity at the identity)
Let W be strictly polyconvex and take $u, \phi : \mathbb{R}^3 \mapsto \mathbb{R}^3$, $u(x) = x$, $\phi(x)_{\partial\Omega} = x$. Then for $F = \nabla u = \mathbb{1}$ and all $\nabla\phi$:

$$\begin{aligned}
W(\nabla\phi) >{}& W(\mathbb{1}) + \langle D_1P(\mathbb{1}, \mathbb{1}, 1), \nabla(\phi - x)\rangle \\
&+ \langle D_2P(\mathbb{1}, \mathbb{1}, 1), [\mathbb{1} \cdot \langle \nabla(\phi - x), \mathbb{1}\rangle - \nabla(\phi - x) + \operatorname{Adj}\,(\nabla(\phi - x))]\rangle \\
&+ D_3P(\mathbb{1}, \mathbb{1}, 1) \cdot [\langle \nabla(\phi - x), \mathbb{1}\rangle + \langle \operatorname{Adj} \nabla(\phi - x), \mathbb{1}\rangle + \det[\nabla(\phi - x)]]\ , \\
W(\mathbb{1} + H) >{}& W(\mathbb{1}) + \langle D_1P(\mathbb{1}, \mathbb{1}, 1), H\rangle + \langle D_2P(\mathbb{1}, \mathbb{1}, 1), [\mathbb{1} \cdot \langle H, \mathbb{1}\rangle - H + \operatorname{Adj} H]\rangle \\
&+ D_3P(\mathbb{1}, \mathbb{1}, 1) \cdot [\langle H, \mathbb{1}\rangle + \langle \operatorname{Adj} H, \mathbb{1}\rangle + \det[H]]\ , \\
W(\mathbb{1} + H) >{}& W(\mathbb{1}) + \langle D_1P(\mathbb{1}, \mathbb{1}, 1), H\rangle + \langle D_2P(\mathbb{1}, \mathbb{1}, 1), \mathbb{1} \cdot \langle H, \mathbb{1}\rangle - H\rangle + D_3P(\mathbb{1}, \mathbb{1}, 1) \cdot [\langle H, \mathbb{1}\rangle] \\
&+ \langle D_2P(\mathbb{1}, \mathbb{1}, 1), \operatorname{Adj} H\rangle + D_3P(\mathbb{1}, \mathbb{1}, 1) \cdot [\langle \operatorname{Adj} H, \mathbb{1}\rangle] \\
&+ D_3P(\mathbb{1}, \mathbb{1}, 1) \cdot \det[H]\,.
\end{aligned}$$

Proof. This is a direct conclusion based on Lemma **??**. ■

3.6 Infinitesimal convexity, polyconvexity, quasiconvexity and rank-one convexity

Let us assume that W is twice Frechet-differentiable. Then it admits the expansion

$$W(F+H) = W(F) + DW(F).H + \frac{1}{2}D^2W(F).(H,H) + o(\|H\|^2)\,.$$

Specifying $F = 1\!\!1$ and assuming that the reference configuration is stress-free, this yields

$$W(1\!\!1+H) = W(1\!\!1) + DW(1\!\!1).H + \frac{1}{2}D^2W(1\!\!1).(H,H) + o(\|H\|^2)\,. \tag{3.13}$$

On the other hand, quasiconvexity at the identity means that forall $u \in C_0^\infty(\Omega,\mathbb{R}^3)$

$$\int_\Omega W(1\!\!1+\nabla u)\ \mathrm{dx} \geq \int_\Omega W(1\!\!1)\ \mathrm{dx}\,. \tag{3.14}$$

Setting $H = \nabla u$, integrating (**??**) and using that

$$\int_\Omega DW(1\!\!1).\nabla u\ \mathrm{dx} = DW(1\!\!1).\int_\Omega \nabla u\,\mathrm{dx} = DW(1\!\!1).(x\otimes u_{\partial\Omega}) = 0$$

yields

$$\int_\Omega W(1\!\!1+\nabla u)\ \mathrm{dx} = \int_\Omega W(1\!\!1)\ \mathrm{dx} + \frac{1}{2}\int_\Omega D^2W(1\!\!1).(\nabla u,\nabla u)\,\mathrm{dx} + o(\|\nabla u\|^2)\,.$$

Combining this with (**??**) yields

$$\int_\Omega W(1\!\!1)\ \mathrm{dx} \leq \int_\Omega W(1\!\!1+\nabla u)\ \mathrm{dx} = \int_\Omega W(1\!\!1)\ \mathrm{dx} + \frac{1}{2}\int_\Omega D^2W(1\!\!1).(\nabla u,\nabla u)\,\mathrm{dx} + o(\|\nabla u\|^2)\,,$$

implying

$$0 \leq \frac{1}{2}\int_\Omega D^2W(1\!\!1).(\nabla u,\nabla u)\,\mathrm{dx} + o(\|\nabla u\|^2)\,.$$

We call W **infinitesimally quasiconvex** if [**?**, p.325]

$$0 \leq \frac{1}{2}\int_\Omega D^2W(1\!\!1).(\nabla u,\nabla u)\,\mathrm{dx}\,,$$

which is a nonlocal condition. Combining the expansion (**??**) with Lemma **??** and equating like powers gives as necessary condition for polyconvexity at the identity, the **infinitesimal polyconvexity** condition

$$\frac{1}{2}D^2W(1\!\!1).(H,H) \geq \langle D_2P(1\!\!1,1\!\!1,1),\operatorname{Adj}H\rangle + D_3P(1\!\!1,1\!\!1,1)\,\langle \operatorname{Adj}H, 1\!\!1\rangle\,. \tag{3.15}$$

Using the relation

$$X^2 - \operatorname{tr}[X]\,X = \operatorname{Cof}X - \operatorname{tr}[\operatorname{Cof}X]\,1\!\!1 \quad \text{(use Cayley-Hamilton)}\,,$$

and since $D_2P(\mathbb{1},\mathbb{1},1)$, $D_3P(\mathbb{1},\mathbb{1},1)$ may be assumed arbitrary, allows to reformulate (**??**) into the equivalent requirement

$$\exists\, Y \in \mathbb{M}^{3\times 3}\ \forall\, H \in \mathbb{M}^{3\times}: \quad \frac{1}{2} D^2W(\mathbb{1}).(H,H) \geq \langle Y, H^2 - \operatorname{tr}[H]\, H\rangle\,. \tag{3.16}$$

Infinitesimal convexity is simply convexity at the identity, thus

$$\forall\, H \in \mathbb{M}^{3\times 3}: \quad D^2W(\mathbb{1}).(H,H) \geq 0\,.$$

Combining (**??**) with the condition that W should be rank-one convex at $F=\mathbb{1}$, i.e. for all $\xi,\eta \in \mathbb{R}^3$

$$W(\mathbb{1}+\xi\otimes\eta) \geq W(\mathbb{1}) + DW(\mathbb{1}).(\xi\otimes\eta)\,,$$

gives

$$\begin{aligned} W(\mathbb{1}) + DW(\mathbb{1}).(\xi\otimes\eta) &\leq W(\mathbb{1}+\xi\otimes\eta) \\ &= W(\mathbb{1}) + DW(\mathbb{1}).(\xi\otimes\eta) \\ &\quad + \frac{1}{2}D^2W(\mathbb{1}).(\xi\otimes\eta,\xi\otimes\eta) + o(\|\xi\|^2\|\eta\|^2)\,, \end{aligned}$$

which implies the local **infinitesimal rank-one convexity** condition

$$\frac{1}{2}D^2W(\mathbb{1}).(\xi\otimes\eta,\xi\otimes\eta) \geq 0\,. \tag{3.17}$$

Open question: is infinitesimal polyconvexity equivalent to infinitesimal rank-one convexity? Essentially, the question is: are there infinitesimal quasiconvex functions that are not infinitesimal polyconvex? It is known that infinitesimal rank-one convexity is equivalent to infinitesimal quasiconvexity, see e.g. [**?**]. Compare also with [**?**, p.128] where the case of quadratic functions W is treated. It is shown there, that there are quadratic functions $Q(F,F)$ which are rank-one convex, but not polyconvex. Taking $W(F)=Q(F,F)$ shows that infinitesimal rank-one convexity does not imply infinitesimal polyconvexity. However, the counterexample is not frame-indifferent.

4 Weak convergence

In this part we would like to motivate the importance of weak convergence and its relation with the minors of a matrix.

Let us start by recalling the notion of weak convergence. We say that a sequence of integrable functions $\xi_k \in L^2$ converges weakly whenever for all fixed $\Psi \in L^2(\Omega)$ it holds

$$\int_\Omega \xi_k(x)\,\Psi(x)\,\mathrm{dV} \to \int_\Omega \xi(x)\,\Psi(x)\,\mathrm{dV}.$$

In this case we write

$$\xi_k \rightharpoonup \xi \quad \text{in } L^2(\Omega)\,.$$

The weak limit, if it exists, is unique. If we choose the testfunction $\Psi \equiv 1$ it holds

$$\int_\Omega \xi_k(x)\, 1\, \mathrm{dV} \to \int_\Omega \xi(x)\, 1\, \mathrm{dV}.$$

Thus weak convergence implies convergence of averages and similarly of all higher moments of ξ_k. However, weak convergence of ξ_k does not imply that ξ_k converges strongly. To see this, consider the (oscillating) sequence

$$\xi_k(x) = \sin(k\, x)\,,$$

on the interval $\Omega = [-\pi, \pi]$. We will show that $\xi_k \rightharpoonup 0$ but that $\forall\, k \in N : \|\xi_k\|_{L^2} \geq c^+ > 0$, excluding strong convergence.

Proof. Take some smooth but otherwise arbitrary $\Psi \in L^2(-\pi, \pi)$ and consider

$$\begin{aligned}\int_{-\pi}^{\pi} \sin(kx)\, \Psi(x)\, \mathrm{dx} &= \int_{-\pi}^{\pi} \frac{\mathrm{d}}{\mathrm{dx}}[-\cos(kx)\frac{1}{k}]\, \Psi(x)\, \mathrm{dx} \\ &= [-\cos(kx)\frac{1}{k}]\, \Psi(x)]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} [-\cos(kx)\frac{1}{k}] \frac{\mathrm{d}}{\mathrm{dx}} \Psi(x)\, \mathrm{dx}\,.\end{aligned}$$

For $k \to \infty$ this converges to zero. However

$$\int_{-\pi}^{\pi} \sin^2(kx)\, \mathrm{dx} = [\frac{1}{2}x]_{-\pi}^{\pi} - [\frac{1}{4k}\sin(2kx)]_{-\pi}^{\pi} = \pi > 0\,.$$

What is missing in obtaining strong convergence? One can show that if in addition to weak convergence, the norms of the sequence converge, i.e.,

$$\|\xi_k\|_{L^2} \to \|\xi\|_{L^2}\,,$$

then we obtain strong convergence $\xi_k \to \xi \in L^2$.

Proof. Assume that $\xi_k \rightharpoonup \xi$ and $\|\xi_k\|_{L^2} \to \|\xi\|_{L^2}$. We write

$$\|\xi_k - \xi\|^2 = \|\xi_k\|^2 - 2\langle \xi_k, \xi\rangle + \|\xi\|^2 \to \|\xi\|^2 - 2\langle \xi, \xi\rangle + \|\xi\|^2 = 0\,,$$

since, in the mixed term, ξ is now treated as a fixed testfunction (assuming the role of Ψ).

4.1 What type of nonlinear functions preserve weak convergence?

This is the question: assume that $\xi_k \rightharpoonup \xi \in L^2(\Omega)$, weakly. For what type of nonlinear functions f is it true that

$$f(\xi_k) \rightharpoonup f(\xi) \quad \text{in } L^2(\Omega)\,.$$

The answer is surprisingly simple: the only functions preserving weak convergence are the affine linear functions $f(x) = a + b\,x$. Thus, weak convergence methods would not be of much help in treating nonlinear problems. However, many sequences arising in applications of continuum mechanics are not arbitrary sequences, but the sequences have the additional, decisive property of being sequences of gradients, i.e. we speak of

$$\xi_k = \nabla\varphi_k \rightharpoonup \xi = \nabla\varphi\,,$$

for a sequence of functions $\varphi_k \to \varphi$. Within this extra structure, there are indeed nonlinear functions, which are preserving weak convergence! Subsequently, we will see that weakly convergent functions are essentially so called null-Lagrangeans.

4.2 How to get weak convergence?

In many cases one can show that some sequence of functions $\xi_k \in L^p(\Omega)$ satisfies a uniform bound of the type

$$\|\xi_k\|_{L^p(\Omega)} = \int_\Omega \|\xi_k(x)\|^p \,\mathrm{dV} \le K\,.$$

If $p > 1$ one can extract a subsequence which converges weakly in $L^p(\Omega)$, i.e.

$$\xi_{k_j} \rightharpoonup \xi \quad \text{in}\, L^p(\Omega)\,, \quad j \to \infty\,.$$

4.3 Null-Lagrangean and weak continuity

Integrands W for which the integral $\int_\Omega W(\nabla u)$ only depends on the boundary values of u are called null Lagrangians, since the Euler-Lagrange equations are automatically satisfied for all functions u. Affine combinations of minors are the only null-Lagrangians and the only functions that preserve weak continuity of sequences of gradients.

Theorem 4.1 (Null-Lagrangean)
Let $\Omega \subset \mathbb{R}^3$ be an open, bounded set and let $\phi \in C_0^\infty$. For constant $A \in \mathbb{M}^{3\times 3}$ define $T(A) = (A, \operatorname{Adj} A, \det[A])$. Then

$$\int_\Omega T(A + \nabla\phi(x))\ \mathrm{dx} = T(A)\cdot|\Omega| \quad \textit{componentwise}\,,$$

especially for $A = 0$

$$\int_\Omega T(\nabla\phi(x))\ \mathrm{dx} = 0\,, \quad \textit{componentwise}\,,$$

this means

$$\begin{aligned}\int_\Omega \nabla\phi(x)\ \mathrm{dx} &= 0\,, \quad \textit{componentwise}\\ \int_\Omega \operatorname{Adj}\nabla\phi(x)\ \mathrm{dx} &= 0\,, \quad \textit{componentwise}\\ \int_\Omega \det[\nabla\phi(x)]\ \mathrm{dx} &= 0\,.\end{aligned}$$

Proof. See, e.g., [**?**], page 193, Theorem 3.2. Let us nevertheless continue. Expanding $T(A+H)$ shows

$$\begin{aligned}
T(A+H) &= (A+H, \operatorname{Adj} A + H, \det[A+H]) \\
&= (A+H, \operatorname{Adj} A + D \operatorname{Adj} A.H + \operatorname{Adj} H, \det[A] + \langle \operatorname{Adj} A, H\rangle + \langle A, \operatorname{Adj} H\rangle + \det[H]) \\
&= (A, \operatorname{Adj} A, \det[A]) + (H, D \operatorname{Adj} A.H, \langle \operatorname{Adj} A, H\rangle) \\
&\quad + (0, \operatorname{Adj} H, \langle A, \operatorname{Adj} H\rangle) + (0,0,\det[H])\,.
\end{aligned}$$

Setting $H = \nabla\phi$ and integrating over Ω shows (since A is constant) the statement once we can show that

$$\begin{aligned}
\int_\Omega \nabla\phi(x)\ \mathrm{dx} &= 0\,, \quad \text{componentwise} \\
\int_\Omega \operatorname{Adj} \nabla\phi(x)\ \mathrm{dx} &= 0\,, \quad \text{componentwise} \\
\int_\Omega \det[\nabla\phi(x)]\ \mathrm{dx} &= 0\,.
\end{aligned}$$

For the first term it is easily seen, by Gauss theorem (partial integration), that

$$\int_\Omega \nabla\phi(x)\ \mathrm{dx} = \int_{\partial\Omega} \phi(x)\otimes \vec{n}\ \mathrm{dS} = 0\,.$$

The other properties hold, since determinants and adjugates of gradients may be written in divergence format. For example consider for $\phi \in C_0^\infty(\mathbb{R}^2, \mathbb{R}^2)$

$$\begin{aligned}
\phi(x_1,x_2) &= \begin{pmatrix} \phi^1(x_1,x_2) \\ \phi^2(x_1,x_2) \end{pmatrix}\,, \quad \nabla\phi(x_1,x_2) = \begin{pmatrix} \phi^1_{x_1} & \phi^1_{x_2} \\ \phi^2_{x_1} & \phi^2_{x_2} \end{pmatrix}\,, \\
\det[\nabla\phi] &= \phi^1_{x_1}\phi^2_{x_2} - \phi^1_{x_2}\phi^2_{x_1}\,, \\
&= \phi^1_{x_1}\cdot{\phi^2}_{x_2} + \phi^1\cdot\phi^2_{x_2,x_1} - \phi^1_{x_2}\cdot{\phi^2}_{x_1} - \phi^1\cdot\phi^2_{x_1,x_2} \\
&= \left[\phi^1\cdot{\phi^2}_{x_2}\right]_{x_1} + \left[-\phi^1\cdot{\phi^2}_{x_1}\right]_{x_2} \\
&= \operatorname{Div}\begin{pmatrix} \phi^1\cdot{\phi^2}_{x_2} \\ -\phi^1\cdot{\phi^2}_{x_1} \end{pmatrix}\,.
\end{aligned}$$

Corollary 4.2 (Simple further null-Lagrangeans)
Let $C \in \mathbb{M}^{3\times 3}$, $C_{ij} = const.$ and $c \in \mathbb{R}$, $c = const.$ and take $\phi \in C_0^\infty(\Omega, \mathbb{R}^3)$. Then

$$\begin{aligned}
\int_\Omega \langle C, \nabla\phi(x)\rangle\ \mathrm{dx} &= \langle C, \int_\Omega \nabla\phi(x)\ \mathrm{dx}\rangle = 0\,, \\
\int_\Omega \langle C, \operatorname{Adj}\nabla\phi(x)\rangle\ \mathrm{dx} &= \langle C, \int_\Omega \operatorname{Adj}\nabla\phi(x)\ \mathrm{dx}\rangle = 0\,, \\
\int_\Omega c\cdot\det[\nabla\phi(x)]\ \mathrm{dx} &= 0\,.
\end{aligned}$$

One of the cornerstones in applying convexity methods to problems in nonlinear elasticity is that weak convergence of the gradient implies weak convergence of the (nonlinear terms) adjugate and determinants.

Theorem 4.3 (Weak continuity of the gradient)
Let $\Omega \subset \mathbb{R}^3$ *be open and bounded,* $p \geq 2$, $\frac{1}{p} + \frac{1}{q} \leq 1$, $(p \geq \frac{q}{q-1})$ *and* $r \geq 1$. *Suppose*

$$\begin{aligned} \phi_k &\rightharpoonup \phi \quad \textit{in } W^{1,p}(\Omega)\,, \\ \operatorname{Adj} \nabla\phi_k &\rightharpoonup H \quad \textit{in } L^q(\Omega)\,, \\ \det[\nabla\phi_k] &\rightharpoonup \delta \quad \textit{in } L^r(\Omega)\,. \end{aligned}$$

Then

$$H = \operatorname{Adj} \nabla\phi\,, \qquad \delta = \det[\nabla\phi]\,.$$

Proof. By the compact embedding $W^{1,2}(\Omega) \subset L^2(\Omega)$ we can immediately assume that $\phi_k \to \phi \in L^2(\Omega)$ strongly. Since smooth functions are dense in $W^{1,2}$ we can consider the smooth case. Let us look at a generic case. The general case follows easily. We consider the weak continuity of the determinant in the planar case. Thus consider a sequence of functions $\phi_k \in C^\infty(\mathbb{R}^2, \mathbb{R}^2)$ such that

$$\phi_k \to \phi \in L^2(\Omega)\,,$$

and the gradients converge weakly, i.e.,

$$\nabla\phi_k \rightharpoonup \nabla\phi \quad \Leftrightarrow \quad \forall\, \Psi \in C_0^\infty(\Omega, \mathbb{M}^{2\times 2}) : \int_\Omega \nabla\phi_k \; \Psi(x,y)\, \mathrm{dV} \to \int_\Omega \nabla\phi \; \Psi(x,y)\, \mathrm{dV}\,.$$

We want to show that the determinant of the gradients converges weakly as well, i.e.,

$$\forall\, \tilde\Psi \in C_0^\infty(\Omega, \mathbb{R}) : \int_\Omega \det[\nabla\phi_k]\, \tilde\Psi(x,y)\, \mathrm{dV} \to \int_\Omega \det[\nabla\phi]\, \tilde\Psi(x,y)\, \mathrm{dV}\,,\ k \to \infty\,,$$

despite the fact that the determinant is a nonlinear function. Fix $\tilde\Psi \in C_0^\infty(\Omega, \mathbb{R})$ and compute (here, subscripts denote partial differentiation)

$$\begin{aligned} \int_\Omega \phi^1_{k,x}\, [\phi^2_{k,y} \tilde\Psi]\, \mathrm{dV} &= -\int_\Omega \phi^1_k\, [\phi^2_{k,y} \tilde\Psi]_x\, \mathrm{dV} = -\int_\Omega \phi^1_k\, [\phi^2_{k,yx} \tilde\Psi + \phi^2_{k,y}\, \tilde\Psi_x]\, \mathrm{dV}\,, \\ \int_\Omega \phi^1_{k,y}\, [\phi^2_{k,x} \tilde\Psi]\, \mathrm{dV} &= +\int_\Omega \phi^1_k\, [\phi^2_{k,x} \tilde\Psi]_y\, \mathrm{dV} = \int_\Omega \phi^1_k\, [\phi^2_{k,xy} \tilde\Psi + \phi^2_{k,x}\, \tilde\Psi_y]\, \mathrm{dV}\,, \end{aligned}$$

since $\tilde\Psi$ has vanishing boundary data. Adding up and using (Theorem of Schwarz on the symmetry of second partial derivatives) that $\phi^2_{k,xy} = \phi^2_{k,yx}$ shows

$$\int_\Omega \det[\nabla\phi_k]\, \tilde\Psi(x,y)\, \mathrm{dV} = \int_\Omega -\phi^1_k\, \phi^2_{k,x}\, \tilde\Psi_x + \phi^2_{k,x}\, \phi^1_k\, \tilde\Psi_y\, \mathrm{dV}\,.$$

Strong convergence of ϕ_k and weak convergence of $\nabla\phi_k$ implies, however,

$$\int_\Omega -\phi^1_k\, \phi^2_{k,x}\, \tilde\Psi_x + \phi^2_{k,x}\, \phi^1_k\, \tilde\Psi_y\, \mathrm{dV} \to \int_\Omega -\phi^1\, \phi^2_x\, \tilde\Psi_x + \phi^2_x\, \phi^1\, \tilde\Psi_y\, \mathrm{dV}\,.$$

Reversing the partial integration shows that

$$\int_\Omega -\phi^1\,\phi^2_x\,\tilde{\Psi}_x + \phi^2_x\,\phi^1\,\tilde{\Psi}_y\,\mathrm{dV} = \int_\Omega \det[\nabla\phi]\,\tilde{\Psi}(x,y)\,\mathrm{dV}\,.$$

This shows the claim for smooth functions. The result follows by density. In the general case note that the entries of the adjugate are determinants themselves and compare with [**?**, p.366]. ■

Example 4.4
With the same assumptions it holds

$$\begin{aligned} &u_\nu \rightharpoonup u \quad \textit{in } W^{1,p}(\Omega) \Longrightarrow \\ &\qquad \int_\Omega \det[\nabla u_\nu]\cdot 1\ \mathrm{dx} \to \int_\Omega \det[\nabla u]\cdot 1\ \mathrm{dx}\,, \\ &\qquad \int_\Omega \mathrm{Adj}\,\nabla u_\nu\ \mathrm{dx} \to \int_\Omega \mathrm{Adj}\,\nabla u\ \mathrm{dx}\,, \quad \textit{componentwise}\,. \end{aligned}$$

Definition 4.5 (Caratheodory-functions)
We say that the function $G : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ is a ***Caratheodory-function****, if*

$$\begin{aligned} &\textit{for every pair } (u_0, F_0) \in \mathbb{R}^3 \times \mathbb{M}^{3\times 3} \textit{ is} \quad x \mapsto G(x, u_0, F_0) \textit{ measurable} \\ &\textit{for almost all } x_0 \in \mathbb{R}^3 \textit{ the function} \quad (u, F) \mapsto G(x_0, u, F) \textit{ is continuous.} \end{aligned}$$

Theorem 4.6 (Nemitsky-operators)
Let $\Omega \subset \mathbb{R}^3$ be a bounded open set and let $G : \Omega \subset \mathbb{R}^3 \times \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ be a Caratheodory-function. Moreover, assume the growth condition

$$\forall\,(x,F) \in \Omega \subset \mathbb{R}^3 \times \mathbb{M}^{3\times 3} \quad G(x,F) \le g(x) + C^+ \cdot \|F\|^{\frac{p}{q}}$$

for a function $g \in L^q(\Omega)$. Then the operator

$$H : L^p(\Omega) \mapsto L^q(\Omega), \quad H(F).(x) = G(x, F(x))$$

is continuous.

Proof. This is well known and shown for example in [**?**]. ■

Remark 4.7
This operator is already continuous, if only $L^p(\Omega)$ is mapped into $L^q(\Omega)$.

Example 4.8 (No growth condition)
The following functions do not satisfy a growth condition:

$$\begin{aligned} F &\mapsto -\ln\det[F]\,, \\ F &\mapsto \frac{1}{\|F\|}\,, \\ F &\mapsto \frac{1}{\det[F]}\,. \end{aligned}$$

In all three cases the energy may be bounded while $\|F\|$ can be large. One cannot expect continuity results between L^p-spaces.

5 A glimpse on the direct method of the calculus of variations

5.1 Weak lower-semicontinuity

Theorem 5.1 (Weak lower semicontinuity and convexity)
Let $G : \mathbb{R}^3 \times \mathbb{R}^N \mapsto \mathbb{R}$ be a Caratheodory-function. Assume that $G(x,T) \geq \beta$ and that $T \mapsto G(x,T)$ is **convex**. *Then*

$$\begin{aligned} &T_n \rightharpoonup T \quad \text{in } L^1 \Rightarrow \\ &\int_\Omega G(x, T(x))\,\mathrm{dx} \leq \liminf_{n\to\infty} \int_\Omega G(x, T_n)\,\mathrm{dx}\,. \end{aligned}$$

Proof. Standard result of convex analysis. See, e.g., [?], Theorem 7.3.1, p. 352. Let us remark that in the presented **convex case one does not need any growth condition** on G in order to conclude the weak lower-semicontinuity. ■

Theorem 5.2 (Weak lower-semicontinuity in $W^{1,p}(\Omega)$ and polyconvexity)
Let $G : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{M}^{3\times3} \mapsto \bar{\mathbb{R}}$ a Caratheodory-function. Let $F \mapsto G(x,u,F)$ be polyconvex for all $(x,u) \in \mathbb{R}^3 \times \mathbb{R}^3$ and let $\Omega \subset \mathbb{R}^3$ be an open bounded set and assume that $u_\nu(x) \to u(x)$ almost everywhere and $\nabla u_\nu \rightharpoonup \nabla u \in L^1(\Omega)$. Moreover, assume there is a function $\Psi \in L^1(\Omega)$ such that

$$G(x, u_\nu(x), \nabla u_\nu(x)) \geq \Psi(x), \qquad G(x, u(x), \nabla u(x)) \geq \Psi(x)\,.$$

Then

$$\int_\Omega G(x, u(x), \nabla u(x))\ \mathrm{dx} \leq \liminf_{\nu\to\infty} \int_\Omega G(x, u_\nu(x), \nabla u_\nu(x))\,\mathrm{dx}\,.$$

Proof. [?],Theorem 5.4, Seite 161. Remark that still no growth condition is needed. ■

Theorem 5.3 (Weak lower-semicontinuity in $W^{1,p}(\Omega)$ and quasiconvexity)
Let $1 \leq p < \infty$ and $G : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{M}^{3\times3} \mapsto \mathbb{R}$ is a Caratheodory-function and assume that $F \mapsto G(x,u,F)$ is **quasiconvex** *for all $(x,u) \in \mathbb{R}^3 \times \mathbb{R}^3$ and $\Omega \subset \mathbb{R}^3$ is an open bounded set. Moreover assume the* **growth condition**

$$\forall\ (x,u,F): \quad 0 \leq G(x,u,F) \leq g(x) + C^+(1 + \|u\|^p + \|F\|^p)\,,$$

for some non-negative locally integrable function $g : \mathbb{R}^3 \mapsto \mathbb{R}$. Then

$$I(u) := \int_\Omega G(x, u, \nabla u)\,\mathrm{dx}$$

is weakly lower-semicontinuous in $W^{1,p}(\Omega)$.

