

# Neue Matrix-Ungleichungen und Anwendungen auf konstitutive Beziehungen in der nichtlinearen Elastizitätstheorie

DIPL.-INF. CHRISTIAN THIEL  
geboren in Remscheid

Dissertation

zur Erlangung des akademischen Grades  
Doktor der Naturwissenschaften (Dr. rer. nat.)  
an der Fakultät für Mathematik  
der Universität Duisburg-Essen

14. September 2016

Tag der mündlichen Prüfung: 13.12.2016

**Erstgutachter:** Prof. Dr. rer. nat. habil. Patrizio Neff  
Fakultät für Mathematik  
Universität Duisburg-Essen

**Zweitgutachter:** Prof. Dr. rer. nat. habil. Krzysztof Chelmiński  
Fakultät für Mathematik und Informatik  
Universität Warschau

**Drittgutachter:** Prof. Dr. rer. nat. habil. Mircea Bîrsan  
Fakultät für Mathematik  
Universität Duisburg-Essen

# Vorwort

Diese Arbeit ist das Ergebnis meiner vierjährigen wissenschaftlichen Tätigkeit auf dem Gebiet der nicht-linearen Elastizitätstheorie.

Für diese Zeit möchte ich Prof. Dr. Patrizio Neff danken. Er bot mir die Chance bei ihm zu promovieren und war als mein Chef, als Leiter unserer Arbeitsgruppe und zu guter Letzt als mein Betreuer für die vorliegende Dissertation stets von ihrem Gelingen überzeugt. Mit seinem Optimismus überbrückte er so manch aufkommenden Zweifel und motivierte mich zu weiterem Fortschritt. Ich danke ihm dafür, dass er mir die zur Fertigstellung nötigen Freiräume geschaffen hat und überdies als Erstgutachter meiner Arbeit auftrat. Den weiteren Gutachtern meiner Arbeit Prof. Krzysztof Chelmiński und Prof. Mircea Bîrsan möchte ich natürlich ebenfalls für ihre Mühen danken.

Besonderer Dank gebührt meinem Bürokollegen Dr. Robert Martin, der mir unaufhörlich ein geduldiger Ansprechpartner war und von dessen mathematischer Expertise ich ein ums andere Mal profitieren konnte. Ich bedanke mich für die vielen Stunden des Diskutierens und betone, dass seine mathematischen Ausführungen stets ein derart hohes Niveau hatten bzw. haben, dass ich ihn fast als mathematischen Mentor bezeichnen kann.

Weiterhin hatte ich das Glück, zu Beginn meiner Zeit in der Arbeitsgruppe auch mit M. Sc. Johannes Lankeit das Büro zu teilen. Inzwischen wissenschaftlicher Mitarbeiter an der Universität Paderborn, lernte ich Johannes zu Beginn des Mathematikstudiums als Kommilitonen kennen: Dabei war er mit seiner Hingabe an die Mathematik ein Vorbild, das aufgrund seiner schier grenzenlosen Auffassungsgabe unerreichbar schien. Aber recht bald erwies sich Johannes als sehr umgänglicher und überaus sympathischer Mensch, woraus eine Freundschaft erwuchs, die bis zum heutigen Tag anhält. Neben all der mathematischen Inspiration und seiner Unterstützung an dem vorliegenden Werk, möchte ich über eben jene Freundschaft meine Dankbarkeit ausdrücken.

Ein Projekt wie eine Dissertationsschrift ist natürlich kaum möglich ohne kompetente Korrekturleser, bei denen ich mich ebenfalls herzlich bedanken möchte:

Neben den oben schon genannten Robert Martin und Johannes Lankeit hatte ich das Glück, auf die geschulten Augen anderer Personen vertrauen zu können, denen kaum ein Schnitzer verborgen blieb. Diese gehörten M. Sc. Niklas Sapountzoglou, der mir schon als Korrekteur im Übungsbetrieb durch seine Gründlichkeit, Unermüdlichkeit und Fachkompetenz auffiel, B. Sc. Jendrik Voss und B. Sc. Eva Schweickert, deren mathematischen Scharfsinn ich sehr schätze, die ihre Zukunft in unserer Arbeitsgruppe noch vor sich haben und denen meine Arbeit hoffentlich als Grundlage für weitere Erkenntnisse dienen kann.

Überdies möchte ich B. Sc. Sonja Uebing danken: Einer gelernten Mathematikerin, die inzwischen ihren Master in der Mechanik macht und so bei der Korrektur neben der mathematischen Sicht die Arbeit auch aus dem Blickwinkel eines Ingenieurs betrachten konnte.

Der abschließende Dank gebührt einem weiteren Ingenieur: meinem Vater Horst Thiel. Ich möchte mich für das Redigieren meiner Arbeit und für sein Bereitstehen als Hilfe bei der einen oder anderen sprachlichen Feinheit bedanken. Ich freue mich sehr über sein Interesse an meiner Dissertation und damit einhergehend an der akademischen Behandlung eines Themengebietes, in dem er lange Jahre beruflich als Konstrukteur tätig und so maßgeblich an unserem technischen Fortschritt beteiligt war.

Essen, im Dezember 2016

# Inhaltsverzeichnis

<b>Vorwort</b>	<b>iii</b>
Einleitung . . . . .	1
Bisherige Publikationen . . . . .	2
<b>I. Mathematische Zusammenhänge</b>	<b>3</b>
<b>1. Grundlagen</b>	<b>4</b>
1.1. Matrizen und Skalarprodukte . . . . .	4
1.2. Differenzierbarkeit . . . . .	8
<b>2. Matrix-Analysis</b>	<b>12</b>
2.1. Mengen auf $\mathbb{R}^{n \times n}$ . . . . .	12
2.1.1. Operatoren auf $\mathbb{R}^{n \times n}$ . . . . .	14
2.2. Frobenius-Skalarprodukt und Frobeniusnorm . . . . .	18
2.3. Tangentialräume . . . . .	25
2.4. Polarzerlegung . . . . .	26
2.5. Isotropie und Koaxialität . . . . .	27
2.5.1. Isotrope Funktionen . . . . .	35
2.5.2. Semi-Invertierbarkeit, Baker-Ericksen und empirischen Ungleichungen . . . . .	38
2.6. Matrixexponential und Matrixlogarithmus . . . . .	43
2.6.1. Exponential . . . . .	43
2.6.2. Logarithmus . . . . .	46
2.7. Spurgleichungen . . . . .	47
2.8. Heron-Verfahren . . . . .	48
2.9. Monotonie und die Tschebyscheff-Summenungleichung . . . . .	48
2.9.1. Partialbruchzerlegung . . . . .	50
2.10. Differenzierbarkeit . . . . .	50
2.10.1. Differentiation von $L: \text{GL}^+(n) \rightarrow \mathbb{R}$ mit $L(QX) = L(X)$ . . . . .	56
2.11. Konvexität . . . . .	57
2.12. Majorization und Schur-Konvexität . . . . .	59
2.13. Nullstellen und Koeffizienten von Polynomen . . . . .	62
2.13.1. Ungleichungen zu elementarsymmetrischen Polynomen . . . . .	68
2.13.2. Eigenschaften von elementarsymmetrischen Polynomen . . . . .	69
<b>3. Sum-of-Squared-Logarithms-Ungleichung</b>	<b>71</b>
3.1. Einleitung . . . . .	71
3.1.1. Die Ungleichung . . . . .	71
3.1.2. Die SSLI als Schlüssel zur Lösung eines Minimierungsproblems . . . . .	72
3.1.3. Chronologie der Lösung . . . . .	73
3.1.4. Ein elementarer Beweis auf Basis der Konvexität für $n = 2$ . . . . .	74
3.1.5. Gescheiterter Beweisversuch auf Basis von Majorization . . . . .	75
3.1.6. Der Beweis für beliebige Dimension . . . . .	75
3.2. Verbindungen der SSLI zur Entropie . . . . .	82
3.3. Andere Bedingungen für die SSLI und die Sum-of-Powered-Logarithms-Ungleichung . . . . .	84
3.4. Verwandte Ungleichungen . . . . .	86
3.5. Neue logarithmische Ungleichungen in der Informationstheorie . . . . .	91

<b>II. Anwendung auf die nichtlineare Elastizitätstheorie</b>	<b>94</b>
<b>4. Einführung in die nichtlineare Elastizitätstheorie</b>	<b>95</b>
4.1. Modellierung über das Kraftgleichgewicht nach Cauchy . . . . .	95
4.2. Modellierung über Energiefunktionen . . . . .	101
4.3. Nichtlineares Elastizitätsgesetz . . . . .	102
4.3.1. Lokalität der Deformation . . . . .	103
4.3.2. Homogenität der Deformation . . . . .	103
4.3.3. Polarzerlegung des Deformationsgradienten . . . . .	103
4.3.4. Konstitutives Gesetz . . . . .	104
4.3.5. Spannungstensoren . . . . .	105
4.3.6. Hyperelastizität . . . . .	106
4.3.7. Objektivität . . . . .	107
4.3.8. Isotropie . . . . .	110
4.3.9. Einige einfache Beispieldeformationen für inkompressible Materialien . . . . .	114
4.3.10. Verschiedene Energiefunktionen und ihr Einfluss auf uniaxiale, biaxiale und Pure-Shear-Deformation . . . . .	116
4.3.11. Darstellung isotroper Energiefunktionen als Funktion in den Invarianten . . . . .	120
4.3.12. Verzerrung . . . . .	121
4.3.13. Lineares Elastizitätsgesetz . . . . .	123
<b>5. Konvexität von Energiefunktionen in <math>\mathcal{C}</math></b>	<b>124</b>
5.1. Anwendungen der Konvexität einer Energiefunktion in $\mathcal{C}$ . . . . .	124
5.2. Die Konvexität der Valanis-Landel-Energien in $\mathcal{C}$ . . . . .	127
5.3. Beispiele für die Konvexität von $\widehat{W}(C) = \alpha \operatorname{tr}(C) + \beta \operatorname{tr}(C)^2 + \delta \operatorname{tr}(C^2) - \gamma \log \det C$ . . . . .	128
5.4. Die Konvexität der Saint-Venant-Kirchhoff-Energie in $\mathcal{C}$ . . . . .	129
5.5. Bestimmung der Approximation bis zur zweiten Ordnung von $\widehat{W}_1, \widehat{W}_2$ und $\widehat{W}_3$ . . . . .	130
<b>6. Semi-Invertierbarkeit und die empirischen Ungleichungen</b>	<b>134</b>
6.1. Semi-Invertierbarkeit von isotropen Elastizitätsgesetzen . . . . .	135
6.1.1. Semi-Invertierbarkeit bei $W(U) = \ \log U\ ^2$ . . . . .	138
6.1.2. Energiefunktionen vom Hencky-Typ . . . . .	141
6.1.3. Energien mit additiv isochor-volumetrischem Split . . . . .	146
6.1.4. Andere Energien . . . . .	150
<b>7. Shear</b>	<b>153</b>
7.1. Shear im linearen Modell – Simple-Shear-Deformationen und Pure-Shear-Spannungen . . . . .	157
7.2. Shear im nichtlinearen Modell – finite Simple-Shear-Deformationen und Pure-Shear-Spannungen . . . . .	159
7.2.1. Die allgemeine Form . . . . .	160
7.2.2. Finite-Simple-Shear-Deformationen . . . . .	165
7.3. Linearisierung von Finite-Simple-Shear . . . . .	169
7.4. Konstitutive Bedingungen damit Pure-Shear-Dehnungen auf Pure-Shear-Spannungen führen . . . . .	171
7.4.1. Pure-Shear-Cauchy-Spannungen . . . . .	172
7.4.2. Pure-Shear-Biot-Spannungen . . . . .	178
7.5. Polarzerlegung des Simple-Shear-Deformationsgradienten . . . . .	179
<b>8. Zusammenfassung und Ausblick</b>	<b>184</b>
8.1. Zusammenfassung der Haupt- und Nebenresultate . . . . .	184
8.2. Ausblick auf weiterführende Forschungsmöglichkeiten . . . . .	186
<b>9. Literaturverzeichnis</b>	<b>187</b>
<b>10. Stichwortverzeichnis</b>	<b>191</b>

## Einleitung

Nach dem Lexikon der Physik des Spektrum Verlags ist Elastizität die „Eigenschaft fester Körper, unter Zug-, Scher- oder Druckbeanspruchung ihre Form vorübergehend zu ändern und nach Wegfall der Belastung ihre ursprüngliche Gestalt wieder einzunehmen.“

Diese Definition beinhaltet auch noch visko-elastisches Verhalten. Dabei treten Hysterese-Effekte auf: Der Rückgang der Deformation nach Belastungsende nimmt Zeit in Anspruch. In dieser Arbeit beschränken wir uns auf ideal-elastisches Verhalten (ohne Hysterese).

Wir unterscheiden in

- linear-elastisches Verhalten: Es tritt vereinfacht bei kleinen Deformationen auf und kann durch den linearen Spannungs-Dehnungszusammenhang  $\varepsilon \mapsto \sigma = \mathbb{C} \cdot \varepsilon$  dargestellt werden (siehe auch Hookesches Gesetz).
- nichtlinear-elastisches Verhalten: Die Spannung hängt mittels  $F \mapsto \sigma(F)$  nichtlinear von der Deformation ab, ein typisches Beispiel ist die Gummielastizität bei größeren Deformationen.

Wir wenden uns in dieser Arbeit dem nichtlinear-elastischen Verhalten zu, welches im Gegensatz zum linear-elastischen Verhalten noch viele offene Fragestellungen beinhaltet. Dieser nichtlineare Fall ist mathematisch zwar wesentlich komplexer, aber mit Werkzeugen der Matrix-Analysis können wir ihn tiefgehend untersuchen.

Diese Arbeit untergliedert sich in zwei Teile. Der erste Teil befasst sich mit mathematischen Zusammenhängen. Dazu werden wir in Kapitel 1 zunächst grundlegende Begriffe und Notationen einführen. Darauf aufbauend widmen wir uns im zweiten Kapitel der Matrix-Analysis. Wichtige Ergebnisse sind hierbei besonders der Zusammenhang zwischen Koaxialität und Isotropie und das Konzept der Semi-Invertierbarkeit von isotropen Funktionen verbunden mit den Baker-Ericksen-Ungleichungen und den empirischen Ungleichungen. Wir führen den Begriff der Konvexität von Tensorfunktionen und der Majorization ein und stellen den Zusammenhang zwischen Nullstellen und Koeffizienten von Polynomen heraus. Letzteren verwenden wir in Kapitel 2, welches die Sum-of-Squared-Logarithms-Ungleichung behandelt. Einen wesentlichen Anteil des Kapitels macht hierbei der Beweis der bis dato nur für  $n \in \{2, 3, 4\}$  gezeigten Ungleichung aus. Zudem ordnen wir unseren Beweis in die Chronologie der bisherigen Beweisfindung ein. Weiterhin untersuchen wir Modifikationen im Hinblick auf Bedingungen der SSLI und stellen die Sum-of-Powered-Logarithms-Ungleichung sowie andere verwandte Ungleichungen vor.

Der zweite Teil der Arbeit beschäftigt sich mit der Anwendung der im ersten Teil herausgestellten mathematischen Zusammenhänge auf die nichtlineare Elastizitätstheorie. Zunächst geben wir eine umfassende Einführung in die Modellierung elastischer Körper durch Elastizitätsgesetze mittels des Kraftgleichgewichtes nach Cauchy und über den Einsatz von Energiefunktionen. In diesem vierten Kapitel verwenden wir unsere Vorarbeit aus Kapitel 2 und können dadurch die Eigenschaften nichtlinearer Elastizitätsgesetze tiefgehend mathematisch untersuchen. Darunter fallen zum Beispiel Hyperelastizität, Objektivität und Isotropie. Darüber hinaus leiten wir die verschiedenen Spannungstensoren – Cauchy-Spannungstensor, erster und zweiter Piola-Kirchhoff-Spannungstensor und Biot-Spannungstensor – her. Wir vergleichen einige der in dieser Arbeit verwendeten Energiefunktionen in ihrer Wirkungsweise auf einfache Beispieldeformationen und stellen diese realen Messdaten gegenüber.

In Kapitel 5 stellen wir eine hinreichende Bedingung vor, unter der eine schwache Lösung der Euler-Lagrange-Gleichungen in der nichtlinearen Elastizität bereits ein globaler Minimierer des zugehörigen Energiefunktionalen ist. Dieses Kriterium lässt sich auf Energien  $W(F) = \widehat{W}(F^T F) = \widehat{W}(C)$ , welche konvex im rechten Cauchy-Green-Tensor  $C = F^T F$  sind, anwenden. Hierbei bezeichnet  $F$  den Deformationsgradienten. Wir geben Beispiele solcher Energien an, die einen Blowup für  $\det F \rightarrow 0$  verursachen.

In Kapitel 6 untersuchen wir die Semi-Invertierbarkeit der Darstellung  $\sigma(B) = \beta_0 \mathbb{1} + \beta_1 B + \beta_{-1} B^{-1}$  der Cauchy-Spannung im Fingertensor  $B = FF^T$ . Wir stellen eine konkrete Formel für die Berechnung der Koeffizientenfunktionen  $\beta_{-1}$ ,  $\beta_0$ ,  $\beta_1$  vor. Wir betrachten Truesdells empirische Ungleichungen  $\beta_{-1} \leq 0$ ,  $\beta_0 \leq 0$ ,  $\beta_1 > 0$  und zeigen, dass die Bedingung an  $\beta_0$  im Hinblick auf die Semi-Invertierbarkeit keine Relevanz hat. Mehr noch, dass relevante semi-invertierbare Elastizitätsgesetze ausschließlich die  $\beta_{-1}$ -Bedingung und  $\beta_1$ -Bedingung erfüllen können. Wir führen eine darauf angepasste schwache Variante der empirischen

Ungleichungen ein und zeigen, dass auch diese den Schluss auf die Baker-Ericksen-Ungleichungen garantieren. Weiterhin nutzen wir in diesem Kapitel unsere Ergebnisse über die SSLI aus dem vorigen Kapitel und beweisen, dass das durch  $W(U) = \|\log U\|^2$  induzierte Elastizitätsgesetz die schwachen empirischen Ungleichungen erfüllt und somit semi-invertierbar ist. Mit diesem Resultat führen wir die Klasse der Energiefunktionen vom Hencky-Typ ein und werden sehen, dass deren typische Vertreter wie die quadratische Hencky-Energie und die Exponentiated-Hencky-Energie ebenfalls die schwachen empirischen Ungleichungen erfüllen. Wir zeigen, dass Semi-Invertierbarkeit nicht Invertierbarkeit impliziert.

In Kapitel 7 zeigen wir, dass unter Voraussetzung des linearen Modells jede Simple-Shear-Deformation von jedem linearen Elastizitätsgesetz auf eine Pure-Shear-Spannung abgebildet wird, aber im finiten Modell jede (nichttriviale) Simple-Shear-Deformation mittels eines isotropen Elastizitätsgesetzes weder auf eine Pure-Shear-Cauchy-Spannung, noch auf eine Pure-Shear-Biot-Spannung führt. Wir präsentieren mit der linken und der rechten Finite-Simple-Shear-Deformation zwei dem Simple-Shear recht ähnliche Deformationen und zeigen, dass die linke Finite-Simple-Shear-Deformation einer Pure-Shear-Cauchy-Spannung, die rechte Finite-Simple-Shear-Deformation einer Pure-Shear-Biot-Spannung entspricht. Zudem stellen wir heraus, dass die Linearisierungen beider Finite-Simple-Shear-Deformationen die klassische Simple-Shear-Deformation ergeben. Weiterhin untersuchen wir in diesem Kapitel Energiefunktionen, die auf Elastizitätsgesetze führen, mit denen jede Pure-Shear-Dehnung auf eine Pure-Shear-Cauchy-Spannung bzw. jede Pure-Shear-Dehnung auf eine Pure-Shear-Biot-Spannung abgebildet wird. Abschließend führen wir noch die rechte und die linke Polarzerlegung des klassischen Simple-Shear-Deformationsgradienten durch.

## Bisherige Publikationen

Größere Teile dieser Arbeit wurden bisher in wissenschaftlichen Zeitschriften veröffentlicht oder sind zur Veröffentlichung akzeptiert. Dies betrifft:

- F. Dannan, P. Neff und C. Thiel. „On the sum of squared logarithms inequality and related inequalities“. *Journal of Mathematical Inequalities* (2015). open access, available at arXiv:1411.1290.
- L. Borisov, P. Neff, S. Sra und C. Thiel. „The sum of squared logarithms inequality in arbitrary dimensions“. to appear in *Linear Algebra and its Applications* (2016). preprint available at arXiv:1508.04039. doi: 10.1016/j.laa.2016.06.026.
- D. Y. Gao, P. Neff, I. Roventa und C. Thiel. „On the convexity of nonlinear elastic energies in the right Cauchy-Green tensor“. to appear in *Journal of Elasticity* (2016). preprint available at arXiv:1508.05721.

Weiterhin sind noch Publikationen in Arbeit, welche die noch nicht eingereichten neu untersuchten Teile dieser Arbeit betreffen:

- C. Thiel, R. J. Martin, P. Neff. „Generalizations of the empirical inequalities“
- C. Thiel, R. J. Martin, P. Neff. „Shear – Pure and Simple“

**Teil I.**

# **Mathematische Zusammenhänge**



# 1. Grundlagen

In diesem Kapitel führen wir zunächst grundlegende mathematische Begriffe und Aussagen ein. Dies soll dazu dienen, dem Leser die Notation der folgenden Kapitel vorzustellen.

## 1.1. Matrizen und Skalarprodukte

**Definitionen 1.1.1.** Sei  $A = (a_{ij})_{i,j \in \{1, \dots, n\}} \in \mathbb{K}^{n \times n}$  eine quadratische  $n \times n$  Matrix über einem Körper  $K \in \{\mathbb{C}, \mathbb{R}\}$  mit

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}. \quad (1.1.1)$$

Wir bezeichnen die Summe der Hauptdiagonalelemente von  $A$  als die *Spur* von  $A$ , geschrieben

$$\operatorname{tr}(A) := \sum_{i=1}^n a_{ii}. \quad (1.1.2)$$

Die Matrix  $A$  heißt *spurfrei*, wenn  $\operatorname{tr}(A) = 0$  gilt.

Die *Determinante* von  $A$  definieren wir über die sogenannte Leibniz-Formel

$$\det A := \sum_{\sigma \in S_n} \left( \operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i, \sigma(i)} \right) \quad (1.1.3)$$

und wird als Summe über alle Permutationen  $\sigma$  der symmetrischen Gruppe  $S_n$  berechnet. Hierbei bezeichnen wir mit  $\operatorname{sgn}(\sigma)$  das Signum der Permutation  $\sigma$ .<sup>1</sup>

Die zu  $A$  *transponierte Matrix*, geschrieben  $A^T \in \mathbb{K}^{n \times n}$ , definieren wir als

$$A^T = \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix}. \quad (1.1.4)$$

So ergibt sich die transponierte Matrix aus  $A$  durch Spiegelung ihrer Elemente an der Hauptdiagonalen  $a_{11}, a_{22}, \dots, a_{nn}$ .

Wir bezeichnen  $A$  als *symmetrisch*, wenn  $A^T = A$ , als *chiefsymmetrisch* wenn  $A^T = -A$  gilt.