Proof. [?], Theorem 2.4. ■

Remark 5.4
In applications to physically realistic material behaviour, which does not satisfy a growth condition, the quasiconvexity condition could not be used up to the present.

In order to put the notions of weak-convergence and convexity into perspective we motivate their usage in the following.

5.2 The direct methods of the calculus of variations in a finite-dimensional nutshell

Let us assume that we have a function $W : \mathbb{R} \mapsto \mathbb{R}$ and we would like to find a global minimizer of this function. The first necessary assumption is that W, the **energy is bounded below**, i.e.,

$$\forall\, x \in \mathbb{R} : \quad W(x) \geq -K\,, \tag{5.18}$$

since otherwise, there does not exist a global minimizer (one can then lower the energy beyond all bounds).

Now we consider **infimizing sequences**: these are sequences of points x_k, whose energy levels approach the lowest possible energy level, i.e.,

$$W(x_k) \to \inf_{a\in\mathbb{R}} W(a)\,.$$

From (**??**) we know that the number $\inf_{a\in\mathbb{R}} W(a)$ exists. Our hope is that the infimizing sequence will somehow approximate the global minimizer, more precisely, if x_k converges to some limit point x, then we hope that this limit is indeed a global minimizer. Assume for the moment that W is continuous and that x_k converges to x. Then

$$\lim_{k\to\infty} W(x_k) = W(\lim_{k\to\infty} x_k) = W(x)$$

by continuity of W. Since on the other hand

$$\lim_{k\to\infty} W(x_k) = \inf_{a\in\mathbb{R}} W(a)\,,$$

we see that x must be a global minimizer (not necessarily unique). We have used continuity of W and convergence of x_k. The continuity of W in this argument is not really necessary. Assume that W is **lower-semicontinuous**. By this we mean that

$$x_k \to x \quad \Rightarrow \quad \liminf_{k\to\infty} W(x_k) \geq W(x)\,.$$

Lower-semicontinuous functions may have jumps, but in jumping, one is only lowering the function value. Now assume again that an infimizing sequence x_k is converging to x. Then

$$\inf_{a\in\mathbb{R}} W(a) \leq W(x) \leq \liminf_{k\to\infty} W(x_k) = \lim_{k\to\infty} W(x_k) = \inf_{a\in\mathbb{R}} W(a)\,.$$

Thus the limit x must be a global minimizer.

Still, it is possible that an infimizing sequence is escaping to infinity, by this we mean

$$|x_k| \to \infty \quad \text{while} \quad W(x_k) \to \inf_{a \in \mathbb{R}} W(a)\,.$$

If this happens, there is no convergence of x_k possible. We need assumptions on W which exclude this behaviour. These conditions are called **coercivity-conditions**. We may assume e.g,

$$W(x) \geq C_1^+ \, |x| - C_2\,. \tag{5.19}$$

Then it is impossible that an infimizing sequence escapes to infinity since

$$\lim_{k\to\infty} W(x_k) = \inf_{a\in\mathbb{R}} W(a)$$

and $|x_k| \to \infty$ is a contradiction due to (**??**).
Finally, we would like to have an infimizing sequence, not escaping to infinity, which converges to some limit point. From **coercivity** we conclude that our **infimizing sequence is necessarily bounded**

$$|x_k| \leq K\,.$$

Since our setting is finite dimensional, we know from Bolzanos Theorem that each bounded sequence of numbers contains at least one convergent subsequence meaning $x_{k_j} \to x$ if $k_j \to \infty$. This property is known as **compactness**. Denoting x_{k_j} again by x_k, (a standard abuse of notation) we have found a convergent, infimizing sequence, and from the preceeding we know that the limit must be a global minimizer. Thus in order to show that global minimizer exist one route is to show that

1. W is bounded below.
2. W admits, therefore, infimizing sequences.
3. W is coercive such that infimizing sequences stay bounded.
4. compactness: the infimizing sequence contains a convergent subsequence.
5. W is lower-semicontinuous: the limit of the subsequence realizes the infimum energy level.

5.3 Function spaces: Lebesgue and Sobolev

Let $1 < p < \infty$ and assume that $\Omega \subset \mathbb{R}^n$ is an open, bounded set. The space $L^p(\Omega)$ of Lebesgue-integrable functions is defined as

$$L^p(\Omega) := \{f : \Omega \subset \mathbb{R}^n \mapsto \mathbb{R}\,, \quad \int_\Omega |f(x)|^p \, \mathrm{dV} < \infty\}\,.$$

This space is a reflexive function space with norm

$$\|f\|^p_{L^p(\Omega)} := \int_\Omega |f(x)|^p \,\mathrm{dV}\,.$$

For $p = 2$ there is a corresponding scalar product defined by

$$\langle f, g\rangle_{L^2(\Omega)} := \int_\Omega f(x)\, g(x) \,\mathrm{dV}\,,$$

for $f, g \in L^2(\Omega)$. The Sobolev space $W^{1,p}(\Omega)$ is defined as

$$W^{1,p}(\Omega) := \{f \in L^p(\Omega)\,, \quad \text{the distributional derivative } \nabla f \text{ satisfies } \int_\Omega |\nabla f(x)|^p \,\mathrm{dV} < \infty\}\,,$$

with norm

$$\|f\|^p_{W^{1,p}(\Omega)} := \int_\Omega |f(x)|^p + |\nabla f(x)|^p \,\mathrm{dV}\,.$$

For $p = 2$ this is a Hilbert space with corresponding scalar product

$$\langle f, g\rangle_{W^{1,2}(\Omega)} := \int_\Omega \langle f(x), g(x)\rangle + \langle \nabla f(x), \nabla g(x)\rangle \,\mathrm{dV}\,.$$

The one property which we need predominantly is that a bounded sequence $f_k \in W^{1,p}(\Omega)$ admits a subsequence converging weakly to some element $f \in W^{1,p}(\Omega)$. Other important properties are so called embedding theorems saying that functions in $W^{1,p}(\Omega)$ are e.g., already continuous, provided certain relations between p and n obtain. For example

$$W^{1,p}(\Omega) \subset C^{1-\frac{n}{p}}(\overline{\Omega})\,,$$

if $p > n$. For the details we refer the reader to [**?**, **?**].

5.4 The direct methods of the calculus of variations in the function space setting

In the infinite-dimensional setting there are several obstructions in repeating the previous method. For purposes of illustration let us specify our domain Ω to be an interval (a, b). Our energy shall be given by

$$I(\varphi) = \int_a^b W(\varphi'(x)) \,\mathrm{dx}\,.$$

We impose Dirichlet boundary conditions, i.e. $\varphi(a) = g$ and assume the **coercivity-condition**

$$W(\xi) \geq C_1^+ \,|\xi|^2 - C_2\,. \tag{5.20}$$

Let us first show that in the space $H^1(a,b)$ of functions defined over the intervall (a,b) with square integrable derivatives, the energy I is bounded below. To this end consider

$$\begin{aligned}\infty > I(\varphi_k) &= \int_a^b W(\varphi'(x))\,\mathrm{dx} \\ &\geq \int_a^b C_1^+\,|\varphi'(x)|^2 - C_2\,\mathrm{dx} \\ &\geq C_1^+ \int_a^b |\varphi'(x)|^2\,\mathrm{dx} - C_2\,[b-a]\,.\end{aligned}$$

Using the one-dimensional Poincaré-inequality and the boundary condition $\varphi(a) = g$ we obtain

$$\infty > I(\varphi_k) \geq C_D^+ \int_a^b |\varphi(x)|^2 + |\varphi'(x)|^2\,\mathrm{dx} - C_2\,[b-a]\,.$$

Hence we conclude that sequences with bounded energy, satisfying the boundary conditions, are necessarily bounded in the space $H^1(a,b)$ of square-integrable functions with square integrable distributional derivatives. This step is usually straight forward, based on **local coercivity conditions**[2] like (**??**) for multi-dimensional situations. As a consequence we obtain that the energy I is bounded below and that, therefore, infimizing sequences of functions $\varphi_k \in H^1(a,b)$ exist. These infimizing sequences stay bounded in $H^1(a,b)$, i.e., for some positive constant,

$$K \geq \int_a^b |\varphi(x)|^2 + |\varphi'(x)|^2\,\mathrm{dx}\,.$$

However, in the infinite-dimensional setting, boundedness of the H^1-norm does not imply that the sequence of functions converges strongly in H^1! For examples we refer to [**?**]. The best one can hope for is weak-convergence, i.e., there exists a subsequence such that

$$\varphi_{k_j} \rightharpoonup \varphi \quad \text{in } H^1(a,b)\,.$$

Our hope is that along weakly convergent, infimizing sequences we may find global minimizers of I. Now we need something to replace lower-semicontinuity (which worked for strong-convergence). We need to require the **weak lower-semicontinuity**, i.e., along weakly convergent sequences

$$\varphi_k \rightharpoonup \varphi \in H^1(a,b) \quad \Rightarrow \quad I(\varphi) \leq \liminf_{k\to\infty} I(\varphi_k)$$

In our one-dimensional setting, a sufficient condition for weak-lower semicontinuity is the **convexity** of the integrand W. Therefore,

$$I(\varphi) \leq \liminf_{k} I(\varphi_k) = \lim_{k\to\infty} I(\varphi_k) = \inf_{\tilde\varphi\in H^1(a,b),\,\tilde\varphi(a)=g} I(\tilde\varphi)\,.$$

---

[2] By this we mean an estimate showing that locally, the energy density $W(F)$ bounds some power of $\|F\|$. Recall, that in linearized elasticity, this is not true any more, since only the symmetric part of F is locally controlled. Coercivity is only established after integration over the domain and using appropriate boundary conditions.

Thus, if W is convex, then existence of global mimimizers is guaranteed, as in the finite-dimensional setting. In this scalar-valued setting ($\varphi(x) \in \mathbb{R}$) the only weakly-lower semi-continuous energies are those based on convex integrands W. Here, **convexity is used to show weak-lower semicontinuity of the problem**.

5.5 How to use polyconvexity to obtain weak lower semicontinuity

In the vectorvalued setting polyconvexity is a readily available tool to obtain weak lower semicontinuity. Without entering into the details let us explain the main idea. Let

$$I(\varphi) = \int_\Omega W(\nabla\varphi)\,\mathrm{dx}$$

be the given stored energy with a polyconvex W, i.e.,

$$W(F) = P(F, \operatorname{Cof} F, \det[F])$$

and P convex. Assume one has already an infimizing sequence which converges weakly. Since this is a sequence of gradients we have,

$$\nabla\varphi_k \rightharpoonup \nabla\varphi \quad \text{in some } L^q(\Omega)\,.$$

By the weak continuity of the minors of gradients we have as well

$$T(\nabla\varphi_k) = (\nabla\varphi_k, \operatorname{Cof}\nabla\varphi_k, \det[\nabla\varphi_k]) \rightharpoonup T(\nabla\varphi) = (\nabla\varphi, \operatorname{Cof}\nabla\varphi, \det[\nabla\varphi])\,.$$

Thus the convexity of P together with the weak convergence of the argument vector $T_k \rightharpoonup T$ ensures

$$\begin{aligned}
I(\varphi) &= \int_\Omega W(\nabla\varphi)\,\mathrm{dx} \\
&= \int_\Omega P(\nabla\varphi, \operatorname{Cof}\nabla\varphi, \det[\nabla\varphi])\;\mathrm{dx} \\
&= \int_\Omega P(T(\nabla\varphi))\,\mathrm{dx} = \int_\Omega P(T)\,\mathrm{dx} \\
&\le \liminf_{k\to\infty} \int_\Omega P(T_k)\,\mathrm{dx} = \int_\Omega P(T(\nabla\varphi_k))\,\mathrm{dx} = \int_\Omega W(\nabla\varphi_k)\,\mathrm{dx}\,.
\end{aligned}$$

Thus, polyconvexity is in fact using the convexity in a nontrivial way.

The major task (not yet completely solved): find manageable conditions (other than polyconvexity) on the integrand W such that weak lower-semicontinuity is satisfied. It is known that Legendre-Hadamard ellipticity is not sufficient in general.

6 Energies

Let us call $W : \mathbb{M}^{3\times 3} \mapsto \mathbb{R}$ the energy density. We start with two properties of the energy, first the frame-indifference condition (a physically must) and second a material symmetry property: isotropy.

6.1 Frame-indifference of the energy

Frame-indifference expresses the requirement that rigidly rotating a body should not alter the stored elastic energy. This translates into the requirement of left-invariance under SO(3):

$$\forall\, Q \in \mathrm{SO}(3): \quad W(Q\,F) = W(F)\,. \tag{6.21}$$

Note that here $Q \in \mathrm{SO}(3)$ is a rigid rotation. It is easily seen that if

$$W(F) = \Psi(F^T F)\,, \tag{6.22}$$

then W is frame-indifferent. The converse can be established assuming that $F \in \mathrm{GL}^+(3)$. Assume that $F = \nabla\varphi$ is homogeneous. Since

$$W(F) = W(QF) = W(QR(F)U(F)) = W(U) = W(\sqrt{F^T F}) = \Psi(F^T F)\,.$$

Here, $F = RU$ is the unique polar-decomposition into proper orthogonal and positive definite symmetric stretch for $F \in \mathrm{GL}^+(3)$, moreover, we have chosen $Q = R^T(F)$.[3]

6.2 Material isotropy

Material isotropy is the property that the material in a specified (reference) configuration has no preferred direction. Imagine a ball of homogeneous material without preferrred direction. Applying a specific load to that ball will give rise to a specific deformation and stress response. If we rotate the ball before loading (material rotation), the response should be unchanged.

In terms of the elastic energy W this translates into the requirement

$$\forall\; Q \in \mathrm{SO}(3): \quad W(F\,Q) = W(F)\,,$$

i.e., right-invariance under special rotations. There are some subtleties involved in also assuming the right-invariance under $O(3)$ which we do not discuss here.

It is useful to remark that isotropy is immediately lost upon deformation: if the ball is already deformed into an ellipse, it will make a difference whether we rotate this ellipse prior to deformation!

Since we are dealing with an energy which is already frame-indifferent we have both left- and right -invariance under SO(3)

$$\forall\; Q_1\,, Q_2 \in \mathrm{SO}(3): \quad W(Q_1 F Q_2) = W(F)\,.$$

---

[3] The reader should note that this reduction of $W(F) = \Psi(F^T F)$ is only valid in first order theories, i.e. energies only depending on F and not on higher gradients.

In terms of the representation of $W(F) = \Psi(F^T F)$ (frame-indifference) we obtain the corresponding isotropy- invariance condition as

$$W(F\,Q) = W(F) \;\Leftrightarrow\; \Psi((FQ)^T(FQ)) = \Psi(Q^T C Q) = \Psi(C)\,.$$

Using this result, we may orthogonally diagonalize $F^T F$

$$F^T F = Q^T \operatorname{diag}(\lambda_1^2, \lambda_2^2, \lambda_3^2) Q\,,$$

where $\lambda_i > 0$ are the eigenvalues of $U = \sqrt{F^T F}$. Note that the numbering of the eigenvalues can be changed by permutating the columns of the matrix Q appropriately. Thus we must have

$$\Psi(F^T F) = \Psi(Q^T F^T F Q) = \Psi(\operatorname{diag}(\lambda_1^2, \lambda_2^2, \lambda_3^2)) = \tilde{\Psi}(\lambda_1, \lambda_2, \lambda_3)\,, \tag{6.23}$$

and $\tilde{\Psi}$ must be invariant under permutation of the entries (called symmetry of $\tilde{\Psi}$). The eigenvalues of U are also called singular values.

Theorem 6.1 (Representation for homogeneous, isotropic energies I)
Let $W \in C^2(\mathbb{M}^{3\times 3}, \mathbb{R})$ be frame-indifferent and isotropic. Assume that the reference configuration is stress free, i.e., $DW(1\!\!1) = 0$ and let the material be homogeneous, meaning that W depends only on F but not on the location $x \in \Omega$. Then there exist constants $\lambda, \mu > 0$, the so called Lamé-constants, such that

$$W(F) = \frac{\mu}{4}\|F^T F - 1\!\!1\|^2 + \frac{\lambda}{8}\langle F^T F - 1\!\!1, 1\!\!1\rangle^2 + o(\|F^T F - 1\!\!1\|^2)\,.$$

Proof. [?], S.156. ■

With the same assumptions we have as well

Corollary 6.2 (Representation for homogeneous, isotropic energies II)
Let $W \in C^2(\mathbb{M}^{3\times 3}, \mathbb{R})$ be frame-indifferent and isotropic. Assume that the reference configuration is stress free, i.e., $DW(1\!\!1) = 0$ and let the material be homogeneous, meaning that W depends only on F but not on the location $x \in \Omega$. Then there exist constants $\lambda, \mu > 0$, the so called Lamé-constants, such that

$$W(F) = \mu\|\sqrt{F^T F} - 1\!\!1\|^2 + \frac{\lambda}{2}\langle \sqrt{F^T F} - 1\!\!1, 1\!\!1\rangle^2 + o(\|\sqrt{F^T F} - 1\!\!1\|^2)\,,$$

where $\sqrt{F^T F} = U$ is the symmetric, positive definite Biot stretch tensor. This representation is usually circumvented since computing first U as a function of F is cumbersome.

6.3 Nonlinear St.Venant-Kirchhoff energy

Let $E = \frac{1}{2}(F^T F - 1\!\!1)$ be the Green strain tensor. A function W of the form

$$\begin{aligned}
\hat{W}(E) &= 4\mu\|E\|^2 + 4\lambda\langle E, 1\!\!1\rangle^2 \\
W(F) := \hat{W}(E) &= \mu\|F^T F - 1\!\!1\|^2 + \lambda\langle F^T F - 1\!\!1, 1\!\!1\rangle^2
\end{aligned}$$

is called energy of St.Venant-Kirchhoff type with the Lamé-constants $\mu, \lambda > 0$.[4] We are interested in the first and second derivatives:

$$\begin{aligned} DW(F).H &= 2\mu \cdot \langle F^T F - \mathbb{1}, F^T H + H^T F\rangle + \\ &\qquad 2\lambda \cdot \langle F^T F - \mathbb{1}, \mathbb{1}\rangle \langle F^T H + H^T F, \mathbb{1}\rangle \\ D^2 W(F).(H,H) &= 2\mu \|F^T H + H^T F\|^2 + 4\mu \langle F^T F - \mathbb{1}, H^T H\rangle + \\ &\qquad 8\lambda \langle F, H\rangle^2 + 4\lambda(\|F\|^2 - 3)\|H\|^2 \,. \end{aligned}$$

We observe that $D^2W(F).(H,H)$ is not positive throughout. Thus the St.Venant-Kirchhoff energy is not **not convex**. Let us proceed to show that this energy is not Legendre-Hadamard elliptic either. To this end we choose a uniform compression $F_0 = \frac{1}{m} \cdot \mathbb{1}$ and consider

$$\begin{aligned} D^2 W(F_0).(H,H) &= 2\mu \|\frac{1}{m} H + \frac{1}{m} H^T\|^2 + 4\mu \langle \frac{1}{m^2}\mathbb{1} - \mathbb{1}, H^T H\rangle + \\ &\quad 8\lambda \langle \frac{1}{m}\mathbb{1}, H\rangle^2 + 4\lambda(\|\frac{1}{m}\mathbb{1}\|^2 - 3)\|H\|^2 \\ &= \frac{1}{m^2}\big(2\mu \|H + H^T\|^2 + 4\mu(1-m^2)\langle \mathbb{1}, H^T H\rangle + \\ &\quad 8\lambda \langle \mathbb{1}, H\rangle^2 + 4\lambda 3(1-m^2)\|H\|^2\big) \\ &\text{take } H = \xi \otimes \eta \text{ to obtain} \\ &= \frac{1}{m^2}\big(2\mu(2\|\xi\|^2\|\eta\|^2 + 2\langle \xi, \eta\rangle^2) + 4\mu(1-m^2)\|\xi\|^2\|\eta^2\| + \\ &\quad 8\lambda \langle \xi, \eta\rangle^2 + 4\lambda 3(1-m^2)\|\xi\|^2\|\eta\|^2\big) \\ &\text{choose } \xi, \eta \text{ such that } \langle \xi, \eta\rangle = 0 \\ &= \frac{1}{m^2}\big(2\mu(2\|\xi\|^2\|\eta\|^2 + 4\mu(1-m^2)\|\xi\|^2\|\eta^2\| + 12\lambda(1-m^2)\|\xi\|^2\|\eta\|^2\big) \\ &= \frac{1}{m^2}\|\xi\|^2\|\eta\|^2\left(4\mu + 4\mu(1-m^2) + 12\lambda(1-m^2)\right) \,. \end{aligned}$$

Taking m sufficiently large, i.e. compressing strongly, it follows

$$D^2 W(F_0).(\xi \otimes \eta, \xi \otimes \eta) < 0 \,.$$

Hence, W is **not elliptic**. Therefore W is not polyconvex/quasiconvex either. This energy is therefore not suited for realistic applications in the nonlinear regime. Its only virtue is that it is invariant under rigid rotations and that it is consistent with isotropic linear elasticity.

6.4 Energies of Green-Naghdi type in elasto-plasticity

While I am not concerned with plasticity theory here, the question still arises, whether the potential energy used for the elastic part of the deformation will lead to an elliptic boundary value problem once the plastic process is kept fixed. This is not an academic

---

[4] This definition of the elastic constants is actually valid only up to constant factors!

problem since usually, in FEM-calculations, the so-called elastic trial step may be far off the final equilibrium position. If, then, the elastic response is not itself well-behaved, the algorithm might run in troubles.

We have already seen that by using the multiplicative decomposition $F = F_e\, F_p$ and assuming a polyconvex potential in the variable F_e (which is not a gradient), the potential as a function of F will be polyconvex at given F_p. In this sense we may say that elastic polyconvexity is preserved under plastic flow. The same holds for convexity, rank-one convexity and quasiconvexity.

Plasticity models do not need to be not be based on the multiplicative decomposition. Green-Naghdi postulate in [?] a split of the total strain tensor $E = F^T F - 1\!\!1$ in an elastic and plastic contribution, such that

$$E = E_e + E_p\,, \quad E_e = E - E_p\,.$$

Here, E_p is the plastic strain which has to be determined by some flow rule. The energy governing the elastic response of the material is defined in terms of E_e, thus

$$W(F, E_p) := \hat{W}(E_e) = \hat{W}(F^T F - 1\!\!1 - E_p)\,.$$

If $\hat{W}$ is a convex function of E_e one may show, similar to the case of the St.Venant-Kirchhoff energy that the ensuing PDE-system at given E_p is not Legendre-Hadamard elliptic.

6.5 Biot-type energy

If we specify the energy to depend only on the stretch part $U = \sqrt{F^T F}$ in the polar decomposition of F,

$$W_{\mathrm{Biot}}(F) = \mu\, \|U - 1\!\!1\|^2 + \frac{\lambda}{2}\, \mathrm{tr}\, [U - 1\!\!1]^2\,,$$

we obtain another frame-indifferent, isotropic replacement of linear elasticity. While this energy has certain compelling features setting it apart from the SVK-energy, one can show that it is still not leading to an elliptic boundary value problem [?]. In the one-dimensional case the energy looks still like a (nonconvex) double well potential.

6.6 Linearized St.Venant-Kirchhoff energy

Setting $F = 1\!\!1 + \nabla u$ implies

$$\begin{aligned} F^T F - 1\!\!1 &= (1\!\!1 + \nabla u)^T (1\!\!1 + \nabla u) - 1\!\!1 \\ &= (1\!\!1 + \nabla u^T + \nabla u + \nabla u^T \nabla u) - 1\!\!1 \\ &= \nabla u^T + \nabla u + \nabla u^T \nabla u\,. \end{aligned}$$

For small displacement gradients ($\nabla u \ll 1$) one may neglect the quadratic term and abbreviating the infinitesimal strain tensor by ε,

$$\varepsilon = \mathrm{sym}(\nabla u) = \frac{1}{2}(\nabla u^T + \nabla u)\,,$$

we obtain

$$\hat{W}(\nabla u) = \mu\|\varepsilon\|^2 + \frac{\lambda}{2}\langle\varepsilon, 1\!\!1\rangle^2 = \frac{\mu}{4}\|\nabla u^T + \nabla u\|^2 + \frac{\lambda}{8}\langle\nabla u^T + \nabla u, 1\!\!1\rangle^2 .$$

The first and second differential w.r.t. ∇u read

$$\begin{aligned} D(\hat{W}(\nabla u)).H &= \mu\langle\nabla u + \nabla u^T, H\rangle + \lambda\cdot\langle\nabla u, 1\!\!1\rangle\cdot\langle 1\!\!1, H\rangle \\ &= \langle 2\mu\varepsilon + \lambda\langle\varepsilon, 1\!\!1\rangle\cdot 1\!\!1, H\rangle , \\ D^2(\hat{W}(\nabla u)).(H,H) &= \frac{1}{2}\mu\|H^T + H\|^2 + \lambda\langle 1\!\!1, H\rangle^2 \geq \frac{1}{2}\mu\|H^T + H\|^2 . \end{aligned}$$

This implies uniform Korn-ellipticity w.r.t.∇u (but not local coercivity). The corresponding Cauchy stress tensor $\sigma(\nabla u)$ verifies

$$\begin{aligned} \langle\sigma(\nabla u), H\rangle &= D(\hat{W}(\varepsilon).H \\ &= \langle 2\mu\varepsilon + \lambda\langle\varepsilon, 1\!\!1\rangle\cdot 1\!\!1, H\rangle \\ \Rightarrow \sigma(\nabla u) &= \mu\nabla u + \mu\nabla u^T + \lambda\cdot\operatorname{tr}(\nabla u)\cdot 1\!\!1 \\ &= \mu\nabla u + \mu\nabla u^T + \lambda\cdot\operatorname{Div} u\cdot 1\!\!1 . \end{aligned}$$

Since $\operatorname{Div}(\nabla u^T) = \nabla(\operatorname{Div} u)$ und $\operatorname{Div}((\operatorname{Div} u)\cdot 1\!\!1) = \nabla(\operatorname{Div} u)$ it follows

$$\operatorname{Div}\ \sigma(\nabla u) = (2\mu + \lambda)\,\nabla(\operatorname{Div} u) + 2\mu\Delta u .$$

This is the well-known expression for the divergence in the **Lam*é*-equations of linearized elasticity**.