So können wir  $A$  in einen *symmetrischen Anteil*  $\operatorname{sym}(X)$  und einen *chiefsymmetrischen Anteil*  $\operatorname{skew}(X)$  aufteilen mittels

$$\operatorname{sym}(A) := \frac{1}{2}(A + A^T) \quad \text{und} \quad \operatorname{skew}(A) := \frac{1}{2}(A - A^T). \quad (1.1.5)$$

Den *deviatorischen Anteil*  $\operatorname{dev}_n(A)$  (spurfreie Anteil) von  $A$  definieren wir als

$$\operatorname{dev}_n(A) := A - \frac{1}{n} \operatorname{tr}(X) \cdot \mathbf{1}. \quad (1.1.6)$$

---

<sup>1</sup>Das Signum einer Permutation  $\sigma$  ist 1, falls sie gerade ist und  $-1$ , falls sie ungerade ist

## 1. Grundlagen

Hierbei bezeichnen wir mit  $\mathbb{1}$  die *Einheitsmatrix*.

Die Einheitsmatrix ist eine Diagonalmatrix mit Einsen auf der Hauptdiagonalen, allgemein definieren wir eine *Diagonalmatrix* mit Hauptdiagonalelementen  $\alpha_1, \dots, \alpha_n$  mittels

$$\text{diag}(\alpha_1, \dots, \alpha_n) := (c_{ij})_{i,j \in \{1, \dots, n\}} \quad \text{mit } c_{ii} = \alpha_i \text{ und } c_{ij} = 0 \text{ für } i \neq j. \quad (1.1.7)$$

**Proposition 1.1.2.** Die Determinante  $\det: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  ist

- i) multilinear in ihren Zeilen bzw. Spalten,
- ii) alternierend in ihren Zeilen bzw. Spalten,
- iii) normiert, d.h.  $\det(\mathbb{1}) = 1$ .

Hierbei betrachten wir die Determinante als Abbildung  $\det: \overbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}^{n \text{ mal}} \rightarrow \mathbb{R}$  mit den Zeilen bzw. Spalten als Argument. Dann bedeuten die Bedingungen i) und ii):

- i) Für alle  $v_1, \dots, v_n, w \in \mathbb{R}^n$  und  $\lambda, \mu \in \mathbb{R}$  gilt

$$\det(v_1, \dots, \lambda v_i + \mu w, \dots, v_n) = \lambda \det(v_1, \dots, v_i, \dots, v_n) + \mu \det(v_1, \dots, w, \dots, v_n).$$

- ii) Für alle  $v_1, \dots, v_n, v \in \mathbb{R}^n$  gilt

$$\det(v_1, \dots, v, \dots, v, \dots, v_n) = 0,$$

d.h. sind zwei Zeilen oder zwei Spalten gleich, so ist die Determinante Null.

**Bemerkung 1.1.3.** Erfüllt eine Abbildung  $f: \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$  die Bedingungen i) und ii), so gilt  $f(X) = \det(X) \cdot f(\mathbb{1})$ . Somit ist  $\det$  schon eindeutig über i), ii) und iii) festgelegt.<sup>2</sup>

**Folgerung 1.1.4.** Für alle  $A, B \in \mathbb{R}^{n \times n}$  gilt  $\det(AB) = \det(A) \det(B)$ .

*Beweis.* Für beliebiges aber festes  $B \in \mathbb{R}^{n \times n}$  definieren wir die Abbildung  $f(X) = \det(XB)$ . Sei  $a_{\cdot i}$  die  $i$ -te Zeile von  $A$ , dann ist  $a_{\cdot i}B$  die  $i$ -te Zeile von  $AB$ , so hängt jede Zeile von  $AB$  auf gleiche Weise linear von der jeweiligen Zeile von  $A$  ab. Somit ist  $f$  multilinear und alternierend und wir können mittels der vorigen Bemerkung schließen

$$\det(AB) = f(A) = \det(A) \cdot f(\mathbb{1}) = \det(A) \det(B). \quad \blacksquare$$

**Lemma 1.1.5.** Seien  $A, B \in \mathbb{R}^{n \times n}$ . Dann gilt  $\text{tr}(AB) = \text{tr}(BA)$ .

*Beweis.* Die Multiplikation  $C = AB$  ist definiert über  $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ , also  $c_{ii} = \sum_{k=1}^n a_{ik} b_{ki}$ . Wir erhalten

$$\text{tr}(AB) = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki} = \sum_{k=1}^n \sum_{i=1}^n b_{ki} a_{ki} = \text{tr}(BA). \quad \blacksquare$$

**Folgerung 1.1.6.** Die Spur ist invariant gegenüber zyklischen Vertauschungen, d.h. für  $A_1, \dots, A_r \in \mathbb{R}^{n \times n}$  gilt  $\text{tr}(A_1 \cdot A_2 \cdot \dots \cdot A_r) = \text{tr}(A_r \cdot A_1 \cdot A_2 \cdot \dots \cdot A_{r-1}) = \text{tr}(A_{r-1} \cdot A_r \cdot A_1 \cdot \dots \cdot A_{r-2}) = \dots = \text{tr}(A_2 \cdot \dots \cdot A_r \cdot A_1)$ .

**Definition 1.1.7.** Sei  $\mathcal{U} \subseteq \mathbb{R}^{n \times n}$  ein Untervektorraum der reellen  $n \times n$ -Matrizen. Eine Abbildung  $\Phi: \mathcal{U} \rightarrow \mathbb{R}^{n \times n}$  nennt sich *linear*, wenn für alle  $A, B \in \mathbb{R}^{n \times n}$  und  $\alpha \in \mathbb{R}$  gilt

$$\Phi(A + B) = \Phi(A) + \Phi(B) \quad \text{und} \quad \Phi(\alpha A) = \alpha \Phi(A). \quad (1.1.8)$$

**Bemerkung 1.1.8.** Die beiden Eigenschaften der Linearität lassen sich äquivalent vereinen zu:

$$\Phi(\alpha A + \beta B) = \alpha \Phi(A) + \beta \Phi(B) \quad \text{für alle } \alpha, \beta \in \mathbb{R} \text{ und alle } A, B \in \mathcal{U}. \quad (1.1.9)$$

---

<sup>2</sup>Näheres hierzu in Bosch [13, S. 140].

## 1. Grundlagen

**Lemma 1.1.9.** Die Spurabbildung  $\text{tr}$  ist linear.

*Beweis.* Für alle  $X, Y \in \mathbb{R}^{n \times n}$  und  $\alpha, \beta \in \mathbb{R}$  gilt

$$\text{tr}(\alpha X + \beta Y) = \sum_{i=1}^n (\alpha X + \beta Y)_{ii} = \alpha \sum_{i=1}^n x_{ii} + \beta \sum_{i=1}^n y_{ii} = \alpha \text{tr}(X) + \beta \text{tr}(Y). \quad \blacksquare$$

**Proposition 1.1.10.** Sei  $\mathcal{U} \subseteq \mathbb{R}^{n \times n}$  ein Untervektorraum der  $n \times n$ -Matrizen. Dann sind die Abbildungen  $\text{sym}, \text{skew}, \text{dev}_n: \mathcal{U} \rightarrow \mathbb{R}^{n \times n}$  linear.

*Beweis.* Für alle  $X, Y \in \mathcal{U}$  und  $\alpha, \beta \in \mathbb{R}$  gilt

$$\begin{aligned} \text{sym}(\alpha X + \beta Y) &= \frac{1}{2}((\alpha X + \beta Y) + (\alpha X + \beta Y)^T) \\ &= \frac{1}{2}(\alpha X + \beta Y + \alpha X^T + \beta Y^T) = \alpha \text{sym}(X) + \beta \text{sym}(Y), \end{aligned}$$

$$\begin{aligned} \text{skew}(\alpha X + \beta Y) &= \frac{1}{2}((\alpha X + \beta Y) - (\alpha X + \beta Y)^T) \\ &= \frac{1}{2}(\alpha X + \beta Y - \alpha X^T - \beta Y^T) = \alpha \text{skew}(X) + \beta \text{skew}(Y), \end{aligned}$$

$$\begin{aligned} \text{dev}_n(\alpha X + \beta Y) &= (\alpha X + \beta Y) - \frac{1}{n} \text{tr}(\alpha X + \beta Y) \cdot \mathbb{1} \\ &= \alpha X + \beta Y - \frac{1}{n}(\alpha \text{tr}(X) \cdot \mathbb{1} + \beta \text{tr}(Y) \cdot \mathbb{1}) = \alpha \text{dev}_n(X) + \beta \text{dev}_n(Y). \quad \blacksquare \end{aligned}$$

**Proposition 1.1.11.** Sei  $\mathfrak{M} \subseteq \mathbb{R}^{n \times n}$  eine isotrope Menge. Dann sind die Abbildungen  $\text{sym}, \text{skew}, \text{dev}_n: \mathfrak{M} \rightarrow \mathbb{R}^{n \times n}$  isotrop.<sup>3</sup>

*Beweis.* Für alle  $X \in \mathbb{R}^{n \times n}$  und  $Q \in O(n)$  gilt

$$\begin{aligned} \text{sym}(Q^T X Q) &= \frac{1}{2}((Q^T X Q) + (Q^T X Q)^T) \\ &= \frac{1}{2}(Q^T X Q + Q^T X^T Q) = Q^T \text{sym}(X) Q, \end{aligned}$$

$$\begin{aligned} \text{skew}(Q^T X Q) &= \frac{1}{2}((Q^T X Q) - (Q^T X Q)^T) \\ &= \frac{1}{2}(Q^T X Q - Q^T X^T Q) = Q^T \text{skew}(X) Q, \end{aligned}$$

$$\begin{aligned} \text{dev}_n(Q^T X Q) &= (Q^T X Q) - \frac{1}{n} \text{tr}(Q^T X Q) \cdot \mathbb{1} \\ &= Q^T X Q - \frac{1}{n} \text{tr}(Q Q^T X) \cdot Q^T \mathbb{1} Q = Q^T \text{dev}_n(X) Q. \quad \blacksquare \end{aligned}$$

**Definition 1.1.12.** Sei  $V$  ein komplexer Vektorraum. Wir nennen die Abbildung  $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$  ein Skalarprodukt, wenn für alle  $u, v, w \in V$  und alle  $\alpha \in \mathbb{R}$  die folgenden vier Bedingungen gelten.  $\langle \cdot, \cdot \rangle$  ist ...

- i) *linear im ersten Argument:*  $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$  und  $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$ .
- ii) *semilinear im zweiten Argument:*  $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$  und  $\langle u, \alpha v \rangle = \bar{\alpha} \langle u, v \rangle$ .
- iii) *hermitesch:*  $\langle u, v \rangle = \overline{\langle v, u \rangle}$ .
- iv) *positiv definit:*  $\langle u, u \rangle \geq 0$  mit  $\langle u, u \rangle = 0$  genau dann, wenn  $u = 0$ .

Hierbei bezeichnet  $\bar{x}$  für jedes  $x \in \mathbb{C}$  die komplex Konjugierte von  $x$ .

Gelten i) und ii), so sagen wir auch,  $\langle \cdot, \cdot \rangle$  ist *sesquilinear*.

<sup>3</sup>siehe auch Definition 2.5.1

## 1. Grundlagen

**Definition 1.1.13.** Sei  $V$  ein reeller Vektorraum. Wir nennen die Abbildung  $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$  ein *Skalarprodukt*, wenn für alle  $u, v, w \in V$  und alle  $\alpha \in \mathbb{R}$  die folgenden drei Bedingungen gelten.  $\langle \cdot, \cdot \rangle$  ist ...

- i) *linear im ersten Argument*:  $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$  und  $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$ .
- ii) *symmetrisch*:  $\langle u, v \rangle = \langle v, u \rangle$ .
- iii) *positiv definit*:  $\langle u, u \rangle \geq 0$  mit  $\langle u, u \rangle = 0$  genau dann, wenn  $u = 0$ .

Aufgrund der Symmetrie ist die Linearität im ersten Argument äquivalent zur Linearität in beiden Argumenten.