6.7 Ogden-material

Let $f \in C^1((0,\infty),\mathbb{R})$ be a convex function. Energies of the form

$$\begin{aligned} W(F) &= \alpha\ \operatorname{tr}(F^TF - 1\!\!1) + \beta\ \operatorname{tr}(\operatorname{Adj} F^TF - 1\!\!1) + f(\det[F]) \\ &= \alpha\|F\|^2 + \beta\|\operatorname{Adj} F\|^2 + f(\det[F]) - 3\alpha - 3\beta \end{aligned}$$

are simplest examples for so-called Ogden-materials [**?**, **?**]. (In this special form they are also called **compressible Mooney-Rivlin energies**). As defined, this energy is obviously polyconvex. Choosing $\alpha, \beta > 0$ and $f'(1)$ allows to adjust to a stress-free reference configuration

$$\forall H \in \mathbb{M}^{3\times 3} : \quad DW(1\!\!1).H = 0 .$$

Often $f(x) = -ln(x)$ or $f(x) = \frac{1}{x^2}$ is used which implies a physical singularity for $\det[F] \to 0$. The first and second derivative are calculated as

$$\begin{aligned} DW(F).H &= 2\alpha\langle F, H\rangle + 2\beta\langle\operatorname{Adj} F, \operatorname{Adj} F\langle F^{-T}, H\rangle - (\operatorname{Adj} F)HF^{-1}\rangle + \\ &\quad f'(\det[F])\det[F]\langle F^{-T}, H\rangle , \\ D^2W(F).(H,H) &= 2\alpha\|H\|^2 + 4\beta\langle\operatorname{Adj} F, \operatorname{Adj} H\rangle + \\ &\quad 2\beta\|D(\operatorname{Adj} F).H\|^2 + f''(\det[F])(\det[F]\langle F^{-T}, H\rangle)^2 + \\ &\quad f'(\det[F])\det[F]\langle F^{-T}, H\rangle^2 - f'(\det[F])\det[F]\langle(F^{-1}H)^T, F^{-1}H\rangle . \end{aligned}$$

For incompressible material ($\det[F] = 1$) this is reduced to

$$\begin{aligned} W(F) &= \alpha\, \|F\|^2 + \beta\, \|F^{-1}\|^2 \\ &= \alpha\, (\lambda_1^2 + \lambda_2^2 + \lambda_3^2) + \beta\, (\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} + \frac{1}{\lambda_3^2})\,, \end{aligned}$$

where $\lambda = \lambda(\sqrt{F^T F})$ are the singular values of F.

The most general form of Ogden type materials is given by

$$W(F) = \sum_i^N\, \alpha_i \|F\|^{p_i} + \sum_j^M \beta_j \|\operatorname{Adj} F\|^{q_j} + f(\det[F])\,,$$

in which the constants $\alpha_i, \beta_j, p_i, q_j$ can be determined from experiment.

6.8 Compressibles Neo-Hookean material

A subclass of Ogden type materials appears by setting the parameter $\beta = 0$ and considering a smooth convex function

$$f(x) = -\eta_1 \cdot (\ln x) + \eta_2 (\ln x)^2 + \eta_3 x^2 - \eta_4 x + \eta_5, \quad \eta_i > 0\,, i = 1 \dots 5\,.$$

This energy is also called to be of **Neo-Hooke type** since the leading term $\alpha \|F\|^2$ alone would lead to a linear relation between stretch and strain.

6.9 Incompressible Neo-Hookean material

The energy looks like

$$W(F) = \|F\|^2\,.$$

However, the corresponding equilibrium system is by no means linear since the nonlinear constraint $\det[F] = 1$ has to be incorporated.

6.10 Polyconvex model energy with singularity for $\det[F] \to 0$

Having in mind that quasiconvexity is only useful for energies satisfying a polynomial growth condition the strength of polyconvexity comes into play when considering energies which do not satisfy such a condition. A very simple polyconvex energy with (a physically meaningful) singularity for $\det[F] \to 0$ is obtained by setting for $F:\ \det[F] > 0$

$$W(F) = \operatorname{tr}(F^T F - 1\!\!1) - \ln \det[(F^T F)] = \|F\|^2 - 2 \ln \det[F] - 3\,.$$

This energy is material frame-indifferent and isotropic. The first and second differential are

$$\begin{aligned} DW(F).H &= 2\langle F, H\rangle - 2\langle F^{-T}, H\rangle\,, \\ D^2W(F).(H,H) &= 2\|H\|^2 + 2\langle (F^{-1}H)^T, F^{-1}H\rangle \\ &= 2\|H\|^2 + 2\langle F^{-T}, H\rangle^2 - \frac{4}{\det[F]}\langle \operatorname{Adj} H, F^T\rangle \\ &\geq 2\|H\|^2 - \frac{4}{\det[F]}\langle \operatorname{Adj} H, F^T\rangle\,. \end{aligned}$$

Thus $DW(1\!\!1).H = 2\langle 1\!\!1, H\rangle - 2\langle 1\!\!1^{-T}, H\rangle \equiv 0$, hence $F = 1\!\!1$ is a stress-free reference configuration. If $\det[F] \geq c^+ > 0$ then $H \mapsto D^2W(F).(H,H)$ is strictly polyconvex and strictly Legendre-Hadamard elliptic, which can be shown by straightforward computations.

6.11 Energies defined on the Hencky-tensor

Let $C = F^TF$. The matrix logarithm $\ln C$ is also called Hencky-tensor. In the engineering literature one may encounter the following isotropic and frame-indifferent energy:

$$\begin{aligned} W(F) &= \langle \ln C, 1\!\!1\rangle^p, \quad p \geq 2 \\ &= 2^p\,(\ln\det[F])^p \quad \text{if } \det[F] > 0\,. \end{aligned}$$

The first and second derivative are given by

$$\begin{aligned} DW(F).H &= p\, 2^p\,(\ln\det[F])^{p-1}\,\langle F^{-T}, H\rangle\,, \\ D^2W(F).(H,H) &= p\,(p-1)\,2^p(\ln\det[F])^{p-2}\,\langle F^{-T}, H\rangle^2 \\ &\quad - p\,2^p\,(\ln\det[F])^{p-1}\,\langle (HF^{-1})^T, HF^{-1}\rangle\,, \end{aligned}$$

and we observe, that $1\!\!1$ is a stress-free reference configuration. However, we show that this energy will not lead to an elliptic system. To this end choose $F_0 = m \cdot 1\!\!1$. Then $\det[F_0] = m^3$ and $F^{-1} = \frac{1}{m} \cdot 1\!\!1$. We have

$$\begin{aligned} D^2W(F_0).(H,H) &= p\,(p-1)\,2^p\,(\ln m^3)^{p-2}\,\frac{1}{m^2}\langle 1\!\!1, H\rangle^2 \\ &\quad - p\,2^p\,(\ln m^3)^{p-1}\,\frac{1}{m^2}\langle H^T, H\rangle \\ &= \frac{p\,2^p}{m^2}\big((p-1)(\ln m^3)^{p-2}\,\langle 1\!\!1, H\rangle^2 - (\ln m^3)^{p-1}\,\langle H^T, H\rangle\big) \\ &\quad \text{choose } H = \eta\otimes\xi \\ D^2W(F_0).(\eta\otimes\xi, \eta\otimes\xi) &= \frac{p\,2^p}{m^2}(\ln m^3)^{p-2}[(p-1) - (\ln m^3)]\,\langle \xi, \eta\rangle^2 \end{aligned}$$

There always exists an $m > 1$ with $D^2W(F_0).(\eta\otimes\xi, \eta\otimes\xi) < 0$. Here, e.g., $\frac{p-1}{3} < \ln m$. Thus this energy W is **not elliptic**. ■

6.12 Sansours energy

Set $C = F^T F$ and let $\Psi : \mathrm{Sym} \mapsto \mathbb{R}$ be isotropic. We consider

$$W(F) = \hat{W}(C) = \Psi(\ln C)\,.$$

It is shown by Sansour in [?] that (see (**??**))

$$D_C\hat{W}(C) = D\Psi(\ln C) \cdot C^{-1}\,.$$

Let us consider more specifically a quadratic Ψ:

$$\begin{aligned}\Psi(X) &:= \lambda\langle X, 1\!\!1\rangle^2 + \mu\|\operatorname{dev}(X)\|^2\,,\\ D\Psi(X).H &= 2\lambda\langle X, 1\!\!1\rangle\langle 1\!\!1, H\rangle + 2\mu\langle \operatorname{dev}(X), \operatorname{dev}(H)\rangle\\ &= 2\lambda\langle X, 1\!\!1\rangle\langle 1\!\!1, H\rangle + 2\mu\langle \operatorname{dev}(X), H\rangle\,,\end{aligned}$$

with constants $\lambda, \mu > 0$. This implies

$$\begin{aligned}W(F) = \hat{W}(C) &= \lambda\langle \ln C, 1\!\!1\rangle^2 + \mu\|\operatorname{dev}(\ln C)\|^2\,,\\ D_F W(F).H = \langle S_1, H\rangle &= \langle F\cdot S_2, H\rangle = \langle F \cdot D_C\hat{W}(F^TF), H\rangle\\ &= \langle F\ D\Psi(\ln C)\ C^{-1}, H\rangle = \langle D\Psi(\ln C), F^T H C^{-1}\rangle\\ &= 2\lambda\langle \ln F^TF, 1\!\!1\rangle\langle 1\!\!1, F^THC^{-1}\rangle + 2\mu\langle \operatorname{dev}(\ln F^TF), F^THC^{-1}\rangle\end{aligned}$$

and we see that $1\!\!1$ is a stress-free reference configuration. Let us show that the energy is not Legendre-Hadamard elliptic. To this end consider the function $h : \mathbb{R} \mapsto \mathbb{R}$

$$\begin{aligned}h(t) &= W(1\!\!1 + t(\eta\otimes\xi))\\ &= \hat{W}((1\!\!1 + t(\eta\otimes\xi))^T(1\!\!1 + t(\eta\otimes\xi)))\\ &= \hat{W}(1\!\!1 + t(\eta\otimes\xi + \xi\otimes\eta) + t^2(\eta\otimes\xi)^T(\eta\otimes\xi))\,.\end{aligned}$$

Taking $\eta\otimes\xi = \begin{pmatrix}0&1&0\\0&0&0\\0&0&0\end{pmatrix}$ implies

$$\begin{aligned}(1\!\!1 + t(\eta\otimes\xi))^T(1\!\!1 + t(\eta\otimes\xi)) &= \begin{pmatrix}1&t&0\\t&1+t^2&0\\0&0&1\end{pmatrix}\,,\\ \det\begin{pmatrix}1-\kappa&t&0\\t&1+t^2-\kappa&0\\0&0&1-\kappa\end{pmatrix} &= (1-\kappa)\cdot(\kappa^2 - \kappa(2+t^2)+1)\,,\end{aligned}$$

and the eigenvalues are

$$\begin{aligned}&\kappa_1 = 1,\quad \kappa_2 = \frac{1}{2}\left(2 + t^2 + t\sqrt{4+t^2}\right),\quad \kappa_3 = \frac{1}{2}\left(2+t^2 - t\sqrt{4+t^2}\right)\,,\\ &\kappa_1\kappa_2\kappa_3 = 1\,,\\ &0 = \ln\kappa_1\kappa_2\kappa_3 = \ln\kappa_2\kappa_3 = \ln\kappa_2 + \ln\kappa_3\,,\\ &\ln\kappa_2 = -\ln\kappa_3\,.\end{aligned}$$

Diagonalising $C = F^T F$ and using the isotropy of the energy we may obtain the following represention of W in terms of the eigenvalues of C:

$$\begin{aligned}
\hat{W}(F^T F) &= \lambda \langle \ln C, 1\!\!1 \rangle^2 + \mu \| \operatorname{dev}(\ln C) \|^2 \\
&= \lambda \left(\ln(\kappa_1 \kappa_2 \kappa_3)\right)^2 + \mu \left((\ln \kappa_1)^2 + (\ln \kappa_2)^2 + (\ln \kappa_3)^2 - \frac{1}{3} (\ln \kappa_1 + \ln \kappa_2 + \ln \kappa_3)^2 \right) .
\end{aligned}$$

This implies for the function h:

$$\begin{aligned}
h(t) &= \lambda \left(\ln(\kappa_1 \kappa_2 \kappa_3)\right)^2 + \mu \left((\ln \kappa_1)^2 + (\ln \kappa_2)^2 + (\ln \kappa_3)^2 - \frac{1}{3} (\ln \kappa_1 + \ln \kappa_2 + \ln \kappa_3)^2 \right) \\
&= 2\mu (\ln \kappa_2)^2 = 2\mu [\ln \left(2 + t^2 + t\sqrt{4 + t^2}\right) - \ln 2]^2 .
\end{aligned}$$

However, this function $h(t)$ is not convex in t. Therefore, W is not rank-one convex, implying that W is not Legendre-Hadamard elliptic. ■

The same argument shows that

$$W(F) = \frac{1}{2} \| \ln F^T F \|^2$$

is not an LH-elliptic energy since in terms of the eigenvalues of $F^T F$ we have

$$W(F) = \frac{1}{2} \left((\ln \kappa_1)^2 + (\ln \kappa_2)^2 + (\ln \kappa_3)^2 \right)$$

and therefore, similar to the above,

$$h(t) = (\ln \kappa_2)^2 = [\ln \left(2 + t^2 + t\sqrt{4 + t^2}\right) - \ln 2]^2$$

is not convex. ■

6.13 Simo/Ortiz energy

Let us consider an energy W of the form

$$W(F) = \frac{1}{q} \|F\|^q - \ln \det[F] + (\ln \det[F])^p, \quad p, q \geq 2 .$$

This isotropic energy has been used for the modelling of compressible material in [**?**]. We show, that this energy is **not polyconvex**. On the one hand it is easily seen that W cannot be identified to be polyconvex simply by looking at the convexity properties in terms of the determinant, since

$$h(y) = -\ln y + (\ln y)^p, \quad p \geq 2 ,$$

is **not convex**. However, this failure does not automatically imply that polyconvexity is lost! (also it seems that it is highly unlikely...). In order to show the lack of polyconvexity let us adapt an ingenious argument which has been first given in [**?**] in another context. We set for $\alpha > 0$:

$$A_\alpha = \alpha\, 1\!\!1, \quad D_\alpha = \alpha \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} .$$

Then it obtains

$$\begin{aligned}
\|A_\alpha\| &= \alpha\,\sqrt{3}, & \det[A_\alpha] &= \alpha^3\,, \\
\|D_\alpha\| &= \alpha\,\sqrt{11}, & \det[D_\alpha] &= 3\,\alpha^3\,, \\
\frac{1}{2}A_\alpha + \frac{1}{2}D_\alpha &= \alpha \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\,, \\
\|\frac{1}{2}A_\alpha + \frac{1}{2}D_\alpha\| &= \alpha\,\sqrt{6}, & \det[(\frac{1}{2}A_\alpha + \frac{1}{2}D_\alpha)] &= 2\,\alpha^3\,, \\
\det[(\frac{1}{2}A_\alpha + \frac{1}{2}D_\alpha)] &= \frac{1}{2}\det[A_\alpha] + \frac{1}{2}\det[D_\alpha] \\
\mathrm{Adj}\,(\frac{1}{2}A_\alpha + \frac{1}{2}D_\alpha) &= \frac{1}{2}\,\mathrm{Adj}\,A_\alpha + \frac{1}{2}\,\mathrm{Adj}\,D_\alpha\,.
\end{aligned}$$

Assuming that W is polyconvex, then there exists a convex function $P : \mathbb{M}^{3\times3} \times \mathbb{M}^{3\times3} \times \mathbb{R}$, such that $W(F) = P(F, \mathrm{Adj}\,F, \det[F])$ and therefore

$$\begin{aligned}
W(\frac{1}{2}A_\alpha + \frac{1}{2}D_\alpha) &= P(\frac{1}{2}A_\alpha + \frac{1}{2}D_\alpha, \mathrm{Adj}\,(\frac{1}{2}A_\alpha + \frac{1}{2}D_\alpha), \det[(\frac{1}{2}A_\alpha + \frac{1}{2}D_\alpha)]) \\
&= P(\frac{1}{2}A_\alpha + \frac{1}{2}D_\alpha, \frac{1}{2}\,\mathrm{Adj}\,A_\alpha + \frac{1}{2}\,\mathrm{Adj}\,D_\alpha, \frac{1}{2}\det[A_\alpha] + \frac{1}{2}\det[D_\alpha]) \\
&\le \frac{1}{2}P(A_\alpha, \mathrm{Adj}\,A_\alpha, \det[A_\alpha]) + \frac{1}{2}P(D_\alpha, \mathrm{Adj}\,D_\alpha, \det[D_\alpha]) \\
&= \frac{1}{2}\left(W(A_\alpha) + W(D_\alpha)\right)\,.
\end{aligned}$$

Let us evaluate this condition for arbitrary $q > 2$ and $p = 2$:

$$\begin{aligned}
W(\alpha \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}) &= \frac{1}{q}\,\alpha^q\,\sqrt{6}^q - \ln 2 - 3\ln\alpha + (\ln(2\alpha^3))^2 \\
&\le \frac{1}{2}\left(\frac{1}{q}\alpha^q\,\sqrt{3}^p - 3\ln\alpha + 9(\ln\alpha)^2 + \right. \\
&\qquad \left. \frac{1}{q}\alpha^q\,\sqrt{11}^q - \ln 3 - 3\ln\alpha + (\ln(3\alpha^3))^2\right)\,.
\end{aligned}$$

However, there is always some $\alpha > 0$, such that this necessary condition for polyconvexity is violated. Therefore, W cannot be polyconvex. It can be shown, moreover, that W is not LH-elliptic either.

6.14 Modified Simo/Ortiz energy

It is very easy to modify the preceeding energy into a polyconvex one. We just add the term $(\det[F]-1)^2$ and consider

$$W(F) = \frac{1}{q}\|F\|^q - \ln\det[F] + (\ln\det[F])^p + (\det[F]-1)^2, \quad p,q \geq 2\,.$$

This energy is polyconvex for $q \geq 2, p = 2$ since it is easy to check that the function $P : \mathbb{M}^{3\times 3} \times \mathbb{R}^+ \mapsto \mathbb{R}$

$$P(\xi, y) = \frac{1}{q}\|\xi\|^q - \ln y + (\ln y)^2 + (y-1)^2$$

is convex.

7 Again polyconvexity

Let $W \in C^2(\mathbb{R}^3 \times \mathbb{M}^{3\times 3}, \mathbb{R})$ be a given scalar valued energy density. We say that

Definition 7.1 (Polyconvexity)
$F \mapsto W(x,F)$ is polyconvex if and only if there exists a function $P : \mathbb{R}^3 \times \mathbb{M}^{3\times 3} \times \mathbb{M}^{3\times 3} \times \mathbb{R} \mapsto \mathbb{R}$ (in general non unique) such that

$$W(x,F) = P(x, F, \operatorname{Adj} F, \det[F])$$

and the function $\mathbb{R}^{19} \mapsto \mathbb{R}$, $(X,Y,Z) \mapsto P(x,X,Y,Z)$ is convex forall $x \in \mathbb{R}^3$. ■

A consequence of this definition for a more restrictive class of energy densities is

Corollary 7.2 (Additive polyconvex functions)
Let $W(x,F) = W_1(x,F) + W_2(x, \operatorname{Adj} F) + W_3(x, \det[F])$. If W_i, $i = 1,2$ are convex in the second variable respectively and $W_3 : \mathbb{R}^3 \times \mathbb{R}^+ \mapsto \mathbb{R}$ is convex in the second variable as well, then W is altogether polyconvex. ■

Remark 7.3 (Construction of polyconvex energies)
The last corollary will be one of our main tools in constructing polyconvex strain energies: we identify functions which are convex on $\mathbb{M}^{3\times 3}$ and $\mathbb{R}$ and then take positive combinations of them. ■

Definition 7.4 (Ellipticity)
We say that the elastic free energy $W \in C^2(\mathbb{M}^{3\times 3}, \mathbb{R})$ leads to a uniformly elliptic equilibrium system whenever the so called uniform Legendre-Hadamard condition

$$\exists\, c^+ > 0\ \forall\, F \in \mathbb{M}^{3\times 3}\ : \forall\, \xi, \eta \in \mathbb{R}^3 : \quad D_F^2 W(F).(\xi \otimes \eta, \xi \otimes \eta) \geq c^+ \cdot \|\xi\|^2 \|\eta\|^2$$

holds. We say that W is strictly elliptic if and only if the strict Legendre-Hadamard condition

$$\forall\, F \in \mathbb{M}^{3\times3} \; : \forall\, \xi, \eta \in \mathbb{R}^3 : \quad D_F^2 W(F).(\xi \otimes \eta, \xi \otimes \eta) > 0$$

holds. We say that the elastic free energy W is strictly rank-one convex if the function $f : \mathbb{R} \mapsto \mathbb{R},\ f(t) = W(F + t\cdot(\xi\otimes\eta))$ is strictly convex forall $F \in \mathbb{M}^{3\times3}$ and all $\xi, \eta \in \mathbb{R}^3$. ■

Definition 7.5 (Quasiconvexity)
We say that the elastic free energy is quasiconvex whenever forall $\Omega \subset \mathbb{R}^3$ and all $F \in \mathbb{M}^{3\times3}$ and all $v \in C_0^\infty(\Omega)$ we have

$$W(F) \cdot |\Omega| = \int_\Omega W(F) \;\mathrm{dx} \leq \int_\Omega W(F + \nabla v(x))\,\mathrm{dx}.$$

This means, that the homogeneous solution $\nabla u = F$ of the homogeneous boundary value problem

$$\mathrm{Div}\, D_F W(\nabla u) = 0\,, \quad u_{|\partial\Omega}(x) = F.x + c\,,$$

is automatically a global minimizer. It is clear that this condition is a nonlocal stability condition and therefore difficult to handle. Every quasiconvex function is automatically elliptic.

The decisive property in the context to be treated here is the following well known property

Theorem 7.6 (Polyconvexity implies quasiconvexity and ellipticity)
Let the stored energy W be sufficiently smooth. Then if W is polyconvex it is quasiconvex and elliptic. Moreover rank-one ellipticity and ellipticity are equivalent.

Proof. Standard result in the calculus of variations [**?**]. We remark that the converse is in general not true. ■

Theorem 7.7 (Rank one convexity and eigenvalues)
Let

$$W(F) = \Phi(\lambda_1, \ldots, \lambda_n)\,,$$

*where Φ is symmetric and λ_i are the eigenvalues of $(F^T F)^{\frac{1}{2}}$ (we have shown in (**??**) that this representation can always be assumed for isotropic and frame-indifferent energies). If W is rank-one convex and $\Phi \in C^2(\mathbb{R}^n)$ then*

$$\frac{\partial^2 \Phi}{\partial \lambda_i^2}(\lambda_1, \ldots \lambda_n) \geq 0.$$

This means that rank-one convexity implies separate convexity (convexity in each argument separately) of the energy in terms of the representation in the eigenvalues λ_i.

Proof. This is Proposition 1.2 of [**?**, p.254]. ■

8 Isochoric and volumetric terms

Let $F \in \mathrm{GL}(3,\mathbb{R})$. Then

$$\det[\left(\frac{F^TF}{(\det[F^TF])^{\frac{1}{3}}}\right)] = 1.$$

It is sometimes preferable to express strain energies as a sum of isochoric (describing only the change of shape) and volumetric terms, i.e

$$W(F) = W_{iso}\left(\frac{F^TF}{(\det[F^TF])^{\frac{1}{3}}}\right) + W_{vol}(\det[F]).$$

We will show that this decomposition is compatible with the requirement of polyconvexity. Let for example $W_1(A) = \langle A, 1\!\!1\rangle$ and define $iso(F) = \frac{F^TF}{(\det[F^TF])^{\frac{1}{3}}}$. Then

$$W_{iso}\left(\frac{F^TF}{(\det[F^TF])^{\frac{1}{3}}}\right) = W_1(iso(F)) = \begin{cases} \frac{\|F\|^2}{\det[F]^{\frac{2}{3}}} & \det[F] > 0 \\ \infty & \det[F] \leq 0 \end{cases} \tag{8.1}$$

and $W(F) = W_1(iso(F))$ is a polyconvex function (however not of the additive type, see Corollary **??**) which we proceed to show in the sequel.

For the remainder let us agree to extend functions W which are naturally only defined on the (nonconvex) set $\det[F] > 0$ to $\mathbb{M}^{3\times 3}$ by setting $W = \infty$ for arguments with $\det[F] \leq 0$ as we did in the last example. It is clear by such an extension that W can never be convex, because it is only supported on a nonconvex set. However, this extension is compatible with the requirement of polyconvexity since

$$P(x) = \begin{cases} f(x) & x > 0 \\ \infty & x \leq 0 \end{cases} \tag{8.2}$$

is a convex function whenever f is convex on $\mathbb{R}^+$.

Lemma 8.1 (Isochoric terms)
Let $W(F) = \frac{\|F\|^2}{\det[F]^{\frac{2}{3}}}$. *Then* W *is polyconvex.*

Proof. We investigate first the convexity of the function $P : \mathbb{R}^+ \times \mathbb{R} \mapsto \mathbb{R}$, $P(x,y) = f(x)\cdot g(y)$. The matrix of second derivatives is of course

$$D^2P(x,y) = \begin{pmatrix} f''(x)\cdot g(y) & f'(x)\cdot g'(y) \\ f'(x)\cdot g'(y) & f(x)\cdot g''(y) \end{pmatrix}.$$

If f, g are positive, smooth and convex then we have $f''(x)\cdot g(y) \geq 0$ and $\det[D^2P(x,y)] = f''(x)g(y)f(x)g''(y) - (f'(x)g'(x))^2$. Observe that P is convex, if D^2P is positive definite by Lemma **??**. In our situation D^2P is positive definite, if $f''(x)\cdot g(y) \geq 0$ and

$\det[D^2P(x,y)] \geq 0$. Thus we must guarantee that $f''(x)\ g(y) f(x)\ g''(y) \geq (f'(x)g'(x))^2$. Let $\alpha > 0$ and $p \geq 2$. We choose $f(x) = x^{-\alpha}$ and $g(y) = y^p$. Then

$$f''(x)\ g(y) f(x)\ g''(y) = \alpha(\alpha+1)x^{-(2+\alpha)}y^p x^{-\alpha} p(p-1)y^{p-2}$$

and

$$(f'(x)g'(x))^2 = (-\alpha x^{-(\alpha+1)} p y^{p-1})^2 = \alpha^2 x^{-2(\alpha+1)} p^2 y^{2(p-1)}.$$

We arrive at the condition that

$$\frac{\alpha+1}{\alpha} \geq \frac{p}{p-1}. \tag{8.3}$$

The larger one chooses p, the better for the choice of α. Notably $P(x,y) = \frac{1}{x^\alpha} \cdot y^p$ is convex for $\alpha = \frac{2}{3}$ and $p = 2$. We set

$$\hat{W}(F,\xi) = P(\xi, \|F\|) = \frac{\|F\|^2}{\xi^{\frac{2}{3}}}.$$

We check the convexity of $\hat{W}(F,\xi)$. Thus

$$\begin{aligned}\hat{W}(\lambda F_1 + (1-\lambda)F_2, \lambda\xi_1 + (1-\lambda)\xi_2) &= P(\lambda\xi_1 + (1-\lambda)\xi_2, \|\lambda F_1 + (1-\lambda)F_2\|) \\ &= \frac{\|\lambda F_1 + (1-\lambda)F_2\|^2}{(\lambda\xi_1 + (1-\lambda)\xi_2)^{\frac{2}{3}}}\end{aligned}$$

and the monotonicity of the square for positive arguments yields

$$\begin{aligned}\hat{W}(\lambda F_1 + (1-\lambda)F_2, \lambda\xi_1 + (1-\lambda)\xi_2) &\leq \frac{(\lambda\|F_1\| + (1-\lambda)\|F_2\|)^2}{(\lambda\xi_1 + (1-\lambda)\xi_2)^{\frac{2}{3}}} \\ &= P(\lambda\xi_1 + (1-\lambda)\xi_2, \lambda\|F_1\| + (1-\lambda)\|F_2\|).\end{aligned}$$

Since by assumption P is convex, we get

$$\begin{aligned}\hat{W}(\lambda F_1 + (1-\lambda)F_2, \lambda\xi_1 + (1-\lambda)\xi_2) &\leq \lambda P(\xi_1, \|F_1\|) + (1-\lambda)P(\xi_2, \|F_2\|) \\ &= \lambda\hat{W}(F_1,\xi_1) + (1-\lambda)\hat{W}(F_2,\xi_2).\end{aligned}$$

The proof is complete. For this proof compare also with [**?**] or [**?**, p. 140]. ■

Lemma 8.2 (Volumetric terms)
Set $C = F^TF$. The following terms are each polyconvex and have each a stress free reference configuration:

1. $F \mapsto \left(\det[C] + \frac{1}{\det[C]} - 2\right)^k$, $k \geq 1$.
2. $F \mapsto \left((\det[C])^p + \frac{1}{(\det[C])^p} - 2\right)^k$, $p \geq k \geq 1, p \geq \frac{1}{2}$.

3. $F \mapsto (\sqrt{\det[C]} - 1)^k,\ k \geq 1.$

4. $F \mapsto (\det[C] - \ln\det[C]).$

5. $F \mapsto (\det[C] - \ln\det[C] + (\ln\det[C])^2).$

Proof. On the natural domain of definition $\det[F] > 0$ the given functions are convex in the variable $\det[F]$. ■

Lemma 8.3 (Convex terms)
Let $X \in \mathbb{M}^{3\times 3}$. Then the following terms are each convex:

1. $X \mapsto [\operatorname{tr}(X^TX)]^k,\ k \geq 1.$

2. $X \mapsto \operatorname{tr}((X^TX)^k),\ k \geq 1.$

and the statements remain true if X is changed into X^T since linear transformations leave convexity properties invariant.

Proof.

1. $[\operatorname{tr}(X^TX)]^k = \|X\|^{2k}$. We compute the second differential:

$$\begin{aligned} D_X\left(\|X\|^{2k}\right).H &= 2k \cdot \|X\|^{2k-2}\langle X, H\rangle \\ D_X^2\left(\|X\|^{2k}\right).(H,H) &= 2k \cdot \|X\|^{2k-2}\langle H, H\rangle + 2k\ (2k-2)\ \|X\|^{2k-4}\langle X, H\rangle^2 > 0. \end{aligned}$$

2. $\operatorname{tr}((X^TX)^k)$. Set $C = X^TX$. Then

$$\begin{aligned} D_C\left(\operatorname{tr}(C^k)\right).H &= D_C\left(\langle C^k, 1\!\!1\rangle\right).H \\ &= k \cdot \langle C^{k-1}, H\rangle \\ D_C^2\left(\operatorname{tr}(C^k)\right).(H,H) &= k(k-1)\langle C^{k-2}H, H\rangle \geq 0. \end{aligned}$$

   Thus $D_C\left(\operatorname{tr}(C^k)\right) = kC^{k-1} \in \mathrm{PSym}$ and $D_C^2\left(\operatorname{tr}(C^k)\right).(H,H) \geq 0$ which allows us to apply Lemma **??**.

A change of X into X^T is permited: convexity is invariant under the transformation $X \mapsto X^T$. ■

Lemma 8.4 (Generic polyconvex terms)
Let $F \in \mathbb{M}^{3\times 3}$. Then the following terms are each polyconvex:

1. $\frac{[\operatorname{tr}(F^TF)]^k}{\det[(F^TF)]^{\frac{1}{3}}},\ k \geq 1.$

2. $\frac{[\operatorname{tr}(\operatorname{Adj}(F^TF))]^k}{\det[(F^TF)]^{\frac{1}{3}}},\ k \geq 1.$

Proof.

1.

$$\frac{[\mathrm{tr}(F^TF)]^k}{\det[(F^TF)]^{\frac{1}{3}}} = \frac{\|F\|^{2k}}{\det[F]^{\frac{2}{3}}}$$

We have already shown (see (**??**)) that the function $P(x,y) = \frac{1}{x^\alpha} \cdot y^p$ is convex provided that $\alpha = \frac{2}{3}$ and $p = 2k \geq 2$. Now define a new function

$$\hat{W}(F,\zeta) := P(\zeta, \|F.\eta\|) = \frac{\|F\|^{2k}}{\zeta^{\frac{2}{3}}}.$$

Observe that by the monotonocity of the $2k$-th power for positive arguments we have the inequality

$$\|\lambda F_1 + (1-\lambda)F_2\|^{2k} \leq (\lambda\|F_1\| + (1-\lambda)\|F_2\|)^{2k} \tag{8.4}$$

It remains to check the convexity of $\hat{W}(F,\zeta)$. To this end

$$\begin{aligned}\hat{W}(\lambda F_1 + (1-\lambda)F_2, \lambda\zeta_1 + (1-\lambda)\zeta_2) &= P(\lambda\zeta_1 + (1-\lambda)\zeta_2, \|\lambda F_1 + (1-\lambda)F_2\|) \\ &= \frac{\|\lambda F_1 + (1-\lambda)F_2\|^{2k}}{(\lambda\zeta_1 + (1-\lambda)\zeta_2)^{\frac{2}{3}}}.\end{aligned}$$

With (**??**) we have

$$\begin{aligned}\hat{W}(\lambda F_1 + (1-\lambda)F_2, \lambda\zeta_1 + (1-\lambda)\zeta_2) &\leq \frac{(\lambda\|F_1\| + (1-\lambda)\|F_2\|)^{2k}}{(\lambda\zeta_1 + (1-\lambda)\zeta_2)^{\frac{2}{3}}} \\ &= P(\lambda\zeta_1 + (1-\lambda)\zeta_2, \lambda\|F_1\| + (1-\lambda)\|F_2\|).\end{aligned}$$

The convexity of P yields

$$\begin{aligned}\hat{W}(\lambda F_1 + (1-\lambda)F_2, \lambda\zeta_1 + (1-\lambda)\zeta_2) &\leq \lambda P(\zeta, \|F_1\|) + (1-\lambda)P(\zeta_2, \|F_2\|) \\ &= \lambda\hat{W}(F_1,\zeta_1) + (1-\lambda)\hat{W}(F_2,\zeta_2).\end{aligned}$$

The proof is finished.

2. $\frac{\mathrm{tr}(\mathrm{Adj}(F^TF))^k}{\det[(F^TF)]^{\frac{1}{3}}} = \frac{\|\,\mathrm{Adj}\,F\|^{2k}}{\det[(F)]^{\frac{2}{3}}}$ and we may proceed like in case one.

■

Corollary 8.5 (Generic exponential polyconvex terms)
Let $F \in \mathbb{M}^{3\times3}$. Then the following terms are each polyconvex:

1. $\exp\left[\frac{\mathrm{tr}(F^TF)^k}{\det[(F^TF)]^{\frac{1}{3}}}\right]$, $k \geq 1$.
2. $\exp\left[\frac{\mathrm{tr}(\mathrm{Adj}(F^TF))^k}{\det[(F^TF)]^{\frac{1}{3}}}\right]$, $k \geq 1$.

3. $\exp\left[W(F)\right]$ *if* $W(F)$ *is polyconvex.*

Proof. By the foregoing lemma each argument of the exponential is polyconvex. Since exp is convex and monotone increasing it preserves the underlying convexity. Hence the composition is polyconvex. Observe however, that these functions alone are not stress free in the reference configuration. ■

Lemma 8.6 (Special polyconvex terms)
Let $F \in \mathbb{M}^{3\times 3}$*. Then the following terms are each polyconvex as functions* $F \mapsto \mathbb{R}$*:*

1. $F \mapsto \left(\frac{\|F\|^2}{\det[F]^{\frac{2}{3}}} - 3\right)^i = \left(\operatorname{tr}\left(\frac{C}{\det[C]^{\frac{1}{3}}}\right) - 3\right)^i, \; i \geq 1.$

2. $F \mapsto \left(\frac{\|\operatorname{Adj} F\|^3}{\det[F]^2} - 3\sqrt{3}\right)^j = \left(\operatorname{tr}\left(\operatorname{Adj}\left(\frac{C}{\det[C]^{\frac{1}{3}}}\right)\right)^{\frac{3}{2}} - 3\sqrt{3}\right)^j, \; j \geq 1.$

Proof.

1. We have already checked in Lemma **??** that the expression $\frac{\|F\|^2}{\det[F]^{\frac{2}{3}}}$ is polyconvex, hence there exists a convex function $P(F, \det[F]) = \frac{\|F\|^2}{\det[F]^{\frac{2}{3}}}$. Observe that by the estimates for the invariants Lemma **??** we know that $P(F, \det[F]) - 3 \geq 0$. We define the function $[a]_+ = \max\{a, 0\}$. Then

$$\left(\frac{\|F\|^2}{\det[F]^{\frac{2}{3}}} - 3\right)^i = [P\underbrace{(F, \det[F])}_{X\in\mathbb{R}^{10}} - 3]_+^i.$$

P is convex in F and $x \mapsto x^i$, $i \geq 1$ is monotone increasing for positive values and convex, hence

$$[P(X) - 3]_+^i$$

is altogether convex in X which is however the polyconvexity of $F \mapsto [P(F, \det[F]) - 3]_+^i$. Since this last expression coincides with

$$\left(\frac{\|F\|^2}{\det[F]^{\frac{2}{3}}} - 3\right)^i$$

the polyconvexity is proved. ■

2. We know already that $\frac{\|\operatorname{Adj} F\|^3}{\det[F]^2} - 3\sqrt{3}$ is polyconvex since the exponents verify the decisive inequality $\frac{\alpha+1}{\alpha} \geq \frac{p}{p-1}$. Moreover, $\frac{\|\operatorname{Adj} F\|^3}{\det[F]^2} - 3\sqrt{3} \geq 0$ with Lemma **??**. Now exactly the same reasoning as before applies.

Corollary 8.7
Let $F \in \mathbb{M}^{3\times3}$. Then the following more general terms are each polyconvex:

1. $F \mapsto \left(\frac{\|F\|^{2k}}{\det[F]^{\frac{2k}{3}}} - 3^k\right)^i, \; i \geq 1, \; k \geq 1.$
2. $F \mapsto \left(\frac{\|\operatorname{Adj} F\|^{3k}}{\det[F]^{2k}} - (3\sqrt{3})^k\right)^j, \; j \geq 1, \; k \geq 1.$
3. $F \mapsto \exp\left[\left(\frac{\|F\|^{2k}}{\det[F]^{\frac{2k}{3}}} - 3^k\right)^i\right] - 1, \; i \geq 1, \; k \geq 1.$
4. $F \mapsto \exp\left[\left(\frac{\|\operatorname{Adj} F\|^{3k}}{\det[F]^{2k}} - (3\sqrt{3})^k\right)^j\right] - 1, \; j \geq 1, \; k \geq 1.$

Proof. Apply the same ideas as above and observe that exp is a convex monotone increasing function, so that we may apply Lemma **??**. ■

Remark 8.8
The above isotropic terms of the type

$$W(F) = \left(\frac{\|F\|^2}{\det[F]^{\frac{2}{3}}} - 3\right)^i, \; i \geq 1$$

have the nice special property that $W(\mathbb{1}) = 0$ and $W(F) \geq 0$. Hence the reference configuration is automatically stress free. This contrasts known polyconvex functions like Ogden-materials, where only by a judicious choice of parameters the reference configuration can be made stress free. We know of no other polyconvex energies which join this feature. The polyconvexity of these terms seems to be gone unnoticed. Of course, the terms are objective and meet various growth conditions which are necessary for the succesful application of the direct methods of variations to prove the existence of solutions to a corresponding finite elasticity problem.

Remark 8.9 (Non ellipticity of mixed terms)
The following terms are non-elliptic hence not polyconvex:

$$W(F) = \left(\frac{\|F\|^2}{\det[F]^{\frac{2}{3}}} - 3\right)^i \cdot \left(\frac{\|\operatorname{Adj} F\|^3}{\det[F]^2} - 3\sqrt{3}\right)^j, \quad i, j \geq 1\,.$$

Moreover, the term

$$F \mapsto \left(\frac{\|\operatorname{Adj} F\|^2}{\det[F]^{\frac{4}{3}}} - 3\right)^i, i \geq 1$$

is non-elliptic, hence cannot be polyconvex, even so $\frac{\|\operatorname{Adj} F\|^2}{\det[F]^{\frac{4}{3}}} - 3 \geq 0$ in light of Lemma **??**. *Here the term $\frac{\|\operatorname{Adj} F\|^2}{\det[F]^{\frac{4}{3}}}$ itself does not have the right exponents [**?**] for polyconvexity.*

Proof. We let $i, j = 1$ and consider the eigenvalue representation of $W(F)$:

$$W(F) = \left(\frac{\lambda_1^2+\lambda_2^2+\lambda_3^2}{(\lambda_1\lambda_2\lambda_3)^{\frac{2}{3}}} - 3\right)^i \left(\frac{(\lambda_1\lambda_2)^2+(\lambda_2\lambda_3)^2+(\lambda_1\lambda_3)^2}{(\lambda_1\lambda_2\lambda_3)^2} - 3\sqrt{3}\right)^j = \Phi(\lambda_1,\lambda_2,\lambda_3).$$

We take the extreme deformation $F = \operatorname{diag}(0.1, 10, t)$ with $t \in \mathbb{R}^+$. If $W(F)$ is rank-one convex, then

$$\Phi(0.1,10,t) = \left(\frac{100.01+t^2}{t^{\frac{2}{3}}} - 3\right)^i \cdot \left(\frac{(1+100.01t^2}{t^2} - 3\sqrt{3}\right)^j$$

should be convex, according to Theorem **??**. However, this is not the case as can easily be verified. Typically, convexity in t (hence ellipticity with respect to F) is lost for extreme deformations only. ■

9 Coercivity for new isotropic formulation

Coercivity is a condition on the growth of the free energy for large deformation gradients. It is a necessary part of the existence proof via the direct methods of variations. More precisely we have

Definition 9.1 (Coercivity)
Let $I(u) = \int_\Omega W(\nabla u)\,\mathrm{dx}$ *for* $W : \mathbb{M}^{3\times3} \mapsto \mathbb{R}$ *be the elastic stored energy functional. We say that* I *is* q*-coercive, whenever*

$$I(u) \le K \Longrightarrow \|u\|_{1,q,\Omega} \le \tilde{K}\,.$$

Lemma 9.2 (Coercivity of special energy)
Let the elastic stored energy density be given by

$$\begin{aligned} W(F) &= \alpha^+ \left(\left[\frac{\|F\|^2}{\det[F]^{\frac{2}{3}}}\right]^3 - 3^3\right) + \beta^+ \left(\frac{\|\operatorname{Adj} F\|^3}{\det[F]^2} - 3\sqrt{3}\right) + \\ &\qquad \delta\left(\det[F]^4 + \frac{1}{\det[F]^4} - 2\right) + c_1\ (\ldots) + c_2\ (\ldots) \\ &= \alpha^+ \left(\left[\operatorname{tr}\left(\frac{C}{\det[C]^{\frac{1}{3}}}\right)\right]^3 - 3^3\right) + \beta^+ \left(\left[\operatorname{tr}\left(\operatorname{Adj}\left(\frac{C}{\det[C]^{\frac{1}{3}}}\right)\right)\right]^{\frac{3}{2}} - 3\sqrt{3}\right) + \\ &\qquad \delta\left(\det[C]^2 + \frac{1}{\det[C]^2} - 2\right) + c_1\ (\ldots) + c_2\ (\ldots)\,. \end{aligned}$$

Then

$$I(u) = \int_\Omega W(\nabla u)\ \mathrm{dx}$$

is coercive for $q = 4$.

Proof.

$$
\begin{aligned}
F &= \frac{F}{\det[F]^{\frac{1}{3}}} \cdot (\det[F]^{\frac{1}{3}})\,, \\
\|F\|^q_{q,\Omega} &= \|\frac{F}{\det[F]^{\frac{1}{3}}} \cdot (\det[F]^{\frac{1}{3}})\|^q_{q,\Omega} \\
&= \int_\Omega \|\frac{F}{\det[F]^{\frac{1}{3}}}\|^q \cdot |\det[F]|^{\frac{q}{3}} \ \mathrm{dx} \\
&\text{apply Youngs inequality with } \frac{1}{a} + \frac{1}{b} = 1 \\
&\le \int_\Omega \frac{1}{a}\|\frac{F}{\det[F]^{\frac{1}{3}}}\|^{qa} + \frac{1}{b}|\det[F]|^{\frac{qb}{3}} \ \mathrm{dx} \\
&\text{taking } a = \frac{3}{2},\ b = 3 \text{ yields} \\
&= \int_\Omega \frac{2}{3}\|\frac{F}{\det[F]^{\frac{1}{3}}}\|^{\frac{3q}{2}} + \frac{1}{3}|\det[F]|^q \ \mathrm{dx} \\
&\text{for } q = 4 \text{ this shows} \\
&= \int_\Omega \frac{2}{3}\|\frac{F}{\det[F]^{\frac{1}{3}}}\|^6 + \frac{1}{3}|\det[C]|^2 \ \mathrm{dx} = \int_\Omega \frac{2}{3}\left[\frac{\|F\|^2}{\det[F]^{\frac{2}{3}}}\right]^3 + \frac{1}{3}|\det[C]|^2 \ \mathrm{dx} \\
&\le \frac{2}{3\alpha}\ I(u) + 3^3 + \frac{1}{3\delta}\ I(u) + 2 \\
&\le \left(\frac{2}{3\alpha} + \frac{1}{3\delta}\right)\ I(u) + 2 + 3^3\,.
\end{aligned}
$$

An application of Poincaré's inequality will end the proof. ■

Theorem 9.3 (Existence of minimizers)
Let the reference configuration $\Omega \subset \mathbb{R}^3$ be a bounded smooth domain and let $\partial\Omega_1$ be a part of the boundary $\partial\Omega$ with nonvanishing Lebesgue measure. Assume that $I(\phi) = \int_\Omega W(\nabla\phi(x))\ \mathrm{dx}$ with W as in Lemma **??** *Let $\phi_0 \in W^{1,4}(\Omega)$ be given with $I(\phi_0) < \infty$. Then the problem*

$$\inf\{\ I(\phi) = \int_\Omega W(\nabla\phi(x))\ \mathrm{dx},\ \ \phi(x) = \phi_0(x),\ x \in \partial\Omega_1,\ \phi \in W^{1,4}(\Omega)\}$$

admits at least one solution. Formally, this solution corresponds to a solution of the boundary value problem

$$
\begin{aligned}
\operatorname{Div} D_F W(\nabla\phi) &= 0\,, \\
\phi(x) &= \phi_0(X),\ x \in \partial\Omega_1\,.
\end{aligned}
$$

Proof. The energy is polyconvex and coercivity in $W^{1,4}(\Omega)$ has been shown in Lemma **??**. Since $I(\phi) \ge 0$ and $I(\phi_0) < \infty$ the infimum exists and the direct methods of variation yield the existence of at least one minimizer. ■

It is our aim to express the parameters $\alpha, \beta, \delta > 0$ in terms of the Lamé constants μ, λ. In fact, by a judicious choice of these parameters all Lamé moduli of the infinitesimal theory can be represented.

Lemma 9.4 (Stress free natural state)
Let W be an isotropic free energy function with stress free natural state, then W admits the following expansion near to $1\!\!1$

$$\begin{aligned} W(F) = W(C) &= \mu \, \|E\|^2 + \frac{\lambda}{2} \, \operatorname{tr}(E)^2 + o(\|E\|^2) \\ &= \frac{\mu}{4} \, \|C - 1\!\!1\|^2 + \frac{\lambda}{8} \, \operatorname{tr}(C - 1\!\!1)^2 + o(\|C - 1\!\!1\|^2) \, , \end{aligned}$$

where $E = \frac{1}{2}(C - 1\!\!1) = \frac{1}{2}(F^T F - 1\!\!1)$ and $\lambda, \mu > 0$ are the so called Lamé-constants.

Proof. See e.g., Ciarlet [?]. ■

Corollary 9.5 (Expansion)

$$\begin{aligned} W(C) &= W(1\!\!1 + (C - 1\!\!1)) \\ &= W(1\!\!1) + D_C W(C_{|1\!\!1}).[C - 1\!\!1] + \frac{1}{2} D_C^2 W(C_{|1\!\!1}).[(C - 1\!\!1), (C - 1\!\!1)] + \dots \\ &= 0 + 0 + \frac{1}{2} D_C^2 W(C_{|1\!\!1}).[(C - 1\!\!1), (C - 1\!\!1)] + \dots \end{aligned}$$

Hence equating like powers shows

$$\begin{aligned} \frac{1}{2} D_C^2 W(1\!\!1).[H, H] &= \frac{\mu}{4} \, \|H\|^2 + \frac{\lambda}{8} \, \operatorname{tr}(H)^2 \\ D_C^2 W(1\!\!1).[H, H] &= \frac{\mu}{2} \, \|H\|^2 + \frac{\lambda}{4} \, \operatorname{tr}(H)^2 . \end{aligned}$$ ■

This forces us to calculate the second differential in $1\!\!1$ of

$$\begin{aligned} W(C) = \alpha^+ & \left(\left[\operatorname{tr}(\frac{C}{\det[C]^{\frac{1}{3}}}) \right]^3 - 3^3 \right) + \beta^+ \left(\left[\operatorname{tr}(\operatorname{Adj}(\frac{C}{\det[C]^{\frac{1}{3}}})) \right]^{\frac{3}{2}} - 3\sqrt{3} \right) + \\ & \delta \left(\det[C]^2 + \frac{1}{\det[C]^2} - 2 \right) + c_1 \, (\dots)^2 + c_2 \, (\dots)^2 + \dots \end{aligned}$$

It is obvious that the higher order terms do not yield any contribution as far as a linearization with respect to $1\!\!1$ is concerned since the brackets itself are stress-free. Thus we are left with

$$\begin{aligned}W(C) &= \alpha^+ \left(\left[\operatorname{tr}(\frac{C}{\det[C]^{\frac{1}{3}}})\right]^3 - 3^3\right) + \beta^+ \left(\left[\operatorname{tr}(\operatorname{Adj}(\frac{C}{\det[C]^{\frac{1}{3}}})\right]^{\frac{3}{2}} - 3\sqrt{3}\right) + \\ &\quad \delta \left(\det[C]^2 + \frac{1}{\det[C]^2} - 2\right) \\ &= \alpha^+ \left(\left[\operatorname{tr}(\frac{C}{\det[C]^{\frac{1}{3}}})\right]^3 - 3^3\right) + \beta^+ \left(\frac{\operatorname{tr}(\operatorname{Adj} C)^{\frac{3}{2}}}{\det[C]} - 3\sqrt{3}\right) + \\ &\quad + \delta \left(\det[C]^2 + \frac{1}{\det[C]^2} - 2\right)\end{aligned}$$

We consider first the contribution of the α-term. The first and second differential are

$$\begin{aligned}D_C W_\alpha(C).H &= 3\left[\operatorname{tr}(\frac{C}{\det[C]^{\frac{1}{3}}})\right]^2 \left[\langle \mathbb{1}, H\rangle \det[C]^{\frac{-1}{3}} + \operatorname{tr}(C)\frac{-1}{3}\det[C]^{\frac{-2}{3}}\langle \operatorname{Adj} C, H\rangle\right] \\ D_C^2 W_\alpha(C).(H,H) &= 6\left[\operatorname{tr}(\frac{C}{\det[C]^{\frac{1}{3}}})\right] \cdot \left[\langle \mathbb{1}, H\rangle \det[C]^{\frac{-1}{3}} + \operatorname{tr}(C)\frac{-1}{3}\det[C]^{\frac{-2}{3}}\langle \operatorname{Adj} C, H\rangle\right]^2 + \\ &\quad 3\left[\operatorname{tr}(\frac{C}{\det[C]^{\frac{1}{3}}})\right]^2 \cdot \Bigg(\langle \mathbb{1}, H\rangle \frac{-1}{3}\det[C]^{\frac{-2}{3}}\langle \operatorname{Adj} C, H\rangle - \frac{1}{3}(\langle \mathbb{1}, H\rangle \det[C]^{\frac{-2}{3}}\langle \operatorname{Adj} C, H\rangle + \\ &\quad \operatorname{tr}(C)\frac{-2}{3}\det[C]^{\frac{-5}{3}}\langle \operatorname{Adj} C, H\rangle^2 + \operatorname{tr}(C)\det[C]^{\frac{-2}{3}}\langle D \operatorname{Adj} C.H, H\rangle)\Bigg)\end{aligned}$$

Thus

$$\begin{aligned}&D_C^2 W_\alpha(\mathbb{1}).(H,H) = 6 \cdot 3\left[\langle \mathbb{1}, H\rangle 1 + 3\frac{-1}{3}1\langle \mathbb{1}, H\rangle\right]^2 + \\ &3 \cdot 3^2 \left(\langle \mathbb{1}, H\rangle \frac{-1}{3}1\langle \mathbb{1}, H\rangle - \frac{1}{3}(\langle \mathbb{1}, H\rangle 1 \langle \mathbb{1}, H\rangle + 3\frac{-2}{3}1\langle \mathbb{1}, H\rangle^2 + 3 \cdot 1(\langle \mathbb{1}, H\rangle^2 - \|H\|^2)\right) \\ &= 18 \cdot 0 + 27\left(\|H\|^2 - \langle \mathbb{1}, H\rangle^2\right) .\end{aligned}$$

Next consider the contribution of the β-term.

$$\begin{aligned}D_C W_\beta(C).H &= \frac{3}{2}\operatorname{tr}(\operatorname{Adj} C)^{\frac{1}{2}} \operatorname{tr}(D. \operatorname{Adj} C.H)\det[C]^{-1} + \operatorname{tr}(\operatorname{Adj} C)^{\frac{3}{2}}(-1)\det[C]^{-2}\langle \operatorname{Adj} C, H\rangle \\ D_C^2 W_\beta(C).(H,H) &= \frac{3}{2}\frac{1}{2}\operatorname{tr}(\operatorname{Adj} C)^{\frac{-1}{2}} \operatorname{tr}(D. \operatorname{Adj} C.H)\det[C]^{-1} + \\ &\quad \frac{3}{2}\operatorname{tr}(\operatorname{Adj} C)^{\frac{1}{2}} \operatorname{tr}(D^2. \operatorname{Adj} C.(H,H))\det[C]^{-1} + \\ &\quad \frac{3}{2}\operatorname{tr}(\operatorname{Adj} C)^{\frac{1}{2}} \operatorname{tr}(D. \operatorname{Adj} C.H)(-1)\det[C]^{-2}\langle \operatorname{Adj} C, H\rangle + \\ &\quad \frac{3}{2}\operatorname{tr}(\operatorname{Adj} C)^{\frac{1}{2}} \langle D \operatorname{Adj} C.H, \mathbb{1}\rangle(-1)\det[C]^{-2}\langle \operatorname{Adj} C, H\rangle + \\ &\quad \operatorname{tr}(\operatorname{Adj} C)^{\frac{3}{2}} 2\det[C]^{-3}\langle \operatorname{Adj} C, H\rangle^2 + \operatorname{tr}(\operatorname{Adj} C)^{\frac{3}{2}}(-1)\det[C]^{-2}\langle D \operatorname{Adj} C.H, H\rangle.\end{aligned}$$