**Definition 1.1.14.** Sei  $V$  ein reeller oder komplexer Vektorraum mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ . Wir nennen  $u, v \in V$  *orthogonal*, geschrieben  $u \perp v$ , wenn  $\langle u, v \rangle = 0$  gilt.

Für  $V = \mathbb{R}^n$  können wir als Summe über die komponentenweise Multiplikation ein Skalarprodukt, das *Standardskalarprodukt*, ausdrücken:

**Definition 1.1.15.** Das *Standardskalarprodukt*  $\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  ist definiert als

$$\langle u, v \rangle := \sum_{i=1}^n u_i v_i.$$

Hierbei bezeichnen  $u_1, \dots, u_n \in \mathbb{R}$  die Komponenten von  $u$  und  $v_1, \dots, v_n \in \mathbb{R}$  die Komponenten von  $v$ .

**Bemerkung 1.1.16.** Wir verwenden für das Standardskalarprodukt dieselbe Notation wie für das allgemeine Skalarprodukt in Definition 1.1.13. Immer, wenn wir in dieser Arbeit ohne zusätzliche Bezeichnung vom Skalarprodukt sprechen, meinen wir das Standardskalarprodukt.

**Definition 1.1.17.** Sei  $V$  ein reeller Vektorraum. Wir nennen die Abbildung  $\|\cdot\|: V \rightarrow \mathbb{R}$  eine *Norm*, wenn für alle  $u, v \in V$  und  $\alpha \in \mathbb{R}$  die folgenden drei Bedingungen gelten.

- i) *Definitheit*: Aus  $\|u\| = 0$  folgt  $u = 0$ .
- ii) *absolute Homogenität*: Es gilt  $\|\alpha \cdot u\| = |\alpha| \cdot \|u\|$ .
- iii) *Subadditivität*: Es gilt die Dreiecksungleichung  $\|u + v\| \leq \|u\| + \|v\|$ .

**Bemerkung 1.1.18.** Aus ii) folgt mit  $\alpha = 0$  und i) die Äquivalenz  $\|u\| = 0 \Leftrightarrow u = 0$ . Setzen wir  $v = -u$  in iii) und nutzen zudem i), so erhalten wir die Nichtnegativität der Norm  $\|u\| \geq 0$ .

**Lemma 1.1.19.** Jedes Skalarprodukt auf einem reellen Vektorraum  $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$  induziert eine Norm  $\|\cdot\|: V \rightarrow \mathbb{R}$ , die sogenannte *Skalarproduktnorm* durch  $\|u\| := \sqrt{\langle u, u \rangle}$ .

**Definition 1.1.20.** Das Standardskalarprodukt induziert die *euklidische Norm*  $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $\|u\| := \sqrt{u_1^2 + \dots + u_n^2}$ .

**Proposition 1.1.21.** Jede durch ein Skalarprodukt induzierte Norm auf den reellen  $n \times n$ -Matrizen erfüllt die *Cauchy-Schwarz-Ungleichung*

$$|\langle X, Y \rangle| \leq \|X\| \cdot \|Y\| \quad \text{für alle } X, Y \in \mathbb{R}^{n \times n}. \quad (1.1.10)$$

*Beweis.* Für  $X = 0$  oder  $Y = 0$  ist die Gleichung trivial. Sei also ohne Einschränkung  $X \neq 0$  und  $Y \neq 0$ . Dann gilt mit den Eigenschaften von Skalarprodukt und Norm

$$\begin{aligned} 0 &\leq \left\langle \frac{X}{\|X\|} \mp \frac{Y}{\|Y\|}, \frac{X}{\|X\|} \mp \frac{Y}{\|Y\|} \right\rangle \\ &= \left\langle \frac{X}{\|X\|}, \frac{X}{\|X\|} \right\rangle \mp 2 \left\langle \frac{Y}{\|Y\|}, \frac{Y}{\|Y\|} \right\rangle + \left\langle \frac{Y}{\|Y\|}, \frac{Y}{\|Y\|} \right\rangle \\ &= \frac{1}{\|X\|^2} \langle X, X \rangle \mp \frac{2}{\|X\| \cdot \|Y\|} \langle X, Y \rangle + \frac{1}{\|Y\|^2} \langle Y, Y \rangle \\ &= 2 \mp \frac{2}{\|X\| \cdot \|Y\|} \langle X, Y \rangle, \end{aligned}$$

## 1. Grundlagen

was äquivalent zu  $\pm\langle X, Y \rangle \leq \|X\| \cdot \|Y\|$  ist. Nun ist  $|\langle X, Y \rangle| = \langle X, Y \rangle$  oder  $|\langle X, Y \rangle| = -\langle X, Y \rangle$ . In beiden Fällen gilt die Aussage. ■

**Definition 1.1.22.** Seien  $A, B \in \mathbb{R}^{n \times n}$ . Wir sagen, dass  $A$  und  $B$  miteinander kommutieren, wenn  $AB = BA$  gilt.

### 1.2. Differenzierbarkeit

**Definition 1.2.1.** Seien  $V$  und  $W$  endlichdimensionale, normierte, reelle Vektorräume und  $\Omega \subseteq V$  eine offene Menge. Wir bezeichnen eine Funktion  $f: \Omega \rightarrow W$  als *differenzierbar an der Stelle*  $x \in \Omega$ , falls es eine lineare Abbildung  $T$  von  $\Omega$  nach  $W$ , und eine Funktion  $R: V \rightarrow W$  mit  $\lim_{h \rightarrow 0} \frac{R(h)}{\|h\|} = 0$  gibt, sodass

$$f(x+h) = f(x) + T.h + R(h) \quad \text{für alle } x \in \Omega \text{ und alle } h \in V \text{ mit } \|h\| \ll 1$$

gilt. Die lineare Abbildung  $T: \Omega \rightarrow W$  nennen wir dann die *Ableitung von  $f$  an der Stelle  $x$*  und bezeichnen sie mit  $Df[x]$ . Ist  $f$  an allen Stellen  $x \in \Omega$  differenzierbar, so bezeichnen wir  $f$  als *differenzierbar* und die Abbildung  $Df: \Omega \rightarrow L(V, W)$  mit  $x \mapsto Df[x]$  als ihre *Ableitungsfunktion*. Ist die Ableitungsfunktion zudem stetig, so schreiben wir  $f \in C^1(\Omega, W)$ . Lässt sich die Ableitungsfunktion wieder stetig differenzieren und sich dieser Prozess insgesamt  $k$ -mal iteriert durchführen, so bezeichnen wir  $f$  als  $k$ -mal stetig differenzierbar. Lässt er sich beliebig oft durchführen, so bezeichnen wir  $f$  als unendlich oft differenzierbar. Die Menge aller  $k$ -mal stetig differenzierbaren Funktionen bezeichnen wir mit  $C^k(\Omega, W)$ , die Menge aller unendlich differenzierbaren Funktionen mit  $C^\infty(\Omega, W)$ . Im Falle von  $W = \mathbb{R}$  schreiben wir auch kurz  $C^k(\Omega)$  für  $k \in \mathbb{N} \cup \{\infty\}$ .

Die Anwendung eines Vektors  $v \in V$  auf  $Df[x]$  nennen wir die *Richtungsableitung* von  $f$  in  $x$  in  $v$ -Richtung, die sich ebenso<sup>4</sup> per Differentialquotient bestimmen lässt, als

$$Df[x].v = \lim_{t \rightarrow 0} \frac{f(x+t \cdot v) - f(x)}{t}. \quad (1.2.1)$$

Ist  $V = \mathbb{R}^n$  mit der kanonischen Basis aus Einheitsnormalenvektoren  $e_1, \dots, e_n$ , so nennen wir die Richtungsableitungen in deren Richtung die *partiellen Ableitungen* von  $f$  an der Stelle  $x$ , geschrieben

$$\frac{\partial f}{\partial x_i}(x) = Df[x].e_i. \quad (1.2.2)$$

Ist  $V = \mathbb{R}^{n \times n}$ , so besteht die kanonische Basis aus  $e_i \otimes e_j$  für  $i, j \in \{1, \dots, n\}$  und wir erhalten die partiellen Ableitungen

$$\frac{\partial f}{\partial x_{ij}}(x) = Df[x].(e_i \otimes e_j). \quad (1.2.3)$$

Im Falle von  $V = \mathbb{R}^n$  und  $W = \mathbb{R}^m$  können wir die Ableitung von  $f: \Omega \rightarrow \mathbb{R}^m$  an der Stelle  $x \in \Omega \subset \mathbb{R}^n$

$$h \mapsto Df[x].h \quad (1.2.4)$$

auch als Matrizenmultiplikation von  $Df[x]$  mit  $h$  auffassen. Hierfür stellen wir  $Df[x]$  als  $m \times n$ -Matrix

$$Df[x] := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix} \quad (1.2.5)$$

dar, die wir als *Jacobi-Matrix* bezeichnen.

<sup>4</sup>wie kurze Umformung von  $f(x+t \cdot v) = f(x) + Df[x].(t.v) + R(t \cdot v)$  und Grenzübergang zeigt.

## 1. Grundlagen

Ist  $f: \Omega \rightarrow W$  zweimal stetig differenzierbar, so ist die partielle Ableitung  $\frac{\partial f}{\partial x_i}$  selbst wieder eine stetig differenzierbare Funktion  $\Omega \rightarrow W$  und für ihre Ableitung gilt  $\frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} = \frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i}$ .<sup>5</sup> Diesen Umstand nutzen wir für die folgende Definition:

**Definition 1.2.2.** Sei  $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$  ein Multiindex mit  $|\alpha| := \alpha_1 + \dots + \alpha_n$  und  $f$  eine  $|\alpha|$ -mal stetig differenzierbare Funktion, dann bezeichnen wir die partielle Ableitung nach  $\alpha$  mit

$$D^\alpha f := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}. \quad (1.2.6)$$

**Definition 1.2.3.** Für  $\Omega \subseteq \mathbb{R}^n$  bezeichnen wir den *Gradienten* von  $f: \Omega \rightarrow \mathbb{R}$  mit

$$\nabla f := \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}. \quad (1.2.7)$$

Für  $\Omega \subseteq \mathbb{R}^{n \times n}$  bezeichnen wir den *Gradienten* von  $f: \Omega \rightarrow \mathbb{R}$  mit

$$\nabla f := \begin{pmatrix} \frac{\partial f}{\partial x_{11}} & \dots & \frac{\partial f}{\partial x_{1n}} \\ \vdots & & \vdots \\ \frac{\partial f}{\partial x_{n1}} & \dots & \frac{\partial f}{\partial x_{nn}} \end{pmatrix}. \quad (1.2.8)$$

Sei  $V$  allgemein ein endlichdimensionaler, normierter, reeller Vektorraum, dann nennen wir den Operator  $\nabla: C^{k+1}(V, \mathbb{R}) \rightarrow C^k(V, V)$  *Nabla-Operator*.

**Bemerkung 1.2.4.** Der Gradient von  $f$  ist für jedes  $x \in \Omega \subseteq V$  eindeutig dadurch bestimmt, dass er

$$\langle \nabla f(x), h \rangle = Df[x].h \quad \text{für alle } h \in V \quad (1.2.9)$$

erfüllt.