Whence

$$\begin{aligned}
D^2_C W_\beta(1\!\!1).(H,H) &= \frac{3}{4}\frac{1}{\sqrt{3}}\,\mathrm{tr}(D.\operatorname{Adj} C.H)1 + \frac{3}{2}\sqrt{3}2\langle \operatorname{Adj} H, 1\!\!1\rangle + \frac{3}{2}\sqrt{3}\,\mathrm{tr}(D.\operatorname{Adj} C.H)(-1)\langle 1\!\!1, H\rangle + \\
&\quad \frac{3}{2}\sqrt{3}\langle D\operatorname{Adj} C.H, 1\!\!1\rangle(-1)\langle 1\!\!1, H\rangle 3^{\frac{3}{2}}2\langle 1\!\!1, H\rangle^2 + 3^{\frac{3}{2}}(-1)\langle D\operatorname{Adj} C.H, H\rangle \\
&= \frac{3}{4}\frac{1}{\sqrt{3}}2\langle 1\!\!1, H\rangle + \frac{3}{2}\sqrt{3}(\langle 1\!\!1, H\rangle^2 - \|H\|^2) + \frac{3}{2}\sqrt{3}2\langle 1\!\!1, H\rangle(-1)\langle 1\!\!1, H\rangle + \\
&\quad \frac{3}{2}\sqrt{3}2\langle 1\!\!1, H\rangle(-1)\langle 1\!\!1, H\rangle 3^{\frac{3}{2}}2\langle 1\!\!1, H\rangle^2 + 3^{\frac{3}{2}}(-1)(\langle 1\!\!1, H\rangle^2 - \|H\|^2) \\
&= \frac{\sqrt{3}}{2}\left(3\|H\|^2 - \langle 1\!\!1, H\rangle^2\right)\,.
\end{aligned}$$

Finally, the δ-term

$$\begin{aligned}
D_C W_\delta(C).H &= 2\det[C]\langle \operatorname{Adj} C, H\rangle - 2\det[C]^{-3}\langle \operatorname{Adj} C, H\rangle \\
D^2_C W_\delta(C).(H,H) &= 2\langle \operatorname{Adj} C, H\rangle^2 + 2\det[C]\langle D\operatorname{Adj} C.H, H\rangle \\
&\quad - 2\left(-3\det[C]^{-4}\langle \operatorname{Adj} C, H\rangle^2 + \det[C]^{-3}\langle D\operatorname{Adj} C.H, H\rangle\right) \\
D^2_C W_\delta(1\!\!1).(H,H) &= 8\langle 1\!\!1, H\rangle^2
\end{aligned}$$

Thus we get altogether

$$\begin{aligned}
D^2_C W(1\!\!1).(H,H) &= 27\alpha(\|H\|^2 - \langle 1\!\!1, H\rangle^2) + \frac{\sqrt{3}}{2}\beta\left(3\|H\|^2 - \langle 1\!\!1, H\rangle^2\right) + 8\delta\langle 1\!\!1, H\rangle^2 \\
&= (27\alpha + \frac{3\sqrt{3}}{2}\beta)\|H\|^2 + (8\delta - \frac{\sqrt{3}}{2}\beta - 27\alpha)\langle 1\!\!1, H\rangle^2 \\
&= \frac{\mu}{2}\|H\|^2 + \frac{\lambda}{4}\langle 1\!\!1, H\rangle^2
\end{aligned}$$

It remains to determine positive parameters $\alpha, \beta, \delta > 0$ such that

$$\begin{aligned}
\frac{\mu}{2} &= 27\alpha + \frac{3\sqrt{3}}{2}\beta \\
\frac{\lambda}{4} &= 8\delta - \frac{\sqrt{3}}{2}\beta - 27\alpha.
\end{aligned}$$

Adding both equations and multiplying by 4 yields

$$\beta = \frac{1}{4\sqrt{3}}(2\mu + \lambda) - 32\delta\,.$$

We choose $\delta = (2\mu + \lambda)\cdot t$ with $t \in \mathbb{R}^+$. Thus $\beta = \frac{1}{4\sqrt{3}}(2\mu + \lambda)[1 - 32t] > 0$. Reinserting into the first equation shows

$$\alpha = \frac{\mu}{2\cdot 27} - \frac{3}{16}(2\mu + \lambda)(1 - 32t) > 0\,.$$

This is true if

$$\frac{8}{3\cdot 27}\frac{\mu}{2\mu + \lambda} > 1 - 32t\,.$$

This is met for e.g. $1 - 32t = \frac{27}{32}\frac{8}{3\cdot 27}\frac{\mu}{2\mu+\lambda}$. This implies $t = \frac{23\mu + 12\lambda}{32\cdot 12(2\mu+\lambda)}$. Hence $\delta = \frac{23\mu+12\lambda}{32\cdot 12}$ and $\beta = \frac{\mu}{48\sqrt{3}}$ and $\alpha = \frac{5\mu}{64\cdot 27}$.

10 The anisotropic polyconvex setting

Lemma 10.1 (Properties of the anisotropy structural tensor M)
Let $\eta \in \mathbb{R}^3$ with $\|\eta\| = 1$ and define $M = \eta \otimes \eta$. Then the following statements hold:

1. $M^T = M$.
2. *M is positive semi-definite.*
3. $M^T M = M$.
4. $\operatorname{tr}(M) = 1$.
5. $M^2 = M$.
6. $\|M\|^2 = 1$.
7. $\|1\!\!1 - M\|^2 = 2$.
8. $(1\!\!1 - M)(1\!\!1 - M) = 1\!\!1 - M$.
9. $(1\!\!1 - M)^T(1\!\!1 - M) = 1\!\!1 - M$.
10. *$(1\!\!1 - M)$ is positive semi-definite.*
11. $rank(M) = 1$.
12. $\operatorname{Adj} M = 0$.
13. $\det[M] = 0$.
14. *$rank(1\!\!1 - M) = 2$ and $\operatorname{Adj}(1\!\!1 - M) \neq 0$.*
15. $\operatorname{Adj}(1\!\!1 - M) = M$.
16. $\langle H, H.M\rangle \geq 0$.

Lemma 10.2 (Convex anisotropic terms)
Let $X \in \mathbb{M}^{3\times 3}$ and $M = \eta \otimes \eta$. Then the following terms are each convex as functions in X:

1. $X \mapsto [\operatorname{tr}(X^T X \cdot M)]^k, \ k \geq 1$.
2. $X \mapsto [\operatorname{tr}(X^T X \cdot (1\!\!1 - M))]^k, \ k \geq 1$.
3. $X \mapsto [\operatorname{tr}(X^T X \cdot M \cdot X^T X \cdot M)]^k, \ k \geq 1$.
4. $X \mapsto [\operatorname{tr}(X^T X)]^2 + \operatorname{tr}(X^T X) \cdot \operatorname{tr}(X^T X M)$.
5. $X \mapsto 2[\operatorname{tr}(X^T X)]^2 + \operatorname{tr}(X^T X) \cdot \operatorname{tr}(X^T X(1\!\!1 - M))$.
6. $X \mapsto \frac{1}{2}[\operatorname{tr}(X^T X)]^2 + \operatorname{tr}(X^T X X^T X M)$,

and the statements remain true if X is changed into X^T since linear transformations leave convexity properties invariant.

Proof.

1. $[\mathrm{tr}(X^T X \cdot M)]^k = \langle X, X \cdot M\rangle^k$. We compute the second differential:

$$\begin{aligned} D_X\left(\langle X, XM\rangle^k\right).H &= k\langle X, XM\rangle^{k-1}\left(\langle X, HM\rangle + \langle H, XM\rangle\right) \\ &= 2k\langle X, XM\rangle^{k-1}\langle X, HM\rangle \\ D_X^2\left(\langle X, XM\rangle^k\right).(H,H) &= 4k(k-1)\langle X, XM\rangle^{k-2} \cdot \langle XM, H\rangle^2 + \\ & \quad 2k\langle X, XM\rangle^{k-1}\langle H, HM\rangle \geq 0. \end{aligned}$$

   in this context see also **??** Equation 15.

2. $[\mathrm{tr}(X^T X \cdot (\mathbb{1} - M))]^k = \left(\|X\|^2 - \|X.\eta\|^2\right)^k$. We may apply the same reasoning as in the previous line. Observe that

$$\left(\|X\|^2 - \|X.\eta\|^2\right) \geq 0 \quad \text{if} \quad \|\eta\| = 1.$$

3.

$$\begin{aligned} (\mathrm{tr}\,[[]X^T X \cdot M \cdot X^T X \cdot M])^k &= \langle X^T X M, M X^T X\rangle^k \\ &= \langle X^T X(\eta \otimes \eta), (\eta \times \eta)X^T X\rangle^k \\ &= \langle X^T(X.\eta \otimes \eta), (\eta \otimes X.\eta)X\rangle^k \\ &= \langle (X.\eta \otimes \eta)X^T, X(\eta \otimes X.\eta)\rangle^k \\ &= \langle (X.\eta \otimes X.\eta), (X.\eta \otimes X.\eta)\rangle^k \\ &= \|(X.\eta \otimes X.\eta)\|^{2k} \\ &= \|X.\eta\|^{4k}. \end{aligned}$$

   Hence, computing the differentials yields

$$\begin{aligned} D_X\left(\|X.\eta\|^{4k}\right).H &= 4k \cdot \|X.\eta\|^{4k-2} \cdot \langle X.\eta, H.\eta\rangle \\ D_X^2\left(\|X.\eta\|^{4k}\right).(H,H) &= 4k\ (4k-2)\ \|X.\eta\|^{4k-4}\langle X.\eta, H.\eta\rangle^2 + \\ & \quad 4k\ \|X.\eta\|^{4k-2}\langle H.\eta, H.\eta\rangle \geq 0. \end{aligned}$$

4. $\mathrm{tr}(X^T X)^2 + \mathrm{tr}(X^T X) \cdot \mathrm{tr}(X^T X M) = \|X\|^4 + \|X\|^2 \cdot \|X.\eta\|^2$. We calculate the second

differential, which yields

$$\begin{aligned}
&D_X^2\left(\|X\|^4+\|X\|^2\cdot\|X.\eta\|^2\right).(H,H)=\\
&8\langle X,H\rangle^2+4\|X\|^2\|H\|^2+8\langle X,H\rangle\langle X.\eta,H.\eta\rangle+\\
&\quad 2\|X.\eta\|^2\|H\|^2+2\|X\|^2\|H.\eta\|^2\\
&\geq 8\langle X,H\rangle^2+4\|X\|^2\|H\|^2-8(\|X\|\|H.\eta\|)(\|H\|\|X.\eta\|)+\\
&\quad 2\|X.\eta\|^2\|H\|^2+2\|X\|^2\|H.\eta\|^2\\
&\geq 8\langle X,H\rangle^2+4\|X\|^2\|H\|^2-4\|X\|^2\|H.\eta\|^2-4\|H\|^2\|X.\eta\|^2+\\
&\quad 2\|X.\eta\|^2\|H\|^2+2\|X\|^2\|H.\eta\|^2\\
&\geq 8\langle X,H\rangle^2+4\|X\|^2\|H\|^2-2\|X\|^2\|H.\eta\|^2-2\|H\|^2\|X.\eta\|^2\\
&\geq 8\langle X,H\rangle^2\geq 0
\end{aligned}$$

where we have used Young's inequality.

5. $X\mapsto 2[\mathrm{tr}(X^TX)]^2+\mathrm{tr}(X^TX)\cdot\mathrm{tr}(X^TX(\mathbb{1}-M))=2\|X\|^4+\|X\|^2\cdot\|X(\mathbb{1}-M)\|^2$. We calculate the second differential, which yields

$$\begin{aligned}
&D_X^2\left(2\|X\|^4+\|X\|^2\cdot\|X(\mathbb{1}-M)\|^2\right).(H,H)=\\
&16\langle X,H\rangle^2+8\|X\|^2\|H\|^2+8\langle X,H\rangle\langle X(\mathbb{1}-M),H(\mathbb{1}-M)\rangle+\\
&\quad 2\|X(\mathbb{1}-M)\|^2\|H\|^2+2\|X\|^2\|H(\mathbb{1}-M\|^2\\
&\geq 16\langle X,H\rangle^2+8\|X\|^2\|H\|^2-8\|X\|\|H\|\|X(\mathbb{1}-M)\|\|H(\mathbb{1}-M)\|+\\
&\quad 2\|X(\mathbb{1}-M)\|^2\|H\|^2+2\|X\|^2\|H(\mathbb{1}-M\|^2\\
&\geq 16\langle X,H\rangle^2+8\|X\|^2\|H\|^2-4\|X\|^2\|H(\mathbb{1}-M)\|^2-4\|H\|^2\|X(\mathbb{1}-M)\|^2+\\
&\quad 2\|X(\mathbb{1}-M)\|^2\|H\|^2+2\|X\|^2\|H(\mathbb{1}-M\|^2\\
&\geq 16\langle X,H\rangle^2+8\|X\|^2\|H\|^2-4\|X\|^2\|H\|^2\|(\mathbb{1}-M)\|^2-4\|H\|^2\|X\|^2\|(\mathbb{1}-M)\|^2+\\
&\quad 2\|X(\mathbb{1}-M)\|^2\|H\|^2+2\|X\|^2\|H(\mathbb{1}-M\|^2\\
&\geq 16\langle X,H\rangle^2+8\|X\|^2\|H\|^2-2\|X\|^2\|H\|^2\|(\mathbb{1}-M)\|^2-2\|H\|^2\|X\|^2\|(\mathbb{1}-M)\|^2\\
&\geq 16\langle X,H\rangle^2+8\|X\|^2\|H\|^2-8\|X\|^2\|H\|^2\\
&=16\langle X,H\rangle^2\geq 0.
\end{aligned}$$

6. $\frac{1}{2}[\mathrm{tr}(X^TX)]^2+\mathrm{tr}(X^TXX^TXM)=\frac{1}{2}\|X\|^4+\|X^TX.\eta\|^2$. Compute the differentials

$$\begin{aligned}
&D_X\left(\frac{1}{2}\|X\|^4+\|X^TX.\eta\|^2\right).H=2\|X\|^2\|H\|^2+\langle X^TX.\eta,(X^TH+H^TX).\eta\rangle\\
&D_X^2\left(\frac{1}{2}\|X\|^4+\|X^TX.\eta\|^2\right).(H,H)=2\|X\|^2\|H\|^2+4\langle X,H\rangle^2+\\
&\quad 2\langle X^TX.\eta,H^TH.\eta\rangle+\|(X^TH+H^TX).\eta\|^2\\
&\geq 2\|X\|^2\|H\|^2+4\langle X,H\rangle^2-2\|X\|^2\|H\|^2\|\eta\|^2+\|(X^TH+H^TX).\eta\|^2\\
&=4\langle X,H\rangle^2+\|(X^TH+H^TX).\eta\|^2\geq 0.
\end{aligned}$$

■

Lemma 10.3 (Generic anisotropic polyconvex terms)
Let $F \in \mathbb{M}^{3\times 3}$ and $M = \eta \otimes \eta$. Then the following terms are each polyconvex for $k \geq 1$:

$$1.\ \frac{[\mathrm{tr}(F^TF)]^k}{(\det[[F^TF]])^{\frac{1}{3}}},\quad 2.\ \frac{[\mathrm{tr}(F^TFM)]^k}{(\det[[F^TF]])^{\frac{1}{3}}},\quad 3.\ \frac{[\mathrm{tr}(F^TF(1\!\!1-M))]^k}{(\det[[F^TF]])^{\frac{1}{3}}}$$
$$4.\ \frac{[\mathrm{tr}(\mathrm{Adj}(F^TF))]^k}{(\det[[F^TF]])^{\frac{1}{3}}},\quad 5.\ \frac{[\mathrm{tr}(\mathrm{Adj}(F^TF)M)]^k}{(\det[[F^TF]])^{\frac{1}{3}}},\quad 6.\ \frac{[\mathrm{tr}(\mathrm{Adj}(F^TF)(1\!\!1-M))]^k}{(\det[[F^TF]])^{\frac{1}{3}}}$$

Proof.

1.

$$\frac{[\mathrm{tr}(F^TF)]^k}{(\det[[F^TF]])^{\frac{1}{3}}} = \frac{\|F\|^{2k}}{(\det[F])^{\frac{2}{3}}}$$

and we may use the same ideas as in the proof to Lemma **??** to conclude that the term is polyconvex.

2.

$$\frac{[\mathrm{tr}(F^TFM)]^k}{(\det[[F^TF]])^{\frac{1}{3}}} = \frac{\langle F, FM\rangle^k}{(\det[F])^{\frac{2}{3}}} = \frac{\langle F, F(\eta\otimes\eta)\rangle^k}{(\det[F])^{\frac{2}{3}}} = \frac{\|F.\eta\|^{2k}}{(\det[F])^{\frac{2}{3}}}.$$

We have already shown (see (**??**)) that the function $P(x, y) = \frac{1}{x^\alpha} \cdot y^p$ is convex provided that $\alpha = \frac{2}{3}$ and $p = 2k \geq 2$. Now define a new function

$$\hat{W}(F, \zeta) := P(\zeta, \|F.\eta\|) = \frac{\|F.\eta\|^{2k}}{\zeta^{\frac{2}{3}}}.$$

Observe that by the monotonicity of the square for positive arguments we have the inequality

$$\|\lambda F_1.\eta + (1-\lambda)F_2.\eta\|^{2k} \leq (\lambda\|F_1.\eta\| + (1-\lambda)\|F_2.\eta\|)^{2k}\ . \tag{10.5}$$

It remains to check the convexity of $\hat{W}(F, \zeta)$. To this end

$$\begin{aligned}\hat{W}(\lambda F_1 + (1-\lambda)F_2, \lambda\zeta_1 + (1-\lambda)\zeta_2) &= P(\lambda\zeta_1 + (1-\lambda)\zeta_2, \|\lambda F_1.\eta + (1-\lambda)F_2.\eta\|)\\ &= \frac{\|\lambda F_1.\eta + (1-\lambda)F_2.\eta\|^{2k}}{(\lambda\zeta_1 + (1-\lambda)\zeta_2)^{\frac{2}{3}}}.\end{aligned}$$

With (**??**) we have

$$\begin{aligned}\hat{W}(\lambda F_1 + (1-\lambda)F_2, \lambda\zeta_1 + (1-\lambda)\zeta_2) &\leq \frac{(\lambda\|F_1.\eta\| + (1-\lambda)\|F_2.\eta\|)^{2k}}{(\lambda\zeta_1 + (1-\lambda)\zeta_2)^{\frac{2}{3}}}\\ &= P(\lambda\zeta_1 + (1-\lambda)\zeta_2, \lambda\|F_1.\eta\| + (1-\lambda)\|F_2.\eta\|).\end{aligned}$$

The convexity of P yields

$$\begin{aligned}\hat{W}(\lambda F_1 + (1-\lambda)F_2, \lambda\zeta_1 + (1-\lambda)\zeta_2) &\leq \lambda P(\zeta, \|F_1.\eta\|) + (1-\lambda)P(\zeta_2, \|F_2.\eta\|)\\ &= \lambda\hat{W}(F_1,\zeta_1) + (1-\lambda)\hat{W}(F_2,\zeta_2).\end{aligned}$$

The proof is finished bearing the correct extension (**??**) in mind.

3. $\dfrac{[\operatorname{tr}(F^TF(1\!\!1 - M))]^k}{(\det[[F^TF]])^{\frac{1}{3}}} = \dfrac{\|F(1\!\!1 - M)\|^{2k}}{(\det[F])^{\frac{2}{3}}}$ and we proceed as in the second case.

4. $\dfrac{[\operatorname{tr}(\operatorname{Adj}(F^TF))]^k}{(\det[[F^TF]])^{\frac{1}{3}}} = \dfrac{\|\operatorname{Adj} F\|^{2k}}{(\det[F])^{\frac{2}{3}}}$ and we proceed as in the first case.

5. $\dfrac{[\operatorname{tr}(\operatorname{Adj}(F^TF)M)]^k}{(\det[[F^TF]])^{\frac{1}{3}}} = \dfrac{\|\operatorname{Adj} F^T.\eta\|^{2k}}{(\det[F])^{\frac{2}{3}}}$ and we proceed as in the second case.

6. $\dfrac{[\operatorname{tr}(\operatorname{Adj}(F^TF)(1\!\!1 - M))]^k}{(\det[[F^TF]])^{\frac{1}{3}}} = \dfrac{\|\operatorname{Adj} F^T(1\!\!1 - M)\|^{2k}}{(\det[F])^{\frac{2}{3}}}$ and we proceed as in the second case.

■

Corollary 10.4 (Generic anisotropic exponential polyconvex terms)
Let $F \in \mathbb{M}^{3\times 3}$ and $M = \eta \otimes \eta$. Then the following terms are each polyconvex for $k \geq 1$:

$$\begin{array}{llll} 1. & \exp\left[\dfrac{[\operatorname{tr}(F^TF)]^k}{(\det[[F^TF]])^{\frac{1}{3}}}\right], & 2. & \exp\left[\dfrac{[\operatorname{tr}(F^TFM)]^k}{(\det[[F^TF]])^{\frac{1}{3}}}\right], \\ 3. & \exp\left[\dfrac{[\operatorname{tr}(F^TF(1\!\!1 - M))]^k}{(\det[[F^TF]])^{\frac{1}{3}}}\right] & 4. & \exp\left[\dfrac{[\operatorname{tr}(\operatorname{Adj}(F^TF))]^k}{(\det[[F^TF]])^{\frac{1}{3}}}\right], \\ 5. & \exp\left[\dfrac{[\operatorname{tr}(\operatorname{Adj}(F^TF)M)]^k}{(\det[[F^TF]])^{\frac{1}{3}}}\right], & 6. & \exp\left[\dfrac{[\operatorname{tr}(\operatorname{Adj}(F^TF)(1\!\!1 - M))]^k}{(\det[[F^TF]])^{\frac{1}{3}}}\right] \end{array}$$

$$7. \quad \exp[W(F)] \quad \textit{if} \quad W(F) \quad \textit{is polyconvex.}$$

Proof. By the foregoing lemma each argument of the exponential is polyconvex. Since exp is convex and monotone increasing it preserves the underlying convexity. Hence the composition is polyconvex. Observe, however, that these functions alone are not stress-free in the reference configuration. ■

Let us show that for any $p > 2$

$$W_{\text{aniso}}(F) = \begin{cases} \exp\left(\|F.a\|^2 - 1\right)^p & \|F.a\|^2 \geq 1\,, \\ 0 & \text{else}\,, \end{cases} \tag{10.6}$$

is Legendre-Hadamard elliptic. This is enough to see that any additive composition of W_{aniso} with an isotropic elliptic energy will also remain Legendre-Hadamard elliptic. To see this, we compute the piecewise second differential. Since

$$D_F[\exp\left(\|F.a\|^2-1\right)^p].H = \exp\left(\|F.a\|^2-1\right)^p \left[p\left(\|F.a\|^2-1\right)^{p-1} 2\langle F.a, H.a\rangle\right] , \tag{10.7}$$

we obtain for the non-zero branch of W_{aniso}

$$\begin{aligned} D_F^2 W_{\text{aniso}}(F).(H,H) &= \exp\left(\|F.a\|^2-1\right)^p [\ldots]^2 \qquad\qquad (10.8)\\ &+ \exp\left(\|F.a\|^2-1\right)^p\, 2\,p\left[(p-1),\left(\|F.a\|^2-1\right)^{p-2} 2\langle F.a, H.a\rangle^2 + \left(\|F.a\|^2-1\right)^{p-1}\langle H.a, H.a\rangle\right] . \end{aligned}$$

This formula tends continuously to zero for $\|F.a\|^2 \to 1$ and is positive for $\|F.a\|^2 \geq 1$. Hence, the complete second differential is always positive and continuous. Thus, convexity of W_{aniso} implies Legendre-Hadamard ellipticity. By continuity, we obtain that W_{aniso} is convex also for $p = 2$.

One might be tempted to use some other ansatz terms in order to construct polyconvex strain energies. However, we have e.g.

Lemma 10.5 (Non-elliptic terms I)
Let $F \in \mathbb{M}^{3\times 3}$ and $M = a \otimes a$. Then the following terms are each non-elliptic, hence non-quasiconvex:

1. $F \mapsto \operatorname{tr}(F^TF\cdot M)\cdot\operatorname{tr}(F^TF) = \operatorname{tr}(CM)\cdot\operatorname{tr}(C)$.
2. $F \mapsto \operatorname{tr}(F^TFF^TF\cdot M) = \operatorname{tr}(C^2M)$.
3. $$F \mapsto \left(\frac{\|\operatorname{Adj}F\|^2}{(\det[F])^{\frac{4}{3}}} - 3\right)^i = \left(\operatorname{tr}(\operatorname{Adj}(\frac{C}{(\det[C])^{\frac{1}{3}}})) - 3\right)^i = \left(\operatorname{tr}(\operatorname{Adj}(\frac{C}{(\det[C])^{\frac{1}{3}}}) - 1\!\!1)\right)^i \quad i \geq 1.$$

Proof.

1. The last equation can be expressed in the form

$$\operatorname{tr}[F^TFM]\operatorname{tr}[F^TF] = \|F\|^2\|F.a\|^2 .$$

Calculating the second differential with respect to the deformation gradient yields

$$\begin{aligned} D_F^2\left(\|F\|^2\cdot\|F.a\|^2\right).(H,H) = 8\langle F, H\rangle\langle F.a, H.a\rangle +\\ 2\|F.a\|^2\|H\|^2 + 2\|F\|^2\|H.a\|^2. \end{aligned}$$

We see that this expression is in general non-positive (take F, H in diagonal form), which excludes convexity. However, it is possible to show the non-ellipticity as well. Take

$$F_n := \begin{pmatrix} \frac{1}{n} & -1 & 0\\ 0 & \frac{1}{n} & 0\\ 0 & 0 & \frac{1}{n}\end{pmatrix}, \quad \xi = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \eta = \begin{pmatrix}1\\n\\0\end{pmatrix}, \quad a = \begin{pmatrix}1\\0\\0\end{pmatrix}$$

and $H = \xi \otimes \eta$. This yields

$$D_F^2\left(\|F\|^2\|F.a\|^2\right).(\xi\otimes\eta,\xi\otimes\eta) = \\ 8\langle F,\xi\otimes\eta\rangle\langle F.a,\xi\otimes\eta.a\rangle + 2\|F.a\|^2\|\xi\otimes\eta\|^2 + 2\|F\|^2\|\xi\otimes\eta.a\|^2 \\ = 8(\frac{1}{n}-n)\frac{1}{n} + 2\frac{1}{n^2}\,(1+n^2) + 2(3\frac{1}{n^2}+1) = \frac{16}{n^2} - 4\,.$$

If we choose $n > 2$, then we get

$$D_F^2\left(\|F\|^2\cdot\|F.a\|^2\right).(\xi\otimes\eta,\xi\otimes\eta) < 0.$$

Thus, the non-ellipticity of this function is shown.