**Definition 1.2.5.** Für  $\Omega \subseteq \mathbb{R}^n$  bezeichnen wir die *Divergenz* von  $f: \Omega \rightarrow \mathbb{R}^n$  mit

$$\operatorname{div} f := \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \dots + \frac{\partial f_n}{\partial x_n}. \quad (1.2.10)$$

Für  $\Omega \subseteq \mathbb{R}^n$  bezeichnen wir die *Divergenz* von  $f: \Omega \rightarrow \mathbb{R}^{n \times n}$  mit

$$\operatorname{div} f := \begin{pmatrix} \frac{\partial f_{11}}{\partial x_1} + \frac{\partial f_{12}}{\partial x_2} + \dots + \frac{\partial f_{1n}}{\partial x_n} \\ \vdots \\ \frac{\partial f_{n1}}{\partial x_1} + \frac{\partial f_{n2}}{\partial x_2} + \dots + \frac{\partial f_{nn}}{\partial x_n} \end{pmatrix}, \quad (1.2.11)$$

und äquivalent, indem wir  $f$  zeilenweise durch  $f_1, \dots, f_n$  ausdrücken,

$$\operatorname{div} f := \begin{pmatrix} \operatorname{div} f_1 \\ \vdots \\ \operatorname{div} f_n \end{pmatrix}. \quad (1.2.12)$$

Den Operator  $\operatorname{div}: C^{k+1}(\mathbb{R}^n, \mathbb{R}^n) \rightarrow C^k(\mathbb{R}^n, \mathbb{R})$  bzw.  $\operatorname{div}: C^{k+1}(\mathbb{R}^n, \mathbb{R}^{n \times n}) \rightarrow C^k(\mathbb{R}^n, \mathbb{R}^n)$  nennen wir den *Divergenz-Operator*.

---

<sup>5</sup>Die Vertauschbarkeit der Reihenfolge der partiellen Ableitungen erfolgt wegen der Stetigkeit der partiellen Ableitungen aus dem Satz von Schwarz

## 1. Grundlagen

**Definition 1.2.6.** Für  $\Omega \subset \mathbb{R}^n$  bezeichnen wir die *Divergenz des Gradienten* von  $f: \Omega \rightarrow \mathbb{R}$  mit

$$\Delta f := \operatorname{div}(\nabla f) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}. \quad (1.2.13)$$

Den Operator  $\Delta: C^{k+2}(\mathbb{R}^n, \mathbb{R}) \rightarrow C^k(\mathbb{R}^n, \mathbb{R})$  nennen wir den *Laplace-Operator*.

**Bemerkung 1.2.7.** Hat eine Funktion mehrere Argumente, so können wir auch partiell nach einzelnen mittels  $\partial$ ,  $\nabla$ ,  $\operatorname{div}$  oder  $\Delta$  ableiten, indem wir die anderen Argumente festhalten und die Funktion als Funktion nur in dem einen Argument nach Definition ableiten. So ist beispielsweise für  $f$  mit  $(a, b, c) \mapsto f(a, b, c)$  dann  $\frac{\partial f}{\partial a}$  (auch geschrieben als  $D_a$ ) oder auch  $\nabla_a f$ ,  $\operatorname{div}_a f$ ,  $\Delta_a f$  möglich.

**Bemerkung 1.2.8.** Für ein Vektorfeld  $f$  bezeichnen wir  $\operatorname{div} f$  auch als *Quellendichte* von  $f$ .

Ein Skalarfeld  $u$  bezeichnen wir auch als *Potential* und  $\Delta u$  als *Quellendichte* des Potentials  $u$ .

Ohne Beweis stellen wir fest:

**Proposition 1.2.9.** Die Ableitungsoperatoren  $D$ ,  $\operatorname{div}$ ,  $\Delta$  und  $\nabla$  sind linear.

**Proposition 1.2.10** (Produktformel). Seien  $V, W$  Vektorräume und  $\Omega \subseteq V$  eine offene Menge. Dazu seien differenzierbare Skalarfelder  $u, v: \Omega \rightarrow \mathbb{R}$  und ein differenzierbares Vektorfeld  $f: \Omega \rightarrow W$  gegeben. Dann gilt

$$\operatorname{div}(uf) = \langle \nabla u, f \rangle + u \operatorname{div} f.$$

Sind  $u$  und  $v$  sogar zweimal differenzierbar, so gilt

$$\Delta(uv) = u \Delta v + 2 \langle \nabla u, \nabla v \rangle + v \Delta u.$$

*Beweis.* Nachrechnen ergibt

$$\operatorname{div}(uf) = \sum_{i=1}^n \frac{\partial(uf_i)}{\partial x_i} = \sum_{i=1}^n \left( \frac{\partial u}{\partial x_i} f_i + \frac{\partial f_i}{\partial x_i} u \right) = \sum_{i=1}^n \frac{\partial u}{\partial x_i} f_i + u \sum_{i=1}^n \frac{\partial f_i}{\partial x_i} = \langle \nabla u, f \rangle + u \operatorname{div} f$$

und

$$\begin{aligned} \Delta(uv) &= \sum_{i=1}^n \frac{\partial^2(uv)}{\partial x_i^2} = \sum_{i=1}^n \left( \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} + \frac{\partial^2 v}{\partial x_i^2} u \right) \\ &= \sum_{i=1}^n \left( \frac{\partial^2 u}{\partial x_i^2} v + \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} + \frac{\partial^2 v}{\partial x_i^2} u + \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} \right) = u \Delta v + 2 \langle \nabla u, \nabla v \rangle + v \Delta u. \quad \blacksquare \end{aligned}$$

**Definition 1.2.11** (Äußeres Normalenfeld). Sei  $V$  ein Vektorraum,  $\Omega \subseteq V$  eine offene Menge und  $A \subseteq \Omega$  eine kompakte Teilmenge mit glattem Rand. Wir nennen das Vektorfeld  $n: \partial A \rightarrow V$  das *äußere Normalenfeld* von  $A$ , wenn für jedes  $x \in \partial A$  der Vektor  $n(x)$  orthogonal und nach außen gerichtet zur Tangentialebene in  $x$  an  $\partial A$  ist. Dann nennen wir  $n(x)$  *äußeren Normalenvektor* an  $\partial A$  in  $x$ .

**Definition 1.2.12.** Sei  $\Omega \subseteq \mathbb{R}^n$  ein Gebiet. Wir nennen eine Funktion  $\psi: \Omega \rightarrow \mathbb{R}^n$  eine *Testfunktion* auf  $\Omega$ , wenn sie unendlich oft differenzierbar ist und einen kompakten Träger<sup>6</sup> in  $\Omega$  besitzt, d.h. außerhalb einer kompakten Menge null ist.

Die Menge aller Testfunktionen bildet mit Addition und Skalarmultiplikation einen Vektorraum, den wir mit

$$C_0^\infty(\Omega) := \{ \psi \in C^\infty(\Omega) \mid \operatorname{supp}(\psi) \subset \Omega \text{ ist kompakt} \}$$

bezeichnen.

<sup>6</sup>Wir bezeichnen den Träger mit  $\operatorname{supp}(\psi) := \overline{\{x \in \Omega \mid \psi(x) \neq 0\}}$ .

## 1. Grundlagen

**Lemma 1.2.13** (Gaußscher Integralsatz). Sei  $\Omega \subseteq \mathbb{R}^n$  eine offene Menge und  $A \subseteq \Omega$  eine kompakte Teilmenge mit glattem Rand. Dazu sei  $f: \Omega \rightarrow V$  ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_{\partial A} \langle f(x), n(x) \rangle dS(x) = \int_A \operatorname{div} f(x) dx.$$

Hierbei ist  $n(x)$  der äußere Normalenvektor an  $\partial A$  in  $x$ .

*Beweis.* Siehe Heuser [38, Abschnitt 210]. ■

**Beispiel 1.2.14** (Wärmeleitungsgleichung). Sei  $\Omega \subseteq \mathbb{R}^3$  eine offene Menge und  $u: \Omega \times \mathbb{R} \rightarrow \mathbb{R}$  ein zeitabhängiges Skalarfeld. Es drückt  $u(x, t)$  die Temperatur am Ort  $x$  zum Zeitpunkt  $t$  aus. Wir fordern aufgrund von physikalischen Erwägungen:

$$\underbrace{\frac{\partial}{\partial t} \int_A u(x) dx}_{\text{zeitliche Ableitung der Gesamtmenge}} = - \int_{\partial A} \underbrace{\langle -\nabla_x u, n \rangle}_{\text{Ausfluss}} dS \quad \text{für alle } A \subseteq \Omega$$

Aufgrund von

$$\int_A u_t(x) dx = \frac{\partial}{\partial t} \int_A u(x) dx \quad \text{und} \quad - \int_{\partial A} \langle -\nabla_x u, n \rangle dS = \int_A \Delta_x u(x) dx$$

gilt damit

$$\int_A u_t(x) dx = \int_A \Delta_x u(x) dx \quad \text{für alle } A \subseteq \Omega.$$

Dies impliziert genau die sogenannte *Wärmeleitungsgleichung*

$$u_t = \Delta_x u. \tag{1.2.14}$$

Es gibt mehrere Formeln, die Greensche Formeln genannt werden. Zwei von ihnen, die wir in dieser Arbeit verwenden werden, stellen wir nun vor.

**Lemma 1.2.15** (Greensche Formel). Sei  $\Omega \subseteq \mathbb{R}^n$  eine offene Menge und  $A \subseteq \Omega$  eine kompakte Menge mit glattem Rand. Zudem sei  $f: \Omega \rightarrow \mathbb{R}^n$  eine stetig differenzierbare Funktion, dann gilt

$$\int_{\partial A} f_i(x) n_i(x) dS(x) = \int_A \frac{\partial f}{\partial x_i}(x) dx. \tag{1.2.15}$$

Seien  $u, v: \Omega \rightarrow \mathbb{R}$  zweimal stetig differenzierbare Skalarfelder, dann gilt

$$\int_A (f(x) \Delta g(x) - g(x) \Delta f(x)) dx = \int_{\partial A} \left( f(Dg[x] \cdot n(x)) - g(Df[x] \cdot n(x)) \right) dx. \tag{1.2.16}$$

*Beweis.* Die Formel (1.2.15) ergibt sich direkt aus dem Gaußschen Integralsatz angewendet auf  $g: \Omega \rightarrow \mathbb{R}^n$  mit  $g_i \equiv f_i$  und  $g_j \equiv 0$  für  $j \neq i$ .

Zum Beweis von Formel (1.2.16):

Sei  $F := f \nabla g - g \nabla f$ , dann ist

$$\operatorname{div} F = \operatorname{div}(f \nabla g - g \nabla f) = f \operatorname{div}(\nabla g) - g \operatorname{div}(\nabla f) = f \Delta g - g \Delta f.$$

Somit gilt für alle  $x \in \partial A$

$$\begin{aligned} \langle F(x), n(x) \rangle &= \langle f(x) \nabla g(x) - g(x) \cdot \nabla f(x), n(x) \rangle \\ &= f \langle \nabla g(x), n(x) \rangle + g \langle \nabla f(x), n(x) \rangle \\ &= f(x) \cdot Dg[x] \cdot n(x) + g(x) \cdot Df[x] \cdot n(x). \end{aligned}$$

Mit Lemma 1.2.13 folgt nun die Aussage. ■



## 2. Matrix-Analysis

In diesem Kapitel der Matrix-Analysis werden erste relevante Ergebnisse dieser Arbeit erzielt. Wir erarbeiten verschiedene mathematische Aussagen, welche die Basis für die folgenden Kapitel bilden. Dabei ist besonderes Augenmerk auf den Zusammenhang zwischen Koaxialität und Isotropie, sowie das Konzept der Semi-Invertierbarkeit von isotropen Funktionen verbunden mit den Baker-Ericksen-Ungleichungen und den empirischen Ungleichungen zu legen. Diese Vorarbeit benötigen wir in Kapitel 6, in welchem wir dies auf hyperelastische Elastizitätsgesetze anwenden. Außerdem führen wir den Begriff der Konvexität von Tensorfunktionen ein, die grundlegend für Kapitel 5 ist. Sowohl die Majorization als auch der Zusammenhang zwischen Nullstellen und Koeffizienten von Polynomen sind wichtige Bausteine für die Arbeit mit der Sum-of-Squared-Logarithms-Ungleichung, die wir in Kapitel 3 thematisieren.