2. The forms of the individual expressions are

$$\operatorname{tr}[F^TFF^TFM] = \|F^TF.a\|^2\,.$$

First we compute the second derivative of the function with respect to F

$$D_F^2\left(\|F^TF.a\|^2\right).(H,H) = 2\langle F^TF.a, H^TH.a\rangle + \|(F^TH+H^TF).a\|^2.$$

Set $H = \xi\otimes\eta$ with $\|\xi\| = \|\eta\| = 1$. This yields after some manipulation

$$D_F^2\left(\|F^TF.a\|^2\right).(\xi\otimes\eta,\xi\otimes\eta) = 2\langle F.a, F\eta\rangle\langle\eta,a\rangle + \langle\eta,a\rangle^2\|F^T\xi\|^2 + \\ \langle F^T\xi,a\rangle^2 + 2\langle F^T\xi,\eta\rangle\langle\eta,a\rangle\langle F^T\xi,a\rangle.$$

Take the explicit expressions

$$F_n := \begin{pmatrix}1&0&0\\0&2&0\\0&0&\frac{1}{n}\end{pmatrix},\quad \xi = \begin{pmatrix}0\\0\\1\end{pmatrix}\quad \|F_n^T\xi\|^2 = \frac{1}{n^2},\quad a = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix},\quad \eta = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\-1\\1\end{pmatrix}\,.$$

This leads to the expressions

$$\langle F_n.a, F_n\eta\rangle = -1 + \frac{1}{3n^2},\qquad \langle a,\eta\rangle = \frac{1}{3}$$

and altogether we have for some reasonable n

$$D_F^2\left(\|F_n^TF_n.a\|^2\right).(\xi\otimes\eta,\xi\otimes\eta) \le (-2+\frac{2}{3n^2})\frac{1}{3} + \frac{4}{n} < 0.$$

Observe, that the isotropic counterpart $\operatorname{tr}[[]\mathbb{C}^2] = \|F^TF\|^2$ is a convex function of F.

3. Even though $\|\operatorname{Adj}F\|^2/(\det[F])^{\frac{4}{3}} - 3 \ge 0$ in light of Lemma **??**, the term $\|\operatorname{Adj}F\|^2/(\det[F])^{\frac{4}{3}}$ alone does not have the right exponents to be polyconvex. Moreover it can be shown that the term is non-elliptic (Dacorogna, 1989). . ■

Lemma 10.6 (Non-elliptic terms II)
Let $F \in \mathbb{M}^{3\times 3}$ and $M = a\otimes a$ with $\|a\| = 1$. Then the following terms are each non-elliptic, hence non-quasiconvex:

1. $F \mapsto \exp\left(\langle \frac{C}{(\det[C])^{\frac{1}{3}}}, a\otimes a\rangle - 1\right) - \langle \frac{C}{(\det[C])^{\frac{1}{3}}}, a\otimes a\rangle.$

2. $F \mapsto \left(\langle \frac{C}{(\det[C])^{\frac{1}{3}}}, a\otimes a\rangle - 1\right)^q, \; q \geq 2.$

Observe that both terms have stress-free reference configuration.

Proof. We show the non-ellipticity of the first expression. The non-ellipticity of the second one follows along the same lines. We calculate

$$\langle \frac{C}{(\det[C])^{\frac{1}{3}}}, a\otimes a\rangle = \langle \frac{F^TF}{(\det[F])^{\frac{2}{3}}}, a\otimes a\rangle = \frac{1}{(\det[F])^{\frac{2}{3}}}\|F.a\|^2.$$

Set $F = F_0 + t\xi\otimes\eta$. This yields

$$\frac{1}{(\det[F])^{\frac{2}{3}}}\|F.a\|^2 = \frac{\|F_0 + t\xi\otimes\eta\|^2}{(\det[[]F_0 + t\xi\otimes\eta])^{\frac{2}{3}}} = \frac{\|F_0.a + t\xi\langle\eta, a\rangle\|^2}{(\det[F]_0 + \langle \operatorname{Adj} F_0^T, t\xi\otimes\eta\rangle + 0 + 0)^{\frac{2}{3}}} =$$
$$\frac{\|F_0.a\|^2 + 2t\langle F_0.a,\xi\rangle\langle\eta,a\rangle + t^2\langle\xi,\xi\rangle\langle\eta,a\rangle}{(\det[F]_0 + t\langle 1\!\!1, \operatorname{Adj} F_0.\xi\otimes\eta\rangle)^{\frac{2}{3}}} = \frac{\|F_0.a\|^2 + 2t\langle F_0.a,\xi\rangle\langle\eta,a\rangle + t^2\langle\xi,\xi\rangle\langle\eta,a\rangle}{(\det[F]_0 + t\langle \operatorname{Adj} F_0.\xi, \eta\rangle)^{\frac{2}{3}}}.$$

Now we choose

$$a = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}, \xi = \begin{pmatrix}0\\1\\0\end{pmatrix}, \eta = \begin{pmatrix}1\\0\\0\end{pmatrix}, F_0^{-1} = \begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & \frac{3\sqrt{2}}{2} & 0\\ 0 & 0 & d\end{pmatrix}.$$

This yields

$$\begin{aligned}
&\|a\| = \|\xi\| = \|\eta\| = 1,\\
&\langle a,\eta\rangle = \langle a,\xi\rangle = \frac{1}{\sqrt{2}},\\
&\langle\eta,\xi\rangle = 0,\\
&F_0^{-1}.\eta = a,\; \det[F]_0 = \frac{1}{d},\; F_0.a = \eta\\
&\det[F]_0\cdot a = \det[F]_0\cdot F_0^{-1}.\eta = \operatorname{Adj} F_0.\eta.
\end{aligned}$$

As a consequence, we get

$$\frac{\|F_0.a\|^2 + 2t\langle F_0.a,\xi\rangle\langle\eta,a\rangle + t^2\langle\xi,\xi\rangle\langle\eta,a\rangle}{(\det[F]_0 + t\langle \operatorname{Adj} F_0.\xi,\eta\rangle)^{\frac{2}{3}}} = \frac{\|\eta\|^2 + t\langle\eta,\xi\rangle\langle\eta,a\rangle + t^2\langle\xi,\xi\rangle\langle\eta,a\rangle}{(\frac{1}{d} + t\langle\xi,\operatorname{Adj} F_0^T.\eta\rangle)^{\frac{2}{3}}}$$

$$= \frac{1+0+t^2\frac{1}{\sqrt{2}}}{(\frac{1}{d} + t\langle\xi,\operatorname{Adj} F_0.\eta\rangle)^{\frac{2}{3}}} = \frac{1+t^2\frac{1}{\sqrt{2}}}{(\frac{1}{d} + t\det[F]_0\langle\xi,a\rangle)^{\frac{2}{3}}} = \frac{1+t^2\frac{1}{\sqrt{2}}}{\frac{1}{d^{\frac{2}{3}}}(1+t\frac{1}{\sqrt{2}})^{\frac{2}{3}}}.$$

Thus

$$h(t) = W(F_0 + t\xi\otimes\eta) = \exp\left(\frac{1+t^2\frac{1}{\sqrt{2}}}{\frac{1}{d^{\frac{2}{3}}}(1+t\frac{1}{\sqrt{2}})^{\frac{2}{3}}} - 1\right) - \frac{1+t^2\frac{1}{\sqrt{2}}}{\frac{1}{d^{\frac{2}{3}}}(1+t\frac{1}{\sqrt{2}})^{\frac{2}{3}}}\,.$$

If we choose $\frac{1}{d^{\frac{2}{3}}} = 3$ it turns out that h is not convex in t, hence (Theorem **??**) W is not elliptic. We remark that the non-ellipticity is mainly due to the fact that

$$\langle\frac{C}{(\det[C])^{\frac{1}{3}}}, a\otimes a\rangle \geq 1$$

is in general not true (consider $F_n = \operatorname{diag}(n, 1, \frac{1}{n})$), whereas

$$\langle\frac{C}{(\det[C])^{\frac{1}{3}}}, 1\!\!1\rangle \geq 3$$

holds by virtue of Lemma **??**. ■

Lemma 10.7 (Non-elliptic terms III)
Let $F\in\mathbb{M}^{3\times3}$ and $M = a\otimes a$ with $\|a\| = 1$. Then the following terms are non-elliptic, hence non-quasiconvex:

1. $F\mapsto W(F) = c_1\ \operatorname{tr}(CM) - c_2\ln\sqrt{\operatorname{tr}(CM)}$.
2. $F\mapsto W(F) = c_1\ \operatorname{tr}(\operatorname{Adj} CM) - c_2\ln\sqrt{\operatorname{tr}(\operatorname{Adj} CM)}$.

with $c_1, c_2 > 0$. Observe that these terms have a physically desired singularity in fiber direction, i.e

$$\begin{aligned} W_S(F) &\to \infty \quad \text{as} \quad F.a \to 0 \\ W_S(F) &\to \infty \quad \text{as} \quad \operatorname{Adj} F.a \to 0. \end{aligned}$$

Proof. We show that the ellipticity condition is in general violated for the first term. We calculate

$$\begin{aligned} W_S(F) &= c_1\ \operatorname{tr}(CM) - c_2\ln\sqrt{\operatorname{tr}(CM)} = c_1\ \|F.a\|^2 - c_2\ln\|F.a\| \\ &= c_1\ \|F.a\|^2 - \frac{c_2}{2}\ln\|F.a\|^2. \end{aligned}$$

Then calculating the first and second differential yields

$$DW_S(F).H = 2c_1\langle F.a, H.a\rangle - \frac{c_2}{\|F.a\|^2} \cdot \langle F.a, H.a\rangle$$
$$D^2W_S(F).(H,H) = 2c_1\|H.a\|^2 - c_2\left(\frac{1}{\|F.a\|^2} \cdot \|H.a\|^2 - 2 \cdot \frac{\langle F.a, H.a\rangle^2}{\|F.a\|^4}\right).$$

Take $H = \xi \otimes \eta$ with $\|\xi\| = \|\eta\| = 1$. This gives

$$\begin{aligned} D^2W_S(F).(\xi \otimes \eta, \xi \otimes \eta) &= 2c_1\|\xi\langle \eta, a\rangle\|^2 - c_2\left(\frac{1}{\|F.a\|^2} \cdot \|\xi\langle\eta, a\rangle\|^2 - 2\frac{\langle F.a, \xi\langle\eta, a\rangle^2\rangle}{\|F.a\|^4}\right) \\ &= 2c_1\langle\eta, a\rangle^2 - c_2\left(\frac{\langle\eta, a\rangle^2}{\|F.a\|^2} - 2\frac{\langle F.a, \xi\rangle\langle\eta, a\rangle^2}{\|F.a\|^4}\right). \end{aligned}$$

Without loss of generality assume that ξ is chosen such that $\langle F.a, \xi\rangle = 0$. It follows that

$$D^2W_S(F).(\xi \otimes \eta, \xi \otimes \eta) = \langle\eta, a\rangle^2\left(2c_1 - \frac{c_2}{\|F.a\|^2}\right).$$

If the deformation F in fiber direction a is such that $\|F.a\|^2 < \frac{c_2}{2c_1}$ then

$$D^2W_S(F).(\xi \otimes \eta, \xi \otimes \eta) = \langle\eta, a\rangle^2\left(2c_1 - \frac{c_2}{\|F.a\|^2}\right) < 0.$$

Observe that the more severe the deformation in fiber direction is, the more the ellipticity condition is violated. It is thus just the physically interesting region $\|F.a\|$ small which fails to be elliptic. Now we consider the second term. We calculate

$$W_S(F) = c_1\ \operatorname{tr}(\operatorname{Adj} CM) - c_2\ln\sqrt{\operatorname{tr}(\operatorname{Adj} CM)} = c_1\|\operatorname{Adj} F^T.a\|^2 - \frac{c_2}{2}\ln\|\operatorname{Adj} F^T.a\|^2$$
$$DW_S(F).H = 2c_1\langle\operatorname{Adj} F^T.a, D\operatorname{Adj} F^T.H^T.a\rangle - \frac{c_2}{\|\operatorname{Adj} F^T.a\|^2}\langle\operatorname{Adj} F^T.a, D\operatorname{Adj} F^T.H^T.a\rangle$$
$$\begin{aligned} &D^2W_S(F).(H,H) = 2c_1[\langle D\operatorname{Adj} F^T.H^T.a, D\operatorname{Adj} F^T.H^T.a\rangle + \\ &\quad \langle\operatorname{Adj} F^T.a, D^2\operatorname{Adj} F^T.(H^T, H^T).a\rangle] - \frac{c_2}{\|\operatorname{Adj} F^T.a\|^2}[\langle D\operatorname{Adj} F^T.H^T.a, D\operatorname{Adj} F^T.H^T.a\rangle + \\ &\quad \langle\operatorname{Adj} F^T.a, D^2\operatorname{Adj} F^T.(H^T, H^T).a\rangle] + \frac{2c_2}{\|\operatorname{Adj} F^T.a\|^4}\langle\operatorname{Adj} F^T.a, D\operatorname{Adj} F^T.H^T.a\rangle^2 \end{aligned}$$

Since $D^2\operatorname{Adj} F.(H,H) = 2\operatorname{Adj} H$ and for $H = \xi \otimes \eta$ we have $\operatorname{Adj} \xi \otimes \eta = 0$, it follows that

$$\begin{aligned} D^2W_S(F).(\xi \otimes \eta, \xi \otimes \eta) &= 2c_1\|D\operatorname{Adj} F^T.H^T.a\|^2 - \frac{c_2}{\|\operatorname{Adj} F^T.a\|^2}\|D\operatorname{Adj} F^T.H^T.a\|^2 \\ &\quad + \frac{2c_2}{\|\operatorname{Adj} F^T.a\|^4}\langle\operatorname{Adj} F^T.a, D\operatorname{Adj} F^T.H^T.a\rangle^2 \\ &= \|D\operatorname{Adj} F^T.H^T.a\|^2[2c_1 - \frac{c_2}{\|\operatorname{Adj} F^T.a\|^2}] + \\ &\quad \frac{2c_2}{\|\operatorname{Adj} F^T.a\|^4}\langle\operatorname{Adj} F^T.a, D\operatorname{Adj} F^T.H^T.a\rangle^2. \end{aligned}$$

Consider $\langle \operatorname{Adj} F^T.a, D \operatorname{Adj} F^T.H^T.a\rangle$. If we choose $F^{-T}.a = s\cdot\xi$ with $s \in \mathbb{R}^+$, then

$$\begin{aligned}
\langle \operatorname{Adj} F^T.a, D \operatorname{Adj} F^T.H^T.a\rangle &= \langle \det[F]F^{-T}.a, \operatorname{Adj} F^T[\langle F^{-1}, H^T\rangle 1\!\!1 - H^T F^{-T}].a\rangle \\
&= \det[F]^2 \langle F^{-T}.a, F^{-T}[\langle F^{-1}, \eta\otimes\xi\rangle 1\!\!1 - (\eta\otimes\xi)F^{-T}].a\rangle \\
&= \det[F]^2 \langle F^{-T}.a, \langle F^{-T}.\eta, \xi\rangle F^{-T}.a - F^{-T}.(\eta\otimes\xi)F^{-T}.a\rangle \\
&= \det[F]^2 \left[\|F^{-T}.a\|^2 \langle F^{-T}.\eta,\xi\rangle - \langle F^{-T}.a, F^{-T}.(\eta\otimes\xi)F^{-T}.a\rangle\right] \\
&= \det[F]^2 s^2 \left[\|\xi\|^2\langle F^{-T}.\eta,\xi\rangle - \langle \xi, (F^{-T}.\eta\otimes\xi).\xi\rangle\right] \\
&= \det[F]^2 s^2 \left[1\cdot\langle F^{-1}.\xi,\eta\rangle - \langle \eta, F^{-1}.\xi\cdot 1\rangle\right] = 0.
\end{aligned}$$

With this choice we get

$$\begin{aligned}
D^2 W_S(F).(\xi\otimes\eta,\xi\otimes\eta) &= \|D \operatorname{Adj} F^T.H^T.a\|^2 [2c_1 - \frac{c_2}{\|\operatorname{Adj} F^T.a\|^2}] \\
&= \|D \operatorname{Adj} F^T.H^T.a\|^2 [2c_1 - \frac{c_2}{\det[F]^2\|F^{-T}.a\|^2}] \\
&= \|D \operatorname{Adj} F^T.H^T.a\|^2 [2c_1 - \frac{c_2}{\det[F]^2\ s^2}]\,.
\end{aligned}$$

Since F can still be chosen with $\det[F] = 1$ taking $s > 0$ sufficiently small finishes the argument. ■

11 Coercivity for metric based anisotropic elasticity

Let us recall

$$\begin{aligned}
g_j &= \operatorname{tr}[G_j] = \operatorname{tr}\left[H_j^T H_j\right] = \|H_j\|^2 > 0 \\
J_{4j} &= \operatorname{tr}[C\,G_j] = \langle F^T F, H_j^T H_j\rangle = \langle F H_j^T, F H_j^T\rangle \\
&= \|F\,H_j^T\|^2 = \|H_j\,F^T\|^2 \geq \lambda_{\min}(H_j^T H_j)\|F^T\|^2 = \lambda_{\min}(G_j)\|F\|^2\,, \\
J_{5j} &= \operatorname{tr}[\operatorname{Cof} C\,G_j] = \langle \operatorname{Cof} F^T \operatorname{Cof} F, H_j^T H_j\rangle = \langle \operatorname{Cof} F\,H_j^T, \operatorname{Cof} F\,H_j^T\rangle = \|\operatorname{Cof} F\,H_j^T\|^2 \\
&= \|H_j\ \operatorname{Cof} F^T\|^2 \geq \lambda_{\min}(H_j^T H_j)\|\operatorname{Cof} F^T\|^2 = \lambda_{\min}(G_j)\|\operatorname{Cof} F\|^2\,, \\
I_3 &= \det[C] = \det[F]^2 \leq \frac{1}{3\sqrt{3}}\|\operatorname{Cof} F\|^3 \quad \text{see (??)}\,. \qquad (11.9)
\end{aligned}$$

Since G_j is always strictly positive definite we know that the smallest eigenvalue $\lambda_{\min}(G_j) > 0$ is strictly positive.

With these preliminaries let us proceed to show that the anisotropic energy Ψ_2^{aniso} satisfies a local coercivity condition, which is needed, together with polyconvexity of Ψ_2^{aniso} to ensure the existence of global energy minimizers. Coercivity is a condition that ensures that the energy growth enough for large deformation gradients F. More precisiely by local coercivity we mean here an estimate of the type, see [**?**, Th.2.2]

$$\forall\, F \in \mathbb{M}^{3\times 3}: \quad \Psi_2^{aniso}(F) \geq C_1\ (\|F\|^p + \|\operatorname{Cof} F\|^q) - C_2\,, \quad p\geq 2\,,\ q\geq \frac{3}{2}\,, \qquad (11.10)$$

with constants $C_1, C_2 \geq 0$ and $C_1 > 0$.

The function Ψ_2^{aniso} has the generic form (taking only the relevant structure into account, i.e. setting $\alpha_{rj} = \alpha$, $\beta_{rj} = \beta$, $\gamma_{rj} = \gamma$, $g_j = g$, $J_{4j} = J_4$, $j_{5j} = J_5$)

$$\Psi_2^{aniso}(F) = \frac{1}{1+\alpha}\,\frac{1}{g^\alpha} J_4^{1+\alpha} + \frac{1}{1+\beta}\,\frac{1}{g^\beta} J_5^{1+\beta} + \frac{g}{\gamma}\, I_3^{-\gamma}\,. \tag{11.11}$$

Thus it follows easily, taking the relations (**??**) into account, that for $\alpha, \beta \geq 0$

$$\begin{aligned}\Psi_2^{aniso}(F) &\geq \frac{1}{1+\alpha}\frac{1}{g^\alpha}\, J_4 + \frac{1}{1+\beta}\frac{1}{g^\beta} J_5 + \frac{g}{\gamma}\, I_3^{-\gamma} \\ &\geq \frac{1}{1+\alpha}\frac{1}{g^\alpha}\, \lambda_{\min}(G_j)\, \|F\|^2 + \frac{1}{1+\beta}\frac{1}{g^\beta}\, \lambda_{\min}(G_j)\, \|\operatorname{Cof} F\|^2 + \frac{g}{\gamma}\, I_3^{-\gamma} \\ &= c_1^+\, \|F\|^2 + c_2^+\, \|\operatorname{Cof} F\|^2 + \frac{g}{\gamma}\, \det[F]^{-2\gamma}\,, \end{aligned} \tag{11.12}$$

for some given constants c_1^+, $c_2^+ > 0$. In case that γ is positive we have shown (**??**) with $C_1 = \min(c_1^+, c_2^+)$, $C_2 = 0$ and $p = q = 2$.

In the case where γ is negative with $0 \geq \gamma \geq -\frac{1}{2}$ we may continue estimating

$$\begin{aligned}\Psi_2^{aniso}(F) &\geq c_1^+\, \|F\|^2 + c_2^+\, \|\operatorname{Cof} F\|^2 + \frac{g}{\gamma}\, \det[F]^{-2\gamma}\,, \\ &\geq c_1^+\, \|F\|^2 + c_2^+\, \|\operatorname{Cof} F\|^2 - c_3^+\, \det[F]^{-2\gamma}\,, \\ &= c_1^+\, \|F\|^2 + c_2^+\, \|\operatorname{Cof} F\|^2 - c_3^+\, (\det[F]^2)^{|\gamma|}\,, \\ &\geq c_1^+\, \|F\|^2 + c_2^+\, \|\operatorname{Cof} F\|^2 - c_3^+\, [\det[F] + 1]\,, \quad 0 \leq |\gamma| \leq \frac{1}{2} \\ &\geq c_1^+\, \|F\|^2 + c_2^+\, \|\operatorname{Cof} F\|^2 - c_3^+\, [\frac{1}{\sqrt{3\sqrt{3}}}\, \|\operatorname{Cof} F\|^{\frac{3}{2}} + 1]\,. \end{aligned} \tag{11.13}$$

It is obvious that for all $k_1, k_2 > 0$ there exist numbers $\tilde{k}_1, \tilde{k}_2 > 0$ such that

$$\forall\, x \in \mathbb{R}^+ : \quad k_1\, x^2 - k_2\, x^{3/2} \geq \tilde{k}_1\, x^{3/2} - \tilde{k}_2\,. \tag{11.14}$$

Applying this reasoning on $x = \|\operatorname{Cof} F\|$ yields the existence of numbers C_2^+, $C_3^+ > 0$ such that

$$\begin{aligned}\Psi_2^{aniso}(F) &\geq c_1^+\, \|F\|^2 + c_2^+\, \|\operatorname{Cof} F\|^2 - c_3^+\, [\frac{1}{\sqrt{3\sqrt{3}}}\, \|\operatorname{Cof} F\|^{\frac{3}{2}} + 1] \\ &\geq c_1^+\, \|F\|^2 + C_2^+\, \|\operatorname{Cof} F\|^{3/2} - C_3^+\,. \end{aligned} \tag{11.15}$$

This shows local coercivity with $C_1 = \min(c_1^+, C_2^+)$ and $p = 2$ and $q = \frac{3}{2}$ also for $0 \geq \gamma \geq -\frac{1}{2}$.

12 Local coercivity of a Cosserat model

Let us assume given a material with rotational degree of freedoms, called a Cosserat material. Here, the tensor $\overline{U} = \overline{R}^T F$ is not necessarily symmetric and $\overline{R} \in \mathrm{SO}(3)$. We assume the strain energy can be written as

$$\begin{aligned} W_{\mathrm{macro}}^{\mathrm{Coss}}(\overline{U}) = 2\alpha \, \| \operatorname{sym} \overline{U} - \mathbb{1} \|^2 + \alpha \operatorname{tr}\left[\overline{U} - \mathbb{1}\right]^2 \\ + 2\beta \, \| \operatorname{sym} \operatorname{Cof} \overline{U} - \mathbb{1} \|^2 + \beta \operatorname{tr}\left[\operatorname{Cof} \overline{U} - \mathbb{1}\right]^2 \\ + \gamma \left((\det[\overline{U}] - 1)^2 + (\frac{1}{\det[\overline{U}]} - 1)^2 \right) , \end{aligned} \tag{12.16}$$

with positive material parameters $\alpha, \beta, \gamma \geq 0$.

Definition 12.1 (Rigidity in Cosserat models)
*Following Truesdell [**?**, p.311] we call the pair* $(\varphi, \overline{R}) \in \mathbb{R}^3 \times \mathrm{SO}(3)$ **rigid**, *whenever* $\nabla\varphi \in \mathrm{SO}(3)$ *and* $\overline{R} = \mathrm{const.}$. *Then* φ *represents a rigid rotation.*

Our first observation is

Lemma 12.2 (Symmetrized control)
Let $X \in \mathbb{M}^{3\times 3}$. *Every two out of the three following conditions imply that* $X \equiv \mathbb{1}$.

$$\operatorname{sym} X = \mathbb{1} \,, \quad \operatorname{sym} \operatorname{Cof} X = \mathbb{1} \,, \quad \det[X] = 1 \,. \tag{12.17}$$

Proof. We start with $\operatorname{sym} X = \mathbb{1}$ and $\operatorname{sym} \operatorname{Cof} X = \mathbb{1}$. Thus, from the first condition $X = \mathbb{1} + A$ with $A \in \mathfrak{so}(3)$ and

$$\begin{aligned} \operatorname{Cof}(\mathbb{1} + A) &= \mathbb{1} + D \operatorname{Cof} \mathbb{1}.A + \operatorname{Cof} A = \mathbb{1} + \left(\langle \mathbb{1}, A\rangle \mathbb{1} - A^T\right) + \operatorname{Cof} A \\ &= \mathbb{1} + A + \operatorname{Cof} A \,. \end{aligned} \tag{12.18}$$

However,

$$\operatorname{sym} \operatorname{Cof}(\mathbb{1} + A) = \mathbb{1} \;\Rightarrow\; \operatorname{sym} \operatorname{Cof} A = 0 \;\Rightarrow\; \operatorname{diag} \operatorname{sym} \operatorname{Cof} A = 0 \,, \tag{12.19}$$

which implies $A = 0$.[5] The second case is $\operatorname{sym} X = \mathbb{1}$ and $\det[X] = 1$. We use $X = \mathbb{1} + A$ and $1 = \det[\mathbb{1} + A] = 1 + \operatorname{tr}[A] + \langle \mathbb{1}, \operatorname{Cof} A\rangle + \det[A] = 1 + \langle \mathbb{1}, \operatorname{Cof} A\rangle = 1 + \|A\|^2$, which

---

[5] Here are some remarkable relations for skew-symmetric matrices:

$$\begin{aligned} A \in \mathfrak{so}(3) : \; \operatorname{Cof} A &= \operatorname{Cof} \begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{pmatrix} = \begin{pmatrix} \gamma^2 & -\beta\gamma & \alpha\gamma \\ -\beta\gamma & \beta^2 & -\beta\alpha \\ \alpha\gamma & -\beta\alpha & \alpha^2 \end{pmatrix} \in \mathrm{Sym} \,, \\ \|A\|^2 &= 2\,(\alpha^2 + \beta^2 + \gamma^2) = 2 \operatorname{tr}[\operatorname{Cof} A] \,, \\ \| \operatorname{Cof} A\|^2 &= \gamma^4 + \beta^4 + \alpha^4 + 2\,(\beta^2\gamma^2 + \alpha^2\gamma^2 + \beta^2\alpha^2) = (\alpha^2 + \beta^2 + \gamma^2)^2 = \operatorname{tr}[\operatorname{Cof} A]^2 \,, \\ A^T A &= \begin{pmatrix} 0 & -\alpha & -\beta \\ \alpha & 0 & -\gamma \\ \beta & \gamma & 0 \end{pmatrix} \begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{pmatrix} = \begin{pmatrix} \alpha^2 + \beta^2 & \beta\gamma & -\alpha\gamma \\ \beta\gamma & \alpha^2 + \gamma^2 & \alpha\beta \\ -\alpha\gamma & \beta\alpha & \beta^2 + \gamma^2 \end{pmatrix} \,, \\ \|A^T A\|^2 &= (\alpha^2 + \beta^2)^2 + (\alpha^2 + \gamma^2)^2 + (\beta^2 + \gamma^2)^2 + 2\,(\beta^2\gamma^2 + \alpha^2\gamma^2 + \alpha^2\beta^2) \\ &= 2\,(\alpha^4 + \beta^4 + \gamma^4) + 4\,(\beta^2\gamma^2 + \alpha^2\gamma^2 + \beta^2\alpha^2) = 2\,(\alpha^2 + \beta^2 + \gamma^2)^2 \\ &= 2 \operatorname{tr}[\operatorname{Cof} A]^2 = \frac{1}{2} \|A\|^4 \,. \end{aligned} \tag{12.20}$$

implies at once $A = 0$. Thence, $X = 1\!\!1$.

In the third case, $\operatorname{sym}\operatorname{Cof} X = 1\!\!1$ and $\det[X] = 1$. Since then $\operatorname{Cof} X = 1\!\!1 + A$ we obtain $\operatorname{Cof} X = \det[X]\, X^{-T} = 1\!\!1 + A$, or, using the determinant constraint, $X^{-T} = 1\!\!1 + A$. Taking the determinant again, implies

$$\frac{1}{\det[X]} = \det[1\!\!1 + A] = 1 + \operatorname{tr}[A] + \operatorname{tr}[\operatorname{Cof} A] + \det[A] = 1 + \operatorname{tr}[\operatorname{Cof} A] \;\Rightarrow$$
$$1 = 1 + \operatorname{tr}[\operatorname{Cof} A] = 1 + \frac{1}{2}\|A\|^2 \,. \tag{12.21}$$

Thus $A = 0$ as in the second case. ■

Corollary 12.3 (Rigidity control)
Zero local energy-level $W^{\mathrm{Coss}}_{\mathrm{macro}}(\overline{U}) \equiv 0$ holds if and only if $\overline{U} \equiv 1\!\!1$.

Proof. The reverse direction is obvious. In the other direction we observe

$$W^{\mathrm{Coss}}_{\mathrm{macro}}(\overline{U}) = 0 \;\Rightarrow\quad \operatorname{sym}\overline{U} = 1\!\!1\,, \quad \operatorname{sym}\operatorname{Cof}\overline{U} = 1\!\!1\,, \quad \det[\overline{U}] = 1\,, \tag{12.22}$$

and use the previous Lemma. ■

Corollary 12.4
Zero local energy-level $W^{\mathrm{Coss}}_{\mathrm{macro}}(\overline{U}) = 0$ implies that $\varphi(x) = \overline{Q}.x + b$ for some constant rotation $\overline{Q} \in \mathrm{SO}(3)$ and some constant translation $b \in \mathbb{R}^3$.

Proof. $W^{\mathrm{Coss}}_{\mathrm{macro}}(\overline{U}) = 0$ implies that $\overline{U} = 1\!\!1$, as seen above. However, $\overline{U} = \overline{R}^T \nabla\varphi = 1\!\!1$ shows that $\nabla\varphi(x) = \overline{R}(x)$. Now we apply the Curl-operator on both sides to see that $0 = \operatorname{Curl}\overline{R}(x)$. This shows that $\overline{R} = \text{const.}$ by using the estimate $\|\operatorname{curl}\overline{R}(x)\|^2 \geq C^+\,\|D_x\overline{R}\|^2$, shown in [?]. Thus $\nabla\varphi = \overline{Q}$ for a constant $\overline{Q}$. Hence φ is a rigid rotation. ■
Let us now show that even the following statement of coercivity is true:

Theorem 12.5 (Local coercivity)
There exist constants $C_1, C_2 > 0$ only depending on the material parameters $\alpha, \beta, \gamma > 0$ such that

$$W^{\mathrm{Coss}}_{\mathrm{macro}}(\overline{U}) \geq C_1^+\left(\|\overline{U}\|^2 + \|\operatorname{Cof}\overline{U}\|^2 + \det[\overline{U}]^2\right) - C_2^+\,. \tag{12.23}$$

Remark 12.6
This result is surprising at first glance since at face value $W^{\mathrm{Coss}}_{\mathrm{macro}}$ controls symmetric parts of the first two terms only.

Proof. To see Theorem **??** we first note that

$$\begin{aligned} W^{\mathrm{Coss}}_{\mathrm{macro}}(\overline{U}) &= 2\alpha\,\|\operatorname{sym}\overline{U} - 1\!\!1\|^2 + \alpha\operatorname{tr}\left[\overline{U} - 1\!\!1\right]^2 \\ &\quad + 2\beta\,\|\operatorname{sym}\operatorname{Cof}\overline{U} - 1\!\!1\|^2 + \beta\operatorname{tr}\left[\operatorname{Cof}\overline{U} - 1\!\!1\right]^2 \\ &\quad + \gamma\left((\det[\overline{U}] - 1)^2 + (\frac{1}{\det[\overline{U}]} - 1)^2\right) \geq 2\alpha\,\|\operatorname{sym}\overline{U} - 1\!\!1\|^2 + \mathcal{E}\,, \end{aligned} \tag{12.24}$$

where

$$\mathcal{E} := \alpha \operatorname{tr}\left[\overline{U} - \mathbb{1}\right]^2 + 2\beta\, \|\operatorname{sym}\operatorname{Cof}\overline{U} - \mathbb{1}\|^2 + \gamma\,(\det[\overline{U}] - 1)^2 . \tag{12.25}$$

Using the result shown in (**??**) we see that for some positive constants $C_1, C_2 > 0$, only depending on the material parameters (subsequently constants may change from line to line)

$$\begin{aligned}
W^{\mathrm{Coss}}_{\mathrm{macro}}(\overline{U}) &\geq 2\alpha\, \|\operatorname{sym}\overline{U} - \mathbb{1}\|^2 + C_1 \left(\operatorname{tr}\left[\overline{U}\right]^2 + \det[\overline{U}]^2 + \|\operatorname{Cof}\overline{U}\|^2\right) - C_2 \\
&= \frac{\alpha}{2}\, \|\overline{U} + \overline{U}^T - 2\,\mathbb{1}\|^2 + C_1 \left(\operatorname{tr}\left[\overline{U}\right]^2 + \det[\overline{U}]^2 + \|\operatorname{Cof}\overline{U}\|^2\right) - C_2 \\
&= \frac{\alpha}{2}\, \left(\|\overline{U} + \overline{U}^T\|^2 - 8\operatorname{tr}\left[\overline{U}\right] + 12\right) + C_1 \left(\operatorname{tr}\left[\overline{U}\right]^2 + \det[\overline{U}]^2 + \|\operatorname{Cof}\overline{U}\|^2\right) - C_2 \\
&= \frac{\alpha}{2}\, \left(2\|\overline{U}\|^2 + 2\operatorname{tr}\left[\overline{U}\right]^2 - 4\langle \operatorname{Cof}\overline{U}, \mathbb{1}\rangle - 8\operatorname{tr}\left[\overline{U}\right] + 12\right) \\
&\quad + C_1 \left(\operatorname{tr}\left[\overline{U}\right]^2 + \det[\overline{U}]^2 + \|\operatorname{Cof}\overline{U}\|^2\right) - C_2 \qquad (12.26) \\
&= \alpha\, \|\overline{U}\|^2 - 2\alpha\, \langle \operatorname{Cof}\overline{U}, \mathbb{1}\rangle - 4\,\alpha \operatorname{tr}\left[\overline{U}\right] \\
&\quad + C_1 \left(\operatorname{tr}\left[\overline{U}\right]^2 + \det[\overline{U}]^2 + \|\operatorname{Cof}\overline{U}\|^2\right) - C_2 .
\end{aligned}$$

The proof is finished by noting that quadratic terms dominate all other contributions. ■

Lemma 12.7
Let $X \in \mathbb{M}^{3\times 3}$. *Then*

$$\begin{aligned}
&\exists\, c^+ > 0 : \quad \forall\, X \in \mathbb{M}^{3\times 3} \\
&\|X + X^T\|^4 + \|\operatorname{Cof} X + (\operatorname{Cof} X)^T\|^2 \geq c^+ \left(\|X\|^4 + \|\operatorname{Cof} X\|^2\right) .
\end{aligned} \tag{12.27}$$

Proof. We proceed by contradiction and compactness. Our first observation is

$$\operatorname{Cof} X = A \in \mathfrak{so}(3) \;\Rightarrow\; \operatorname{Cof} X = 0 . \tag{12.28}$$

To see this, we compute (first for invertible $X \in GL(3)$)

$$\begin{aligned}
\operatorname{Cof} X = \det[X]\, X^{-T} \;&\Rightarrow\; \operatorname{Cof}\operatorname{Cof} X = \det[X]\, X , \\
\det[\operatorname{Cof} X] = \det[\det[X]\, X^{-T}] &= \det[X]^3 \det[X^{-T}] = \det[X]^2 , \\
\operatorname{Cof} X = A \in \mathfrak{so}(3) \;&\Rightarrow\; \det[\operatorname{Cof} X] = \det[A] \;\Rightarrow\; \det[X]^2 = 0 , \\
\operatorname{Cof} X = A \in \mathfrak{so}(3) \;&\Rightarrow\; \operatorname{Cof}\operatorname{Cof} X = \operatorname{Cof} A \;\Rightarrow\; \det[X]\, X = \operatorname{Cof} A \;\Rightarrow \\
\operatorname{tr}\left[\det[X]\, X\right] = \operatorname{tr}\left[\operatorname{Cof} A\right] &= \frac{1}{2}\|A\|^2 \;\Rightarrow\; 0 = \|A\|^2 .
\end{aligned} \tag{12.29}$$

Since every non-invertible matrix X can be approximated by invertible matrices all obtained formulas that make sense for non-invertible matrices are valid as well.

Next, observe that

$$\|X + X^T\|^4 + \|\operatorname{Cof} X + (\operatorname{Cof} X)^T\|^2 = 0 \;\Rightarrow\; \|X\|^4 + \|\operatorname{Cof} X\|^2 = 0 . \tag{12.30}$$

For this statement, we do not really need (**??**) since the cofactor of a skew-symmetric matrix in $\mathbb{M}^{3\times 3}$ is symmetric.

Now, assume that (**??**) does not hold. Then there is a sequence $X_k \in \mathbb{M}^{3\times 3}$ such that

$$\|X_k + X_k^T\|^4 + \|\operatorname{Cof} X_k + (\operatorname{Cof} X_k)^T\|^2 \to 0 \quad \text{but } \|X_k\|^4 + \|\operatorname{Cof} X_k\|^2 \geq \varepsilon > 0\,. \tag{12.31}$$

Thus, dividing by $\|X_k\|^4 + \|\operatorname{Cof} X_k\|^2$ we get

$$\begin{aligned}
&\frac{1}{\|X_k\|^4 + \|\operatorname{Cof} X_k\|^2}\left(\|X_k + X_k^T\|^4 + \|\operatorname{Cof} X_k + (\operatorname{Cof} X_k)^T\|^2\right) \to 0 \;\Leftrightarrow \\
&\|\frac{X_k}{(\|X_k\|^4 + \|\operatorname{Cof} X_k\|^2)^{(1/4)}} + \frac{X_k^T}{(\|X_k\|^4 + \|\operatorname{Cof} X_k\|^2)^{(1/4)}}\|^4 \\
&\quad + \|\frac{1}{\sqrt{\|X_k\|^4 + \|\operatorname{Cof} X_k\|^2}}(\operatorname{Cof} X_k + (\operatorname{Cof} X_k)^T)\|^2 \to 0 \;\Leftrightarrow \\
&\|\frac{X_k}{(\|X_k\|^4 + \|\operatorname{Cof} X_k\|^2)^{(1/4)}} + \frac{X_k^T}{(\|X_k\|^4 + \|\operatorname{Cof} X_k\|^2)^{(1/4)}}\|^4 \\
&\quad \|\operatorname{Cof} \frac{X_k}{(\|X_k\|^4 + \|\operatorname{Cof} X_k\|^2)^{(1/4)}} + \operatorname{Cof} \frac{X_k^T}{(\|X_k\|^4 + \|\operatorname{Cof} X_k\|^2)^{(1/4)}}\|^2 \to 0\,.
\end{aligned} \tag{12.32}$$

However,

$$\frac{X_k}{(\|X_k\|^4 + \|\operatorname{Cof} X_k\|^2)^{(1/4)}} \tag{12.33}$$

is a bounded sequence. Thus we may extract a convergent subsequence (not relabelled) such that

$$\frac{X_k}{(\|X_k\|^4 + \|\operatorname{Cof} X_k\|^2)^{(1/4)}} \to \frac{X}{(\|X\|^4 + \|\operatorname{Cof} X\|^2)^{(1/4)}}\,, \tag{12.34}$$

and by continuity of the cofactor

$$\begin{aligned}
&\|\frac{X}{(\|X\|^4 + \|\operatorname{Cof} X\|^2)^{(1/4)}} + \frac{X^T}{(\|X\|^4 + \|\operatorname{Cof} X\|^2)^{(1/4)}}\|^4 \\
&\quad + \|\operatorname{Cof} \frac{X}{(\|X\|^4 + \|\operatorname{Cof} X\|^2)^{(1/4)}} + \operatorname{Cof} \frac{X^T}{(\|X\|^4 + \|\operatorname{Cof} X\|^2)^{(1/4)}}\|^2 = 0\,.
\end{aligned} \tag{12.35}$$

This is the contradiction. ■

Corollary 12.8
Let $X \in \mathbb{M}^{3\times 3}$ *and* $p \geq 1$. *Then the local estimate holds*

$$\begin{aligned}
&\exists\, c^+ > 0 : \quad \forall\, X \in \mathbb{M}^{3\times 3} : \\
&\quad \|X + X^T\|^{2p} + \|\operatorname{Cof} X + (\operatorname{Cof} X)^T\|^p \geq c^+ \left(\|X\|^{2p} + \|\operatorname{Cof} X\|^p\right)\,.
\end{aligned} \tag{12.36}$$

Remark 12.9
The estimate $\|\operatorname{Cof} X + \operatorname{Cof} X^T\|^2 \geq c^+ \|\operatorname{Cof} X\|^2$ *does not hold as can be seen for* $X_k = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & k \end{pmatrix}$. *The occuring problem is related to the fact that for a not necessarily convergent sequence* Z_k, *the sequence* $\operatorname{Cof} Z_k$ *may converge to some limit* Y, *which is not a cofactor. Note, however, that* $\operatorname{sym}\operatorname{Cof} X = 0 \Rightarrow \operatorname{Cof} X = A \in \mathfrak{so}(3) \Rightarrow \operatorname{Cof} X = 0$ *as has been shown in* (**??**).

12.1 Estimates on $\operatorname{Cof} X$ in terms of $\operatorname{sym}\operatorname{Cof} X$, $\operatorname{tr}[X]$ and $\det[X]$

In the existence theorem one may observe that

$$4\|\operatorname{sym}\operatorname{Cof} X - 1\!\!1\|^2 = \|\operatorname{Cof} X + (\operatorname{Cof} X)^T - 2 1\!\!1\|^2 \tag{12.37}$$

remains bounded when integrated over the domain. It is clear that this is not sufficient to directly obtain a bound of $\|\operatorname{Cof} X\|^2$. However, such a bound may be established, provided we know already independent bounds on $\operatorname{tr}[X]$ and $\det[X]$. To see this, we compute for invertible X

$$\begin{aligned}
&\|\operatorname{Cof} X + (\operatorname{Cof} X)^T - 2 1\!\!1\|^2 \\
&= \|\operatorname{Cof} X + (\operatorname{Cof} X)^T\|^2 - 8\langle \operatorname{Cof} X, 1\!\!1\rangle + 12 \\
&= 2\|\operatorname{Cof} X\|^2 + 2\,\langle \operatorname{Cof} X, (\operatorname{Cof} X)^T\rangle - 8\langle \operatorname{Cof} X, 1\!\!1\rangle + 12 \\
&= 2\|\operatorname{Cof} X\|^2 + 2\,\langle \det[X]\, X^{-T}, \det[X]\, X^{-1}\rangle - 8\langle \operatorname{Cof} X, 1\!\!1\rangle + 12 \\
&= 2\|\operatorname{Cof} X\|^2 + 2\det[X]^2 \langle X^{-T}, X^{-1}\rangle - 8\langle \operatorname{Cof} X, 1\!\!1\rangle + 12 \\
&= 2\|\operatorname{Cof} X\|^2 + 2\det[X]^2 \langle 1\!\!1, (X^{-1})^2\rangle - 8\,\langle \operatorname{Cof} X, 1\!\!1\rangle + 12 \\
&\qquad \text{use } 2\operatorname{tr}\left[Y^2\right] = 2\operatorname{tr}[Y]^2 - 4\,\langle 1\!\!1, \operatorname{Cof} Y\rangle \\
&= 2\|\operatorname{Cof} X\|^2 + 2\det[X]^2\left(2\operatorname{tr}\left[X^{-1}\right]^2 - 4\,\langle 1\!\!1, \operatorname{Cof}(X^{-1})\rangle\right) - 8\,\langle \operatorname{Cof} X, 1\!\!1\rangle + 12 \\
&= 2\|\operatorname{Cof} X\|^2 + 2\det[X]^2\left(2\operatorname{tr}\left[X^{-1}\right]^2 - 4\det[X^{-1}]\langle 1\!\!1, (X^{-1})^T\rangle\right) - 8\,\langle \operatorname{Cof} X, 1\!\!1\rangle + 12 \\
&= 2\|\operatorname{Cof} X\|^2 + 2\det[X]^2\left(2\operatorname{tr}\left[X^{-1}\right]^2 - \frac{4}{\det[X]}\,\langle 1\!\!1, X\rangle\right) - 8\,\langle \operatorname{Cof} X, 1\!\!1\rangle + 12 \\
&= 2\|\operatorname{Cof} X\|^2 + \underbrace{4\det[X]^2 \operatorname{tr}\left[X^{-1}\right]^2}_{\geq 0} - 8\det[X]\langle 1\!\!1, X\rangle - 8\,\langle \operatorname{Cof} X, 1\!\!1\rangle + 12 \\
&\geq 2\|\operatorname{Cof} X\|^2 - 8\det[X]\langle 1\!\!1, X\rangle - 8\,\langle \operatorname{Cof} X, 1\!\!1\rangle + 12 \qquad (12.38) \\
&\geq 2\|\operatorname{Cof} X\|^2 - 8\det[X]\langle 1\!\!1, X\rangle - 8\sqrt{3}\|\operatorname{Cof} X\| + 12 \\
&\quad \text{use Young's inequality on the middle term} \\
&\geq 2\|\operatorname{Cof} X\|^2 - 4\left(\varepsilon^2\det[X]^2 + \frac{1}{\varepsilon^2}\langle 1\!\!1, X\rangle^2\right) - 8\sqrt{3}\|\operatorname{Cof} X\| + 12\,.
\end{aligned}$$

This shows that

$$\begin{aligned}
\mathcal{E} &= \alpha\,\mathrm{tr}\,[X-1\!\!1]^2 + 2\beta\,\|\,\mathrm{sym}\,\mathrm{Cof}\,X - 1\!\!1\|^2 + \gamma\,(\det[X]-1)^2\\
&= \alpha\,\mathrm{tr}\,[X-1\!\!1]^2 + 2\beta\,\|\,\mathrm{sym}\,\mathrm{Cof}\,X - 1\!\!1\|^2 + \gamma\det[X]^2 - 2\gamma\det[X] + \gamma\\
&\geq \alpha\,\mathrm{tr}\,[X-1\!\!1]^2 + \beta\left(\|\,\mathrm{Cof}\,X\|^2 - 2\,(\varepsilon^2\det[X] + \frac{1}{\varepsilon^2}\langle 1\!\!1, X\rangle^2) - 4\sqrt{3}\|\,\mathrm{Cof}\,X\| + 6\right)\\
&\quad + \gamma\det[X]^2 - 2\gamma\det[X]^2 + \gamma\\
&= \alpha\,\mathrm{tr}\,[X]^2 - 6\alpha\,\mathrm{tr}\,[X] + 9\alpha + \beta\left(\|\,\mathrm{Cof}\,X\|^2 - 2\,(\varepsilon^2\det[X] + \frac{1}{\varepsilon^2}\langle 1\!\!1, X\rangle^2) - 4\sqrt{3}\|\,\mathrm{Cof}\,X\| + 6\right)\\
&\quad + \gamma\det[X]^2 - 2\,\gamma\det[X] + \gamma\\
&= \frac{\alpha}{2}\,\mathrm{tr}\,[X]^2 + (\frac{\alpha}{2} - \frac{2\beta}{\varepsilon^2})\,\mathrm{tr}\,[X]^2 - 6\alpha\,\mathrm{tr}\,[X] + 9\alpha + \beta\,(\|\,\mathrm{Cof}\,X\|^2 - 4\sqrt{3}\|\,\mathrm{Cof}\,X\| + 6)\\
&\quad + \frac{\gamma}{2}\det[X]^2 + (\frac{\gamma}{2} - 2\beta\varepsilon^2)\det[X]^2 - 2\gamma\det[X] + \gamma\\
&\qquad \text{if } (\frac{\alpha}{2} - \frac{2\beta}{\varepsilon^2}) \geq 0 \text{ and } (\frac{\gamma}{2} - 2\beta\varepsilon^2) \geq 0 \text{ we obtain}\\
&\geq \frac{\alpha}{2}\mathrm{tr}\,[X]^2 - 6\alpha\,\mathrm{tr}\,[X] + 9\alpha + \beta\,(\|\,\mathrm{Cof}\,X\|^2 - 4\sqrt{3}\|\,\mathrm{Cof}\,X\| + 6) + \frac{\gamma}{2}\det[X]^2 - 2\gamma\det[X] + \gamma\\
&= \frac{\alpha}{2}\mathrm{tr}\,[X]^2 + \frac{\gamma}{2}\det[X]^2 + \beta\,\|\,\mathrm{Cof}\,X\|^2 - 6\alpha\,\mathrm{tr}\,[X] - 4\sqrt{3}\,\beta\,\|\,\mathrm{Cof}\,X\| - 2\gamma\det[X]\\
&\quad + (9\alpha + 6\,\beta + \gamma)\\
&\geq \min\{\frac{\alpha}{2}, \frac{\gamma}{2}, \beta\}\,(\mathrm{tr}\,[X]^2 + \det[X]^2 + \|\,\mathrm{Cof}\,X\|^2)\\
&\quad - \max\{6\alpha, 4\sqrt{3}\beta, 2\gamma\}\,(|\mathrm{tr}\,[X]| + \|\,\mathrm{Cof}\,X\| + |\det[X]|) + (9\alpha + 6\,\beta + \gamma)\\
&\quad \text{abbreviate } \xi = (|\mathrm{tr}\,[X]|, |\det[X]|, \|\,\mathrm{Cof}\,X\|)\\
&\geq a^+\|\xi\|_2^2 - b^+\|\xi\|_1 - c^+ \geq a^+\|\xi\|_2^2 - \tilde{b}^+\|\xi\|_2 + c^+\,, \qquad (12.39)
\end{aligned}$$

where we have used the estimate of the 1-norm in terms of the 2-norm. Hence, in terms of generic constants a^+, b^+, c^+ only depending on the material parameters α, β, γ we obtain that

$$\xi^2 \leq K(\alpha,\beta,\gamma) + \mathcal{E} \;\Rightarrow\; \mathrm{tr}\,[X]^2 + \det[X]^2 + \|\,\mathrm{Cof}\,X\|^2 \leq K + \mathcal{E}\,. \qquad (12.40)$$

The result has been shown using the invertibility of X. Since every matrix X can be approximated by invertible matrices the result holds true also for non-invertible matrices.

12.2 Other useful observations

Lemma 12.10
Let $F \in \mathrm{GL}(3)$. Then

$$\forall\, N \in \mathbb{S}^2: \quad \|F.N\|^2\,\|F^{-T}.N\|^2 \geq 1\,. \qquad (12.41)$$

Proof. We rewrite the inequality in terms of $C = F^TF$. This yields

$$\|F.N\|^2\,\|F^{-T}.N\|^2 = \langle C.N, N\rangle\,\langle C^{-1}.N, N\rangle\,. \qquad (12.42)$$

After diagonalizing C we obtain with $C = Q^T D Q$ and $\eta = Q.N$

$$\langle C.N, N\rangle \, \langle C^{-1}.N, N\rangle = \langle D.\eta, \eta\rangle \, \langle D^{-1}.\eta, \eta\rangle \,. \tag{12.43}$$

Let $\lambda_1, \lambda_2, \lambda_3$ be the diagonal entries of D then

$$\begin{aligned} \langle D.\eta, \eta\rangle \, \langle D^{-1}.\eta, \eta\rangle &= (\lambda_1 \, \eta_1^2 + \lambda_2 \, \eta_2^2 + \lambda_3 \, \eta_3^2) \, (\frac{1}{\lambda_1} \, \eta_1^2 + \frac{1}{\lambda_2} \, \eta_2^2 + \frac{1}{\lambda_3} \, \eta_3^2) \\ &= (\eta_1^2 + \eta_2^2 + \eta_3^2)^2 + \left(\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1} - 2\right) \, \eta_1^2 \eta_2^2 \\ &+ \left(\frac{\lambda_1}{\lambda_3} + \frac{\lambda_3}{\lambda_1} - 2\right) \, \eta_1^2 \eta_3^2 + \left(\frac{\lambda_3}{\lambda_2} + \frac{\lambda_2}{\lambda_3} - 2\right) \, \eta_2^2 \eta_3^2 \,. \end{aligned} \tag{12.44}$$

Since

$$\left(\frac{\lambda_i}{\lambda_j} + \frac{\lambda_j}{\lambda_i} - 2\right) = \frac{(\lambda_i - \lambda_j)^2}{\lambda_i \, \lambda_j} \geq 0 \,, \tag{12.45}$$

the result follows by observing that $(\eta_1^2 + \eta_2^2 + \eta_3^2)^2 = 1$. ■

Corollary 12.11 (Nearly conformal energy)

$$\forall \; N \in \mathbb{S}^2 : \quad \|F.N\|^2 \, \| \operatorname{Cof} F.N\|^2 \geq \det[F]^2 \,. \tag{12.46}$$

Moreover

$$\forall \, N \in \mathbb{S}^2 : \quad W_{\text{conf}}(F, N) := \frac{\|F.N\|^2 \, \| \operatorname{Cof} F.N\|^2}{\det[F]^2} - 1 \geq 0 \,, \tag{12.47}$$

and

$$W_{\text{conf}}(\mathbb{R}^+ \, \mathrm{SO}(3) \, , N) = 0 \,, \quad \textit{nearly conformal local energy} \,.$$

Lemma 12.12 (Antmann's relation)
For $F \in \mathrm{GL}(3)$ *let* $C = F^T F$. *If* $\det[C] = 1$ *and* $\operatorname{tr}[C] = 3$ *then* $C = 1\!\!1$ *and* $F \in \mathrm{SO}(3)$.