### 2.1. Mengen auf $\mathbb{R}^{n \times n}$

**Definitionen 2.1.1.** Die Menge der reellen  $n \times n$ -Matrizen mit Determinante ungleich Null bilden mit der Matrizenmultiplikation eine Gruppe. Wir bezeichnen sie mit

$$\mathrm{GL}(n) := \{X \in \mathbb{R}^{n \times n} \mid \det X \neq 0\},$$

der *allgemeinen linearen Gruppe*, der Gruppe der regulären, also invertierbaren Matrizen.

Eine Untergruppe hiervon wird durch die Menge der Matrizen mit positiver Determinante

$$\mathrm{GL}^+(n) := \{X \in \mathbb{R}^{n \times n} \mid \det X > 0\}$$

und von  $\mathrm{GL}^+(n)$  durch die Menge der Matrizen mit Determinante gleich 1 gebildet:

$$\mathrm{SL}(n) := \{X \in \mathbb{R}^{n \times n} \mid \det X = 1\}.$$

Letztere nennen wir auch die *spezielle lineare Gruppe*.

Ebenfalls eine Untergruppe von  $\mathrm{GL}(n)$  wird durch die Menge aller orthogonalen Matrizen gebildet. Wir bezeichnen sie mit

$$\mathrm{O}(n) := \{X \in \mathbb{R}^{n \times n} \mid X^T X = \mathbb{1}\},$$

der Gruppe der *Drehspiegel-* und *Drehmatrizen*. Die Determinante ihrer Elemente ist  $-1$  für die Drehspiegelmatrizen und  $+1$  für die Drehmatrizen.

Eine Untergruppe sowohl von  $\mathrm{SL}(n)$  als auch von  $\mathrm{O}(n)$  wird durch die Menge aller Drehmatrizen gebildet. Wir bezeichnen sie mit

$$\mathrm{SO}(n) := \{X \in \mathbb{R}^{n \times n} \mid X^T X = \mathbb{1}, \det X = 1\}.$$

Die Menge der *symmetrischen*  $n \times n$ -Matrizen bezeichnen wir mit

$$\mathrm{Sym}(n) := \{X \in \mathbb{R}^{n \times n} \mid X^T = X\},$$

die Menge der *schief-symmetrischen*  $n \times n$ -Matrizen bezeichnen wir mit

$$\mathfrak{so}(n) := \{X \in \mathbb{R}^{n \times n} \mid X^T = -X\}.$$

Die Menge der *spurfreien*  $n \times n$ -Matrizen bezeichnen wir mit

$$\mathfrak{sl}(n) := \{X \in \mathbb{R}^{n \times n} \mid \mathrm{tr}(X) = 0\}.$$

## 2. Matrix-Analysis

$\text{Sym}(n)$ ,  $\mathfrak{so}(n)$  und  $\mathfrak{sl}(n)$  bilden zusammen mit der Matrizenmultiplikation jeweils keine Gruppe.<sup>1</sup> Aber sie sind alle Untervektorräume des  $\mathbb{R}^{n \times n}$ .

Die Teilmenge von  $\text{Sym}(n)$  der *positiv definiten symmetrischen Matrizen*<sup>2</sup>, die wir mit

$$\text{Sym}^+(n) := \{X \in \mathbb{R}^{n \times n} \mid X^T = X, \langle v, Xv \rangle > 0 \forall v \in \mathbb{R}^n\}$$

bezeichnen, ist allerdings kein Untervektorraum des  $\mathbb{R}^{n \times n}$ .

**Bemerkung 2.1.2.** Insgesamt gelten folgende Teilmengenbeziehungen:

$$\text{SO}(n) \subseteq \begin{matrix} \text{SL}(n) \subseteq \text{GL}^+(n) \\ \text{O}(n) \end{matrix} \subseteq \text{GL}(n) \subseteq \mathbb{R}^{n \times n}, \quad (2.1.1)$$

sowie

$$\text{Sym}^+(n) \subseteq \begin{matrix} \text{GL}^+(n) \subseteq \text{GL}(n) \\ \text{Sym}(n) \end{matrix} \subseteq \mathbb{R}^{n \times n} \quad \text{und} \quad \mathfrak{so}(n) \subseteq \mathfrak{sl}(n) \subseteq \mathbb{R}^{n \times n}. \quad (2.1.2)$$

**Proposition 2.1.3.** Die Menge  $\text{GL}(n)$  liegt dicht in  $\mathbb{R}^{n \times n}$ , d.h. für jedes  $A \in \mathbb{R}^{n \times n}$  existiert eine Folge  $(A_n)_{n \in \mathbb{N}}$  mit  $A_n \in \text{GL}(n)$ , die gegen  $A$  konvergiert.

*Beweis.* Sei  $n \in \mathbb{N}$ , dann ist  $\frac{1}{n}$  genau dann Eigenwert von  $A$ , wenn  $\det(A - \frac{1}{n} \cdot \mathbb{1}) = 0$  ist. Hat  $A$  einen kleinsten Eigenwert  $\tilde{\lambda}$  größer Null, wähle  $n_0 > \frac{1}{\tilde{\lambda}}$ , ansonsten wähle  $n_0 = 1$ . Dann gilt  $\det(A - \frac{1}{n} \cdot \mathbb{1}) \neq 0$  für alle  $n \geq n_0$ . Damit ist  $A_n$  invertierbar für alle  $n \geq n_0$ . Zudem ist  $\lim_{n \rightarrow \infty} \|A_n - A\| = \lim_{n \rightarrow \infty} \frac{1}{n} \|\mathbb{1}\| = 0$ . ■

**Folgerung 2.1.4.** Für  $V \in \{\mathbb{R}, \mathbb{R}^n, \mathbb{R}^{n \times n}\}$  seien  $f, g: \mathbb{R}^{n \times n} \times \dots \times \mathbb{R}^{n \times n} \rightarrow V$  stetige Funktionen. Gilt  $f(X) = g(X)$  für alle  $X \in \text{GL}(n) \times \dots \times \text{GL}(n)$ , so folgt  $f = g$ .

**Proposition 2.1.5.** Seien  $p_1, \dots, p_n$  die Spaltenvektoren von  $P \in \mathbb{R}^{n \times n}$ . Dann ist  $P \in \text{O}(n)$  genau dann, wenn

$$\langle p_i, p_j \rangle = \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j \end{cases}$$

für alle  $i, j \in \{1, \dots, n\}$ .

*Beweis.* Nach Definition der Matrizenmultiplikation gilt  $(P^T P)_{ij} = \sum_{k=1}^n p_{ki} p_{kj} = \langle p_i, p_j \rangle$ . ■

**Proposition 2.1.6.** Für alle  $P \in \mathbb{R}^{n \times n}$  ist  $P^T P = \mathbb{1}$  äquivalent zu  $PP^T = \mathbb{1}$ .

*Beweis.* Gilt  $PP^T = \mathbb{1}$  oder  $P^T P = \mathbb{1}$ , so ist  $\det P \cdot \det P^T = 1$ , also  $\det P \neq 0$ . Somit ist  $P$  invertierbar und Links- und Rechtsinverses sind identisch. ■

**Proposition 2.1.7.** Sei  $X \in \mathbb{R}^{n \times n}$ . Dann ist

- i)  $\text{sym } X \in \text{Sym}(n)$ ,
- ii)  $\text{skew } X \in \mathfrak{so}(n)$ ,
- iii)  $\text{dev}_n X \in \mathfrak{sl}(n)$ ,
- iv)  $\text{dev}_n \text{sym } X \in \text{Sym}(n) \cap \mathfrak{sl}(n)$ ,
- v)  $X = \text{sym } X + \text{skew } X = \text{dev}_n \text{sym } X + \text{skew } X + \frac{1}{n} \text{tr}(\text{sym } X) \cdot \mathbb{1}$ .

<sup>1</sup>Als Beispiele betrachten wir  $\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{\in \text{Sym}(2)} \cdot \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}}_{\in \text{Sym}(2)} = \underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}}_{\notin \text{Sym}(2)}$  und  $\underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{\in \mathfrak{so}(2)}^2 = \underbrace{\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}}_{\notin \mathfrak{sl}(2)}$  (mit  $\mathfrak{so}(2) \subseteq \mathfrak{sl}(2)$ ).

<sup>2</sup>So ist  $\mathbb{1} \in \text{Sym}^+(n)$ , aber  $-1 \cdot \mathbb{1} \notin \text{Sym}^+(n)$

## 2. Matrix-Analysis

*Beweis.* Für alle  $X \in \mathbb{R}^{n \times n}$  gilt

- i)  $(\text{sym } X)^T = \frac{1}{2}(X^T + X) = \text{sym } X,$
- ii)  $(\text{skew } X)^T = \frac{1}{2}(X^T - X) = -\frac{1}{2}(X - X^T) = -(\text{skew } X),$
- iii)  $\text{tr dev}_n(X) = \text{tr}(X - \frac{1}{n} \text{tr}(X) \cdot \mathbb{1}) = \text{tr } X - \frac{1}{n} \text{tr}(\text{tr}(X) \cdot \mathbb{1}) = \text{tr } X - \frac{1}{n} n \text{tr}(X) = 0,$
- iv)  $(\text{dev}_n \text{sym } X)^T = (\text{sym } X - \frac{1}{n} \text{tr}(X) \cdot \mathbb{1})^T = (\text{sym } X)^T - \frac{1}{n} \text{tr}(X) \cdot \mathbb{1}^T = \text{dev}_n \text{sym } X,$
- v) 
$$\begin{aligned} X &= X + \frac{1}{2}X^T - \frac{1}{2}X^T = \frac{1}{2}(X + X^T) + \frac{1}{2}(X - X^T) = \text{sym } X + \text{skew } X \\ &= \text{sym } X - \frac{1}{n} \text{tr}(\text{sym } X) \cdot \mathbb{1} + \frac{1}{n} \text{tr}(\text{sym } X) \cdot \mathbb{1} + \text{skew } X \\ &= \text{dev}_n \text{sym } X + \frac{1}{n} \text{tr}(\text{sym } X) \cdot \mathbb{1} + \text{skew } X. \end{aligned}$$

■

### 2.1.1. Operatoren auf $\mathbb{R}^{n \times n}$

**Definition 2.1.8.** Sei  $A \in \mathbb{R}^{n \times n}$  und sei  $A_{ij} \in \mathbb{R}^{(n-1) \times (n-1)}$  für  $i, j \in \{1, \dots, n\}$  die Matrix, die aus  $A$  durch Streichung ihrer  $i$ -ten Zeile und  $j$ -ten Spalte hervorgeht. Wir bezeichnen  $\tilde{a}_{ij}$  als die *Kofaktoren* von  $A$ , wenn gilt

$$\tilde{a}_{ij} = (-1)^{i+j} \det(A_{ij}). \quad (2.1.3)$$

Die Matrix aller Kofaktoren bezeichnen wir als *Kofaktor* von  $A$ , geschrieben

$$\text{Cof}(A) := (\tilde{a}_{ij})_{i,j \in \{1, \dots, n\}}. \quad (2.1.4)$$

Für  $n = 3$  gilt

$$\text{Cof}(A) := \begin{pmatrix} a_{22} a_{33} - a_{23} a_{32} & a_{23} a_{31} - a_{21} a_{33} & a_{21} a_{32} - a_{22} a_{31} \\ a_{32} a_{13} - a_{33} a_{12} & a_{33} a_{11} - a_{31} a_{13} & a_{31} a_{12} - a_{32} a_{11} \\ a_{12} a_{23} - a_{13} a_{22} & a_{13} a_{21} - a_{11} a_{23} & a_{11} a_{22} - a_{12} a_{21} \end{pmatrix}, \quad (2.1.5)$$

bzw. zählen wir die Indizes modulo 3, so gilt für die Komponenten

$$(\text{Cof } A)_{ij} = a_{i+1, j+1} a_{i+2, j+2} - a_{i+1, j+2} a_{i+2, j+1}. \quad (2.1.6)$$