Proof. We may orthogonally diagonalize C and obtain for the eigenvalues that

$$\begin{aligned} 1 &= \det[C] = \lambda_1 \, \lambda_2 \, \lambda_3 \,, \\ 3 &= \operatorname{tr}[C] = \lambda_1 + \lambda_2 + \lambda_3 \,. \end{aligned} \tag{12.48}$$

The arithmetic-geometric inequality $(\lambda_1 \, \lambda_2 \, \lambda_3)^{1/3} \leq \frac{1}{3}(\lambda_1 + \lambda_2 + \lambda_3)$ evaluated for (**??**) yields, however, equality. This is only possible if $\lambda_i = 1$. Thus $C = 1\!\!1$. ■

Remark 12.13
Note that it is decisive that $C = F^T F$ *as shows the example*

$$X = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \,, \quad \det[X] = 1 \,, \quad \operatorname{tr}[X] = 3 \,, \quad \textit{but } X \neq 1\!\!1 \,. \tag{12.49}$$

Lemma 12.14 (Ball's observation)

$$W(F) := \frac{1}{2}\|F\|^2 + \frac{1}{\det[F]} - \frac{5}{2} \geq 0\,, \quad W(\mathrm{SO}(3)) = 0\,. \tag{12.50}$$

Proof. We show first criticality of SO(3).

$$DW(F).H := \langle F, H\rangle - \det[F]^{-2}\,\langle \mathrm{Cof}\, F, H\rangle = 0 \quad \text{for } F \in \mathrm{SO}(3)\,. \tag{12.51}$$

The equation for critical points yields

$$\begin{aligned} F = \det[F]^{-2}\ \mathrm{Cof}\, F &= \frac{1}{\det[F]}\, F^{-T} \ \Rightarrow \\ FF^T = \frac{1}{\det[F]}\, 1\!\!1 \ &\Leftrightarrow\ F = \beta\, R\,, \quad \beta > 0\,. \end{aligned} \tag{12.52}$$

In this case, the energy level reads $\frac{3}{2}\,\beta^2 + \frac{1}{\beta^3}$. This is minimal w.r.t. positive β exactly for$\beta = 1$. Thus among all critical points $F = R$ is minimal. ■

Remark 12.15
*Note that Ball's energy (**??**) satisfies the assumptions of Lemma **??** but is polyconvex.*

Conjecture 12.16
Assume that $W : \mathrm{GL}(3) \mapsto \mathbb{R}$ *is left-*O(3) *indifferent and*

$$\min_{F\in \mathrm{GL}(3)} W(F) = W(1\!\!1)\,, \quad W(F) > W(1\!\!1) \quad \forall\, F \notin \mathrm{O}(3)\,. \tag{12.53}$$

Then W *is not Legendre-Hadamard elliptic? The idea would be to generalize the case of energies of non Legendre-Hadamard ellipticity given in the strain-type form*

$$W(F) = \|F^TF - 1\!\!1\|^2\,. \tag{12.54}$$

Lemma 12.17 (Quasiconformal energy)
Let $F \in \mathrm{GL}(3)$. *Then*

$$\forall\, F \in \mathrm{GL}(3) : \quad W(F) := \|F\|^2\, \|F^{-1}\|^2 - 9 \geq 0\,, \tag{12.55}$$

and $W(\mathbb{R}^+\ \mathrm{SO}(3)) = 0$.

Proof. First we observe

$$\|F\|^2\, \|F^{-1}\|^2 - 9 = \|F\|^2\, \|F^{-T}\|^2 = \mathrm{tr}\,[C]\, \mathrm{tr}\,\left[C^{-1}\right] - 9\,. \tag{12.56}$$

In terms of the eigenvalues of C we have after diagonalization

$$\begin{aligned} &(\lambda_1 + \lambda_2 + \lambda_3)\left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_2}\right) - 9 \\ &= 1 + \frac{\lambda_1}{\lambda_2} + \frac{\lambda_1}{\lambda_3} + 1 + \frac{\lambda_2}{\lambda_1} + \frac{\lambda_2}{\lambda_3} + 1 + \frac{\lambda_3}{\lambda_1} + \frac{\lambda_3}{\lambda_2} - 9 \\ &= \left(\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1} - 2\right) + \left(\frac{\lambda_1}{\lambda_3} + \frac{\lambda_3}{\lambda_1} - 2\right) + \left(\frac{\lambda_2}{\lambda_3} + \frac{\lambda_3}{\lambda_3} - 2\right)\,. \end{aligned} \tag{12.57}$$

Using the binomial formula we see that each bracket is individually positive. ■

13 Relaxed M-condition

For the time being the differences between various forms of Ogden-type energies and the proposed classical energies with voumetric/isochoric split remained unclear since both ansatzes were (uniformly) polyconvex and verified various growth conditions sufficient for the solvability of the boundary value problem. The first appearing difference is that only a judicious choice of positive parameters made Ogden type energies stress free whereas energies with vol/iso-split allowed for arbitrary combinations of positive parameters. As example for Ogden type energie consider e.g.

$$\begin{aligned} W_{Ogden}(F) &= c_1 \cdot \|F\|^2 + c_2 \cdot \|\operatorname{Adj} F\|^2 + c_3 \cdot (\det[F])^2 - c_4 \cdot \ln \det[F] \\ &= c_1 \cdot \operatorname{tr}(C) + c_2 \cdot \operatorname{tr}(\operatorname{Cof} C) + c_3 \cdot \det[C] - \frac{c_4}{2} \cdot \ln \det[C] \\ &= c_1 \cdot \|U\|^2 + c_2 \cdot \|\operatorname{Adj} U\|^2 + c_3 \cdot (\det[U])^2 - c_4 \cdot \ln \det[U] \end{aligned}$$

where $F = R \cdot U$ and $C = F^T F$. In [**?**] the author introduced the so called M-condition. In a simplified setting for hyperelastic materials the M-condition amounts to the requirement of convexity of the strain energy W as a function of U. Here we want to relax this physically appealing requirement to

Definition 13.1 (Relaxed M-condition, RM)
We say that the strain energy W fulfills the relaxed M-condition whenever

$$\begin{aligned} &\forall\, H \in \mathrm{PSym} \quad D_U^2 W(U).(H,H) \geq 0 \\ &\forall\, (U_1 - U_2) \in \mathrm{PSym} \quad \langle D_U W(U_1) - D_U W(U_2), U_1 - U_2 \rangle \geq 0 \end{aligned}$$

■

Remark 13.2
This new condition is still another expression of the intuitive idea that increased strain should lead to increased stress; here increased strain refers to the quantity $U_1 - U_2 \in \mathrm{PSym}$, i.e the two local deformation states differ only by a pure multi-axial stretch and the corresponding stresses $D_U W$ should then increase. If $U_1 - U_2 \notin \mathrm{PSym}$ we cannot truly compare the two local states; in some directions there may be an increase of stretch in others stretch may decrease. The RM-condition does not impose any restriction on those (indefinite) states whereas the original M-condition does. It is clear that RM does not imply M but M is sufficient for RM. Even so neither the original M-condition nor the relaxed M-condition are sufficient to garantee the existence of solutions to corresponding boundary value problems the latter RM-condition may be used to distinguish Ogden ansatzes and volumetric/isochoric splits.

Theorem 13.3 (Neo-Hooke and M-condition)
Let

$$W_{Neo-Hooke}(F) = c_1 \cdot \|U\|^2 - c_4 \cdot \ln \det[U]\,,$$

where $c_i > 0$. Then $W_{Neo-Hooke}$ verifies the original M-condition.

Proof. We simply compute the second derivative for $W_{Neo-Hooke}$:

$$\begin{aligned} D_U W_{Neo-Hooke}(U).H &= c_1\langle U, H\rangle - c_4\langle U^{-1}, H\rangle \\ D_U^2 W_{Neo-Hooke}(U).(H,H) &= c_1\|H\|^2 + c_4\langle U^{-1}HU^{-1}, H\rangle > 0 \end{aligned}$$

if $H \in \operatorname{Sym}$ since $U \in \operatorname{PSym}$. ■

Theorem 13.4 (Relaxed M-condition)
Let

$$\begin{aligned} W_{Ogden}(F) &= c_1 \cdot \|U\|^2 + c_2 \cdot \|\operatorname{Adj} U\|^2 + c_3 \cdot (\det[U])^2 - c_4 \cdot \ln\det[U]\,, \\ W_{vol,1}(F) &= \det[U]^4 - 4\ln\det[U]\,, \\ W_{vol,2}(F) &= \det[U]^4 + \frac{1}{\det[U]^4} - 2\,, \end{aligned}$$

where $c_i > 0$. Then W_{Ogden} and $W_{vol,i}$ verify the relaxed M-condition.

Proof. We “simply” compute the second derivative for W_{Ogden}:

$$\begin{aligned} D_U W_{Ogden}(U).H &= c_1\langle U, H\rangle + c_2\langle \operatorname{Adj} U, D.\operatorname{Adj} U.H\rangle + 2c_3\langle \operatorname{Adj} U^T, H\rangle - c_4\frac{1}{\det[U]}\langle \operatorname{Adj} U^T, H\rangle \\ &= c_1\langle U, H\rangle + c_2\langle \operatorname{Adj} U, D.\operatorname{Adj} U.H\rangle + 2c_3\langle \operatorname{Adj} U^T, H\rangle - c_4\langle U^{-1}, H\rangle \\ D_U^2 W_{Ogden}(U).(H,H) &= c_1\|H\|^2 + c_2\langle D\operatorname{Adj} U.H, D.\operatorname{Adj} U.H\rangle + c_2\langle \operatorname{Adj} U, D^2.\operatorname{Adj} U.(H,H)\rangle + \\ &\quad 2c_3\langle D\operatorname{Adj} U^T.H, H\rangle + c_4\langle U^{-1}HU^{-1}, H\rangle \\ &= c_1\|H\|^2 + c_2\|D\operatorname{Adj} U.H\|^2 + c_2\langle \operatorname{Adj} U, 2\operatorname{Adj} H\rangle + \\ &\quad 2c_3\langle D\operatorname{Adj} U.H, H\rangle + c_4\langle HU^{-1}, U^{-1}H\rangle \\ &= c_1\|H\|^2 + c_2\|D\operatorname{Adj} U.H\|^2 + 2c_2\langle \operatorname{Adj} U, \operatorname{Adj} H\rangle + \\ &\quad 2c_3\langle \operatorname{Adj} U[\langle U^{-1}, H\rangle \mathbb{1} - HU^{-1}], H\rangle + c_4\langle HU^{-1}, U^{-1}H\rangle \\ &= c_1\|H\|^2 + c_2\|D\operatorname{Adj} U.H\|^2 + 2c_2\langle \operatorname{Adj} U, \operatorname{Adj} H\rangle + \\ &\quad 2c_3\det[U]\left[\langle U^{-1}, H\rangle^2 - \langle HU^{-1}, U^{-1}H\rangle\right] + c_4\langle HU^{-1}, U^{-1}H\rangle \\ &= c_1\|H\|^2 + c_2\|D\operatorname{Adj} U.H\|^2 + 2c_2\langle \operatorname{Adj} U, \operatorname{Adj} H\rangle + \\ &\quad 2c_3\det[U]\left[\langle U^{-1}, H\rangle^2 - \langle HU^{-1}, (HU^{-1})^T\rangle\right] + c_4\langle HU^{-1}, U^{-1}H\rangle \\ &= c_1\|H\|^2 + c_2\|D\operatorname{Adj} U.H\|^2 + 2c_2\langle \operatorname{Adj} U, \operatorname{Adj} H\rangle + \\ &\quad 2c_3 \cdot 2\langle \operatorname{Adj} H, U\rangle + c_4\langle HU^{-1}, U^{-1}H\rangle > 0\,, \end{aligned}$$

since every summand is positive for $H \in \operatorname{PSym}$ and $U \in \operatorname{PSym}$. We repeat the same procedure for $W_{vol,1}$. This yields

$$\begin{aligned} D_U W_{vol,1}(U).H &= 4\det[U]^3\langle \operatorname{Adj} U, H\rangle - 4\frac{1}{\det[U]}\langle \operatorname{Adj} U, H\rangle \\ &= 4\det[U]^3\langle \operatorname{Adj} U, H\rangle - 4\langle U^{-1}, H\rangle\,, \\ D_U^2 W_{vol,1}.(H,H) &= 12\det[U]^2\langle \operatorname{Adj} U, H\rangle^2 + 4\det[U]^3\langle D(\operatorname{Adj} U).H, H\rangle + 4\langle U^{-1}HU^{-1}, H\rangle \\ &= 12\det[U]^2\langle \operatorname{Adj} U, H\rangle^2 + 4\det[U]^3 \cdot 2\langle \operatorname{Adj} H, U\rangle + 4\langle HU^{-1}, U^{-1}H\rangle > 0 \end{aligned}$$

if $H \in \mathrm{PSym}$ and $U \in \mathrm{PSym}$. Now the case for $W_{vol,2}$:

$$\begin{aligned}
D_U W_{vol,2}(F) &= 4\det[U]^3 \langle \operatorname{Adj} U, H\rangle - 4\det[U]^{-5}\langle \operatorname{Adj} U, H\rangle \\
&= 4\det[U]^3 \langle \operatorname{Adj} U, H\rangle - 4\det[U]^{-4}\langle U^{-1}, H\rangle\,, \\
D_U^2 W_{vol,2}.(H,H) &= 12\det[U]^2 \langle \operatorname{Adj} U, H\rangle^2 + 4\det[U]^3\langle D(\operatorname{Adj} U).H, H\rangle + \\
&\quad 16\det[U]^{-5}\langle \operatorname{Adj} U, H\rangle\langle U^{-1}, H\rangle - 4\det[U]^{-4}\langle -U^{-1}HU^{-1}, H\rangle \\
&= 12\det[U]^2 \langle \operatorname{Adj} U, H\rangle^2 + 8\det[U]^3\langle \operatorname{Adj} H, U\rangle + \\
&\quad 16\det[U]^{-5}\langle \operatorname{Adj} U, H\rangle\langle U^{-1}, H\rangle + 4\det[U]^{-4}\langle U^{-1}HU^{-1}, H\rangle > 0
\end{aligned}$$

if $H \in \mathrm{PSym}$ and $U \in \mathrm{PSym}$. ■

Lemma 13.5 (Failure of strict RM)
Consider as examples

$$\begin{aligned}
W_1(F) &= \frac{\mu}{4}\|F^TF - 1\!\!1\|^2 + \frac{\lambda}{8}\operatorname{tr}(F^TF - 1\!\!1)^2 = \frac{\mu}{4}\|U^2 - 1\!\!1\|^2 + \frac{\lambda}{8}\operatorname{tr}(U^2 - 1\!\!1)^2\,, \\
W_2(F) &= \left(\left[\frac{\|F\|^2}{\det[F]^{\frac{2}{3}}}\right]^3 - 3^3\right) = \left(\left[\frac{\|U\|^2}{\det[U]^{\frac{2}{3}}}\right]^3 - 3^3\right)\,, \\
W_3(F) &= (\ln\det[U])^{2p}\,.
\end{aligned}$$

Then W_1, W_2 and W_3 fail to verify the strict RM-condition.

Proof. In both cases we compute the second differential.

$$\begin{aligned}
D_U^2 W_1(U).(H,H) = 2\frac{\mu}{4}\left(\|UH + HU\|^2 + \langle U^2 - 1\!\!1, 2H^2\rangle\right) + \\
+\frac{\lambda}{8}2\left(\langle UH + HU, 1\!\!1\rangle^2 + 4\langle U^2 - 1\!\!1, 1\!\!1\rangle\,\langle H^2, 1\!\!1\rangle\right)\,.
\end{aligned}$$

Choose $U = m \cdot 1\!\!1$ and $H = \xi \otimes \xi$ with $\|\xi\| = 1$. Then $H^2 = \xi\otimes\xi$ and

$$\begin{aligned}
D^2 W_1(U).(H,H) &= \frac{\mu}{4}2\left(4m^2\|\xi\otimes\xi\|^2 + 2\langle m^2 1\!\!1 - 1\!\!1, \xi\otimes\xi\rangle\right) + \frac{\lambda}{8}2\left(4m^2 + 4\cdot 3(m^2-1)\right) \\
&= \frac{\mu}{4}2\left(4m^2\|\xi\otimes\xi\|^2 + 2\langle m^2 1\!\!1 - 1\!\!1, \xi\otimes\xi\rangle\right) + \frac{\lambda}{8}2\left(4m^2 + 4\cdot 3(m^2-1)\right) \\
&= \frac{\mu}{4}2\left(4m^2\cdot 1 + 2(m^2-1)\cdot 1\right) + \frac{\lambda}{8}2\left(4m^2 + 4\cdot 3(m^2-1)\right) \\
&< 0
\end{aligned}$$

if $m \in \mathbb{R}$ is sufficiently small. For the second term:

$$
\begin{aligned}
D_U W_2(U).H &= 3\left[\frac{\|U\|^2}{\det[U]^{\frac{2}{3}}}\right]^2 \Bigg[2\langle U, H\rangle \det[U]^{\frac{-2}{3}} - \|U\|^2 \frac{2}{3}\det[U]^{\frac{-5}{3}}\langle \operatorname{Adj} U, H\rangle\Bigg] \\
D_U^2 W_2(U).(H,H) &= 6\left[\frac{\|U\|^2}{\det[U]^{\frac{2}{3}}}\right] \cdot \Bigg[2\langle U, H\rangle \det[U]^{\frac{-2}{3}} - \|U\|^2 \frac{2}{3}\det[U]^{\frac{-5}{3}}\langle \operatorname{Adj} U, H\rangle\Bigg]^2 + \\
&\quad 3\left[\frac{\|U\|^2}{\det[U]^{\frac{2}{3}}}\right]^2 \Bigg[2\|H\|^2 \det[U]^{\frac{-2}{3}} - 2\langle U, H\rangle \frac{2}{3}\det[U]^{\frac{-5}{3}}\langle \operatorname{Adj} U, H\rangle - \\
&\quad 2\langle U, H\rangle \frac{2}{3}\det[U]^{\frac{-5}{3}}\langle \operatorname{Adj} U, H\rangle - \|U\|^2 \frac{2}{3}\frac{-5}{3}\det[U]^{\frac{-8}{3}}\langle \operatorname{Adj} U, H\rangle^2 - \\
&\quad \|U\|^2 \frac{2}{3}\det[U]^{\frac{-5}{3}}\langle D \operatorname{Adj} U.H, H\rangle\Bigg] \\
&= 6\left[\frac{\|U\|^2}{\det[U]^{\frac{2}{3}}}\right] \cdot 4\det[U]^{\frac{-4}{3}} \Bigg[\langle U, H\rangle - \|U\|^2 \frac{1}{3}\langle U^{-1}, H\rangle\Bigg]^2 + \\
&\quad 3\left[\frac{\|U\|^2}{\det[U]^{\frac{2}{3}}}\right]^2 \det[U]^{\frac{-2}{3}} \Bigg[2\|H\|^2 - 2\langle U, H\rangle \frac{2}{3}\langle U^{-1}, H\rangle - 2\langle U, H\rangle \frac{2}{3}\langle U^{-1}, H\rangle - \\
&\quad \|U\|^2 \frac{2}{3}\frac{-5}{3}\langle U^{-1}, H\rangle^2 \\
&\quad -\|U\|^2 \frac{2}{3}\det[U]^{\frac{-3}{3}} \underbrace{\det[U]\left(\langle U^{-1}, H\rangle^2 - \langle U^{-1}H, HU^{-1}\rangle\right)}_{\langle D \operatorname{Adj} U.H, H\rangle = 2\langle \operatorname{Adj} H, U\rangle}\Bigg] \\
&= 24\frac{\|U\|^2}{\det[U]^2}\Bigg[\langle U, H\rangle - \frac{1}{3}\|U\|^2\langle U^{-1}, H\rangle\Bigg]^2 + \\
&\quad 3\left[\frac{\|U\|^2}{\det[U]^{\frac{2}{3}}}\right]^2 \det[U]^{\frac{-2}{3}} \Bigg[2\|H\|^2 - \frac{4}{3}\langle U, H\rangle\langle U^{-1}, H\rangle - \\
&\quad \frac{4}{3}\langle U, H\rangle\langle U^{-1}, H\rangle + \frac{10}{9}\|U\|^2\langle U^{-1}, H\rangle^2 - \frac{4}{3}\|U\|^2\langle \operatorname{Adj} H, \operatorname{Adj} U^{-1}\rangle\Bigg] \\
&= 24\frac{\|U\|^2}{\det[U]^2}\Bigg[\langle U, H\rangle - \frac{1}{3}\|U\|^2\langle U^{-1}, H\rangle\Bigg]^2 + \\
&\quad 3\left[\frac{\|U\|^2}{\det[U]^{\frac{2}{3}}}\right]^2 \det[U]^{\frac{-2}{3}} \Bigg[2\|H\|^2 - \frac{8}{3}\langle U, H\rangle\langle U^{-1}, H\rangle \\
&\quad +\frac{10}{9}\|U\|^2\langle U^{-1}, H\rangle^2 - \frac{4}{3}\|U\|^2\langle \operatorname{Adj} H, \operatorname{Adj} U^{-1}\rangle\Bigg] \\
&= 24\frac{\|U\|^2}{\det[U]^2}\Bigg[\langle U, H\rangle - \frac{1}{3}\|U\|^2\langle U^{-1}, H\rangle\Bigg]^2 +
\end{aligned}
$$

$$
\begin{aligned}
&6\frac{\|U\|^4}{\det[U]^2}\left[\|H\|^2-\frac{4}{3}\langle U,H\rangle\langle U^{-1},H\rangle\right.\\
&\quad\left.+\frac{5}{9}\|U\|^2\langle U^{-1},H\rangle^2-\frac{4}{3}\|U\|^2\langle \operatorname{Adj}H,\operatorname{Adj}U^{-1}\rangle\right]\\
&=6\frac{\|U\|^2}{\det[U]^2}\left(4\left[\langle U,H\rangle-\frac{1}{3}\|U\|^2\langle U^{-1},H\rangle\right]^2+\right.\\
&\quad\|U\|^2\|H\|^2-\frac{4}{3}\|U\|^2\langle U,H\rangle\langle U^{-1},H\rangle\\
&\quad\left.+\frac{5}{9}\|U\|^4\langle U^{-1},H\rangle^2-\frac{2}{3}\|U\|^4\langle \operatorname{Adj}H,\operatorname{Adj}U^{-1}\rangle\right)\\
&=6\frac{\|U\|^2}{\det[U]^2}\left(4\langle U,H\rangle^2-\frac{4\cdot 2}{3}\|U\|^2\langle U,H\rangle\langle U^{-1},H\rangle+\frac{4}{9}\|U\|^4\langle U^{-1},H\rangle^2+\right.\\
&\quad\left.\|U\|^2\|H\|^2-\frac{4}{3}\|U\|^2\langle U,H\rangle\langle U^{-1},H\rangle+\frac{5}{9}\|U\|^4\langle U^{-1},H\rangle^2-\frac{2}{3}\|U\|^4\langle \operatorname{Adj}H,\operatorname{Adj}U^{-1}\rangle\right)\\
&=6\frac{\|U\|^2}{\det[U]^2}\left(\|U\|^2\|H\|^2+4\langle U,H\rangle^2-4\|U\|^2\langle U,H\rangle\langle U^{-1},H\rangle+\|U\|^4\langle U^{-1},H\rangle^2-\right.\\
&\quad\left.\frac{2}{3}\|U\|^4\langle \operatorname{Adj}H,\operatorname{Adj}U^{-1}\rangle\right)
\end{aligned}
$$

Observe that the admissible choice $H=p^+\cdot U$ yields $p^+\cdot D^2_U W_2(U).(U,U)=0$. Now take $H=U+X$ with $X\in\mathrm{PSym}$. For the inequality we need only consider the bracket. After some algebraic manipulations we get for the bracket

$$
\begin{aligned}
&0-2\|U\|^2\left[\langle U,X\rangle-\frac{1}{3}\|U\|^2\langle U^{-1},X\rangle\right]+\\
&\left(\|U\|^2\|X\|^2+4\langle U,X\rangle^2-4\|U\|^2\langle U,X\rangle\langle U^{-1},X\rangle+\|U\|^4\langle U^{-1},X\rangle^2-\right.\\
&\quad\left.\frac{2}{3}\|U\|^4\langle \operatorname{Adj}X,\operatorname{Adj}U^{-1}\rangle\right)\ .
\end{aligned}
$$

Now assume $X=t\eta\otimes\eta$ where η is some unit eigenvector of U and $t\in\mathbb{R}^+$. Thus $H\in\mathrm{PSym}$ and we infer for the bracket

$$
\begin{aligned}
&=-2t\|U\|^2\left[\langle U.\eta,\eta\rangle-\frac{1}{3}\|U\|^2\langle U^{-1}.\eta,\eta\rangle\right]+\\
&t^2\left(\|U\|^2+4\langle U.\eta,\eta\rangle^2-4\|U\|^2\langle U.\eta,\eta\rangle\langle U^{-1}.\eta,\eta\rangle+\|U\|^4\langle U^{-1}.\eta,\eta\rangle^2\right)\\
&=-2t\|U\|^2\left[\lambda_1-\frac{1}{3}(\lambda_1^2+\lambda_2^2+\lambda_3^2)\frac{1}{\lambda_1}\right]+\\
&t^2\left(\|U\|^2+4\langle U.\eta,\eta\rangle^2-4\|U\|^2\langle U.\eta,\eta\rangle\langle U^{-1}.\eta,\eta\rangle+\|U\|^4\langle U^{-1}.\eta,\eta\rangle^2\right)\\
&=-\frac{2}{3\lambda_1}t\|U\|^2\left[3\lambda_1^2-(\lambda_1^2+\lambda_2^2+\lambda_3^2)\right]+\\
&t^2\left(\|U\|^2+4\langle U.\eta,\eta\rangle^2-4\|U\|^2\langle U.\eta,\eta\rangle\langle U^{-1}.\eta,\eta\rangle+\|U\|^4\langle U^{-1}.\eta,\eta\rangle^2\right).
\end{aligned}
$$

If U is such that $\lambda_1 > \max\{\lambda_2, \lambda_3\}$ then for $t > 0$ small enough the expression will be negative. Remark that if $U = p^+ \cdot 1\!\!1$, i.e. a pure pressure state we have $D^2_U W_2(U).(H,H) = p^2 \|\operatorname{dev} H\|^2 \geq 0$. Equally, no such violation is possible if H, U are both diagonal (are coaxial). Finally, the calculation for W_3:

$$\begin{aligned}
D_U W_3(F).H &= 2p\left(\ln\det[U]\right)^{2p-1}\langle U^{-1}, H\rangle \\
D^2_U W_3(F).(H,H) &= 2p(2p-1)\left(\ln\det[U]\right)^{2p-2}\langle U^{-1}, H\rangle^2 - -2p\left(\ln\det[U]\right)^{2p-1}\langle U^{-1}HU^{-1}, H\rangle \\
&= 2p\left(\ln\det[U]\right)^{2p-2}\Big[(2p-1)\langle U^{-1}, H\rangle^2 - \ln\det[U]\langle U^{-1}HU^{-1}, H\rangle\Big] .
\end{aligned}$$

Choose $U = m \cdot 1\!\!1$. Hence

$$D^2_U W_3(m \cdot 1\!\!1).(H,H) = 2p\left(3\ln m\right)^{2p-2}\frac{1}{m^2}\left[(2p-1)\langle 1\!\!1, H\rangle^2 - 3\ln m\|H\|^2\right] .$$

If m is large enough, this expression will be negative. ■