**Proposition 2.1.9.** Für eine offene Menge  $\Omega \subseteq \mathbb{R}^n$  sei  $f: \Omega \rightarrow \mathbb{R}^n$  ein zweimal differenzierbares Vektorfeld. Dann ist der Kofaktor der Ableitungsmatrix von  $f$  divergenzfrei. Für  $n = 3$  bedeutet dies

$$\text{div}(\text{Cof}(Df[x])) = \text{div } \text{Cof} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \frac{\partial f_1}{\partial x_3}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \frac{\partial f_2}{\partial x_3}(x) \\ \frac{\partial f_3}{\partial x_1}(x) & \frac{\partial f_3}{\partial x_2}(x) & \frac{\partial f_3}{\partial x_3}(x) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{für alle } x \in \Omega. \quad (2.1.7)$$

*Beweis.* Wir führen den Beweis zur Veranschaulichung für  $n = 3$  durch Nachrechnen. Unter Zählung der Indizes modulo 3 erhalten wir

$$(\text{Cof } Df)_{ij} = \frac{\partial f_{i+1}}{\partial x_{j+1}} \frac{\partial f_{i+2}}{\partial x_{j+2}} - \frac{\partial f_{i+1}}{\partial x_{j+2}} \frac{\partial f_{i+2}}{\partial x_{j+1}}.$$

Dies führt auf die partiellen Ableitungen

$$\frac{\partial}{\partial x_j} (\text{Cof } Df)_{ij} = \frac{\partial^2 f_{i+1}}{\partial x_{j+1} \partial x_j} \frac{\partial f_{i+2}}{\partial x_{j+2}} - \frac{\partial^2 f_{i+1}}{\partial x_{j+2} \partial x_j} \frac{\partial f_{i+2}}{\partial x_{j+1}} + \frac{\partial f_{i+1}}{\partial x_{j+1}} \frac{\partial^2 f_{i+2}}{\partial x_{j+2} \partial x_j} - \frac{\partial f_{i+1}}{\partial x_{j+2}} \frac{\partial^2 f_{i+2}}{\partial x_{j+1} \partial x_j}.$$

Die Divergenz bestimmt sich nun zeilenweise für alle  $i \in \{1, \dots, n\}$  zu

$$\text{div}(\text{Cof } Df)_{i,\cdot} = \sum_{j=1}^n \frac{\partial}{\partial x_j} (\text{Cof } Df)_{ij} = 0,$$

wie einfaches Nachrechnen bestätigt. ■

## 2. Matrix-Analysis

Das *Vektor-* oder *Kreuzprodukt* ist eine Abbildung von  $n - 1$  Vektoren eines  $n$ -dimensionalen Vektorraums in denselben. Für den  $\mathbb{R}^n$  definieren wir

**Definition 2.1.10.** Das *Vektor-* oder *Kreuzprodukt*  $\times: \mathbb{R}^{n(n-1)} \rightarrow \mathbb{R}^n$  mit  $(v_1, \dots, v_{n-1}) \mapsto v_1 \times v_2 \times \dots \times v_{n-1}$  ist definiert als

$$v_1 \times v_2 \times \dots \times v_{n-1} := \sum_{i=1}^n e_i \cdot \det(e_i | v_1 | v_2 | \dots | v_{n-1}). \quad (2.1.8)$$

**Bemerkung 2.1.11.** Im Fall  $n = 3$  lässt sich das Kreuzprodukt direkt angeben. Es ist für  $u, v \in \mathbb{R}^3$

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \times \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{pmatrix}. \quad (2.1.9)$$

**Proposition 2.1.12.** Für alle Einheitsvektoren  $e_i \in \mathbb{R}^n$  mit  $i \in \{1, \dots, n\}$  gilt

$$\langle e_i, v_1 \times v_2 \times \dots \times v_n \rangle = \det(e_i | v_1 | v_2 | \dots | v_{n-1}). \quad (2.1.10)$$

*Beweis.* Es ist

$$\langle e_i, v_1 \times v_2 \times \dots \times v_n \rangle = \langle e_i, \sum_{j=1}^n e_j \cdot \det(e_j | v_1 | v_2 | \dots | v_{n-1}) \rangle = \det(e_i | v_1 | v_2 | \dots | v_{n-1}),$$

da  $\langle e_j, e_i \rangle = 0$  für  $j \neq i$  und  $\langle e_j, e_i \rangle = 1$  für  $j = i$ . ■

**Folgerung 2.1.13.** Für alle  $A \in \mathbb{R}^{n \times n}$  gilt

$$\langle a_1, a_2 \times \dots \times a_n \rangle = \det A, \quad (2.1.11)$$

wobei  $a_1, \dots, a_n$  die Spalten von  $A$  sind.

*Beweis.* Sei  $a_1 = \alpha_1 e_1 + \dots + \alpha_n e_n$  mit  $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ . Linearität von Skalarprodukt und Determinante liefert

$$\begin{aligned} \langle a_1, a_2 \times \dots \times a_n \rangle &= \sum_{i=1}^n \alpha_i \langle e_i, a_2 \times \dots \times a_n \rangle \\ &= \sum_{i=1}^n \alpha_i \det(e_i | a_2 \times \dots \times a_n) = \det A. \end{aligned} \quad \blacksquare$$

Das Kreuzprodukt ist orthogonal zu all seinen Argumenten:

**Folgerung 2.1.14.** Seien  $v_1, \dots, v_{n-1} \in \mathbb{R}^n$ . Dann ist  $\langle v_i, v_1 \times \dots \times v_{n-1} \rangle = 0$  für alle  $i \in \{1, \dots, n-1\}$ .

*Beweis.* Mit dem Wissen, dass die Determinante einer Matrix mit zwei identischen Spalten null ist, sehen wir leicht:

$$\langle v_i, v_1 \times \dots \times v_{n-1} \rangle = \det(v_i | v_1 | \dots | v_{n-1}) = 0. \quad \blacksquare$$

**Proposition 2.1.15.** Für alle  $A \in \mathbb{R}^{n \times n}$  und  $v_1, \dots, v_{n-1} \in \mathbb{R}^n$  gilt

$$A v_1 \times A v_2 \times \dots \times A v_{n-1} = (\text{Cof } A)(v_1 \times v_2 \times \dots \times v_{n-1}). \quad (2.1.12)$$

*Beweis.* Wir definieren die stetigen Abbildungen  $\mathcal{F}, \mathcal{G}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  mit

$$\mathcal{F}(A) := A v_1 \times A v_2 \times \dots \times A v_{n-1} \quad \text{und} \quad \mathcal{G}(A) := (\text{Cof } A)(v_1 \times v_2 \times \dots \times v_{n-1}).$$

## 2. Matrix-Analysis

Für alle  $A \in \text{GL}(n)$  gilt

$$\begin{aligned}
 \mathcal{F}(A) &= Av_1 \times Av_2 \times \dots \times Av_{n-1} \\
 &= \sum_{i=1}^n e_i \det(e_i | Av_1 | \dots | Av_{n-1}) \\
 &= \sum_{i=1}^n e_i \det(A) \cdot \det(A^{-1} e_i | v_1 | \dots | v_{n-1}) \\
 &= \sum_{i=1}^n e_i \det A \langle A^{-1} e_i, v_1 \times \dots \times v_{n-1} \rangle \\
 &= \sum_{i=1}^n e_i \langle (\text{Cof } A)^T e_i, v_1 \times \dots \times v_{n-1} \rangle \\
 &= \sum_{i=1}^n e_i \langle e_i, (\text{Cof } A)(v_1 \times \dots \times v_{n-1}) \rangle \\
 &= (\text{Cof } A)(v_1 \times \dots \times v_{n-1}) = \mathcal{G}(A).
 \end{aligned}$$

Mit Folgerung 2.1.4 folgt  $\mathcal{F}(A) = \mathcal{G}(A)$  für alle  $A \in \mathbb{R}^{n \times n}$ . ■

**Bemerkung 2.1.16.** Für  $u, v \in \mathbb{R}^3$  und  $A \in \mathbb{R}^{3 \times 3}$  gilt  $Au \times Av = (\text{Cof } A)(u \times v)$ .

**Proposition 2.1.17.** Die Determinantenabbildung  $\det: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  ist differenzierbar, ihre Ableitung ist die Kofaktorabbildung, d.h. es gilt für alle  $A \in \mathbb{R}^{n \times n}$

$$D \det(A) = \text{Cof}(A) = (\tilde{a}_{ij})_{i,j \in \{1, \dots, n\}} \quad \text{mit} \quad \tilde{a}_{ij} = (-1)^{i+j} M_{ij}. \quad (2.1.13)$$

Hierbei bezeichnet  $M_{ij}$  den  $i$ - $j$ -ten Minor von  $A$ : Die Determinante der Matrix, die aus  $A$  durch Streichung der  $i$ -ten Zeile und  $j$ -ten Spalte hervorgeht. Wir nennen die  $\tilde{a}_{ij}$  die Kofaktoren von  $A$ , die Matrix  $(\tilde{a}_{ij})$  ist die Kofaktormatrix  $\text{Cof } A$  (oder kurz: der Kofaktor von  $A$ ).

*Beweis.* Die Determinantenabbildung  $\det: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  ist ein Polynom in den Komponenten ihres Arguments und somit differenzierbar. Für ihre partiellen Ableitungen betrachten wir die Taylorentwicklung in  $A$  in Richtung  $e_i \otimes e_j$ , nutzen die Linearität der Determinante aus und entwickeln in einem zweiten Schritt nach Laplace in der  $j$ -ten Spalte:

$$\begin{aligned}
 \det(A + t \cdot e_i \otimes e_j) &= \det \begin{pmatrix} a_{11} & \dots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & \dots & a_{i,j-1} & a_{ij} + t & a_{i,j+1} & \dots & a_{in} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{pmatrix} \\
 &= \det A + \det \begin{pmatrix} a_{11} & \dots & a_{1,j-1} & 0 & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & \dots & a_{i,j-1} & t & a_{i,j+1} & \dots & a_{in} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & 0 & a_{n,j+1} & \dots & a_{nn} \end{pmatrix} \\
 &= \det(A) + \tilde{a}_{ij} \cdot t.
 \end{aligned}$$

Insgesamt ergibt sich also  $\frac{\partial \det}{\partial a_{ij}}(A) = \tilde{a}_{ij}$  und damit  $D \det(A) = \text{Cof}(A)$ . ■

**Proposition 2.1.18.** Sei  $A \in \mathbb{R}^{n \times n}$ . Dann gilt

$$A \cdot \text{Cof}(A)^T = \det A \cdot \mathbb{1}.$$

## 2. Matrix-Analysis

*Beweis.* Für beliebig vorgegebene  $i, j \in \{1, \dots, n\}$  sei  $B \in \mathbb{R}^{n \times n}$  mit

$$B := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{array}{l} \leftarrow i\text{-te Zeile} \\ \\ \leftarrow j\text{-te Zeile} \end{array}$$

die Matrix, die aus  $A$  durch Ersetzen der  $j$ -ten Zeile durch die  $i$ -te Zeile entsteht. Für  $i = j$  ist  $B = A$  und somit  $\det B = \det A$ . Für  $i \neq j$  hat  $B$  zwei identische Zeilen, die Eigenschaft des Alternierens der Determinante liefert  $\det B = 0$ .

In jedem Fall sind alle Elemente von  $B$  in der  $i$ -ten Zeile zu denen von  $A$  identisch, d.h.  $b_{ik} = a_{ik}$  für alle  $k \in \{1, \dots, n\}$ . Für die Kofaktoren von  $B$  in der  $i$ -ten Zeile kommt die spezielle Konstruktion zum Tragen: Es gilt  $\tilde{b}_{ik} = \tilde{a}_{jk}$  für alle  $k \in \{1, \dots, n\}$ .

Wir berechnen die Determinante von  $B$  durch Entwicklung in der  $i$ -ten Zeile:

$$\det B = \sum_{k=1}^n b_{ik} \tilde{b}_{ik} = \sum_{k=1}^n a_{ik} \tilde{a}_{jk} = (A \cdot \text{Cof } A^T)_{ij}.$$

Somit ist  $(A \cdot \text{Cof } A^T)_{ij} = \det A$  für  $i = j$  und  $(A \cdot \text{Cof } A^T)_{ij} = 0$  für  $i \neq j$ .

Da  $i, j$  beliebig gewählt waren, erhalten wir  $A \cdot \text{Cof}(A)^T = \det A \cdot \mathbb{1}$ . ■

**Folgerung 2.1.19.** Sei  $A \in \text{GL}(n)$ . Dann gilt

$$\text{Cof } A = \det A \cdot A^{-T}. \tag{2.1.14}$$

**Folgerung 2.1.20.** Für alle Drehmatrizen  $Q \in \text{SO}(n)$  gilt  $\text{Cof } Q = Q$  und für alle Drehspiegelmatrizen  $Q \in \text{O}(n) \setminus \text{SO}(n)$  gilt  $\text{Cof}(Q) = -Q$ .

*Beweis.* Sei  $Q \in \text{O}(n)$ , dann gilt  $Q \text{Cof}(Q)^T = \det(Q) \cdot \mathbb{1}$ , also  $\text{Cof}(Q)^T = \det(Q) \cdot Q^T$  und damit

$$\text{Cof}(Q) = \det(Q) \cdot Q. \quad \blacksquare$$

**Definition 2.1.21.** Wir nennen eine Matrix  $A \in \mathbb{R}^{n \times n}$  *regulär* oder *invertierbar*, wenn es eine Matrix  $B \in \mathbb{R}^{n \times n}$  gibt, für die  $AB = BA = \mathbb{1}$  erfüllt ist. Wir nennen  $B$  dann die *Inverse* von  $A$ .

**Proposition 2.1.22.** Seien  $A, B \in \mathbb{R}^{n \times n}$  mit  $AB = \mathbb{1}$ . Dann gilt  $BA = \mathbb{1}$ .

*Beweis.* Es ist  $B = B \cdot \mathbb{1} = B(AB) = (BA)B$ . Wegen der Eindeutigkeit des Einselements<sup>3</sup> gilt damit  $BA = \mathbb{1}$ . ■

**Bemerkung 2.1.23.** Mit Proposition 2.1.22 reicht  $AB = \mathbb{1}$  aus, um nachzuweisen, dass  $B$  zu  $A$  invers ist.

**Proposition 2.1.24.** Eine Matrix  $A \in \mathbb{R}^{n \times n}$  ist genau dann invertierbar, wenn  $\det A \neq 0$  gilt.

*Beweis.* Mit  $\det A \neq 0$  und Proposition 2.1.18 gilt  $A \cdot \frac{\text{Cof}(A)^T}{\det A} = \mathbb{1}$ , damit ist  $A^{-1} := \frac{\text{Cof}(A)^T}{\det A}$  die Inverse von  $A$ .

Mit der Existenz von  $A^{-1}$  gilt  $1 = \det \mathbb{1} = \det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1})$ , also  $\det A \neq 0$ . ■

<sup>3</sup>Seien  $\mathbb{1}_1$  und  $\mathbb{1}_2$  Einselemente, dann gilt  $\mathbb{1}_1 = \mathbb{1}_1 \mathbb{1}_2 = \mathbb{1}_2$ .

## 2. Matrix-Analysis

**Folgerung 2.1.25.** Ist  $A \in \text{GL}(n)$ , dann können wir die Inverse von  $A$  mittels Determinante und Kofaktor bestimmen:

$$A^{-1} = \frac{1}{\det A} \cdot \text{Cof}(A)^T. \quad (2.1.15)$$

**Bemerkung 2.1.26.** Die Transponierte der Kofaktormatrix  $\text{Cof}(A)^T$  wird auch *Adjunkte* genannt.

**Proposition 2.1.27.** Die Spurabbildung  $\text{tr}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  ist differenzierbar, ihre Ableitung ist die Identität, d.h. es gilt  $D \text{tr}(A) = \mathbb{1}$  für alle  $A \in \mathbb{R}^{n \times n}$ .

*Beweis.* Die Spurabbildung  $\text{tr}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  ist ein Polynom in den Komponenten ihres Arguments und somit differenzierbar. Für ihre partiellen Ableitungen betrachten wir die Taylorentwicklung in  $A$  in Richtung  $e_i \otimes e_j$ , nutzen die Linearität der Spur aus:

$$\text{tr}(A + t \cdot e_i \otimes e_j) = \text{tr}(A) + t \cdot \text{tr}(e_i \otimes e_j),$$

und damit  $\frac{\partial \text{tr}}{\partial a_{ij}}(A) = 1$  für  $i = j$  und  $\frac{\partial \text{tr}}{\partial a_{ij}}(A) = 0$  für  $i \neq j$ , was auf  $D \text{tr}(A) = \mathbb{1}$  führt. ■

**Proposition 2.1.28.** Seien  $A, B \in \mathbb{R}^{n \times n}$ . Dann gilt  $\text{Cof}(AB) = \text{Cof}(A) \text{Cof}(B)$ .

*Beweis.* Wir definieren die stetigen Abbildungen  $\mathcal{F}, \mathcal{G}: \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  mit

$$\mathcal{F}(A, B) := \text{Cof}(AB) \quad \text{und} \quad \mathcal{G}(A, B) := \text{Cof}(A) \text{Cof}(B).$$

Für alle  $A, B \in \text{GL}(n)$  ist auch  $AB \in \text{GL}(n)$  und es gilt mit Folgerung 2.1.25 und dem Determinantenproduktsatz

$$\begin{aligned} \mathcal{F}(A, B) &= \text{Cof}(AB) = \det(AB)(AB)^{-T} = \det(A) \det(B) A^{-T} B^{-T} \\ &= \det(A) A^{-T} \cdot \det(B) B^{-T} = \text{Cof}(A) \text{Cof}(B) = \mathcal{G}(A, B). \end{aligned}$$

Mit Folgerung 2.1.4 folgt  $\mathcal{F}(A) = \mathcal{G}(B)$  für alle  $A \in \mathbb{R}^{n \times n}$ . ■

**Folgerung 2.1.29.** Seien  $A \in \mathbb{R}^{n \times n}$  und  $Q \in \text{O}(n)$ . Dann gilt

$$\text{Cof}(QA) = \det Q \cdot Q \text{Cof}(A) \quad \text{und} \quad \text{Cof}(AQ) = \det Q \cdot \text{Cof}(A) Q. \quad (2.1.16)$$

*Beweis.* Die Aussage folgt direkt aus der vorangegangenen Proposition und Folgerung 2.1.20, welches  $\text{Cof}(Q) = \det Q \cdot Q$  sicherstellt. ■

**Folgerung 2.1.30.** Die Kofaktorabbildung ist isotrop<sup>4</sup>, d.h. für  $A \in \mathbb{R}^{n \times n}$  und  $Q \in \text{O}(n)$  gilt

$$\text{Cof}(Q A Q^T) = Q \text{Cof}(A) Q^T. \quad (2.1.17)$$

## 2.2. Frobenius-Skalarprodukt und Frobeniusnorm

Fassen wir den  $\mathbb{R}^{m \times n}$  als  $\mathbb{R}^{m \cdot n}$  auf, so ist das folgende *Frobenius-Skalarprodukt* identisch mit dem Standardskalarprodukt und wird deshalb auch so genannt:

**Definition 2.2.1.** Das *Frobenius-Skalarprodukt*  $\langle \cdot, \cdot \rangle: \mathbb{R}^{m \times n} \times \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$  ist definiert als

$$\langle A, B \rangle := \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij}. \quad (2.2.1)$$

**Bemerkung 2.2.2.** Fassen wir den  $\mathbb{R}^{m \times n}$  als  $\mathbb{R}^{m \cdot n}$  auf, so ist das Frobenius-Skalarprodukt identisch mit dem Standardskalarprodukt und wird deshalb ebenfalls im  $\mathbb{R}^{m \times n}$  als *Standardskalarprodukt* bezeichnet.

---

<sup>4</sup>siehe auch Definition 2.5.1

## 2. Matrix-Analysis

**Folgerung 2.2.3.** Seien  $A, B \in \mathbb{R}^{m \times n}$ . Für das Frobenius-Skalarprodukt gilt  $\langle A, B \rangle = \text{tr}(A^T B)$ .

*Beweis.* Ist  $C := A^T B$ , so ist  $c_{jj} = \sum_{i=1}^m a_{ij} b_{ij}$ . Damit erhalten wir

$$\text{tr } A^T B = \sum_{j=1}^n c_{jj} = \sum_{j=1}^n \sum_{i=1}^m a_{ij} b_{ij} = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij} = \langle A, B \rangle. \quad \blacksquare$$

**Proposition 2.2.4.** Für alle  $A, B \in \mathbb{R}^{m \times n}$ ,  $X \in \mathbb{R}^{m \times m}$  und  $Y \in \mathbb{R}^{n \times n}$  gilt

$$\langle XA, B \rangle = \langle A, X^T B \rangle \quad \text{und} \quad \langle AY, B \rangle = \langle A, BY^T \rangle. \quad (2.2.2)$$

*Beweis.* Wir rechnen unter Zuhilfenahme von Folgerung 2.2.3 nach. Es ist

$$\langle XA, B \rangle = \text{tr}((XA)^T B) = \text{tr}(A^T X^T B) = \langle A, X^T B \rangle.$$

Unter Ausnutzer der Invarianz der Spur unter zyklischen Vertauschungen gilt zudem

$$\langle AY, B \rangle = \text{tr}((AY)^T B) = \text{tr}(Y^T A^T B) = \text{tr}(A^T B Y^T) = \langle A, B Y^T \rangle. \quad \blacksquare$$

**Folgerung 2.2.5.** Für alle  $A, B \in \mathbb{R}^{m \times n}$ ,  $Q \in O(m)$  und  $R \in O(n)$  gilt

$$\langle QA, QB \rangle = \langle AR, BR \rangle = \langle A, B \rangle. \quad (2.2.3)$$

**Folgerung 2.2.6.** Für alle  $A, B \in \mathbb{R}^{n \times n}$ ,  $Q \in O(n)$  gilt

$$\langle Q^T A Q, Q^T B Q \rangle = \langle A, B \rangle. \quad (2.2.4)$$

**Proposition 2.2.7.** Gegeben sei eine symmetrische spurfreie Matrix  $X \in \text{Sym}(n) \cap \mathfrak{sl}(n)$ , eine schiefssymmetrische Matrix  $Y \in \mathfrak{so}(n)$  und eine Diagonalmatrix mit identischen Einträgen  $Z \in \mathbb{R} \cdot \mathbb{1}$ . Dann sind  $X$ ,  $Y$  und  $Z$  paarweise orthogonal.

*Beweis.* Wir haben  $X, Y, Z \in \mathbb{R}^{n \times n}$  mit

$$X^T = X, \quad \text{tr}(X) = 0, \quad Y^T = -Y \quad \text{und} \quad Z = \lambda \cdot \mathbb{1}$$

für ein  $\lambda \in \mathbb{R}$ . Dann ist

$$\begin{aligned} \langle X, Y \rangle &= \langle Y, X \rangle = \text{tr}(Y^T X) = \text{tr}(-Y X^T) = -\text{tr}(X^T Y) = -\langle X, Y \rangle, \\ \langle Y, Z \rangle &= \text{tr}(Y^T \lambda \cdot \mathbb{1}) = \text{tr}(-Y \lambda \cdot \mathbb{1}) = -\text{tr}((\lambda \cdot \mathbb{1})^T Y) = -\langle Z, Y \rangle = -\langle Y, Z \rangle \end{aligned}$$

und

$$\langle Z, X \rangle = \text{tr}((\lambda \cdot \mathbb{1})^T X) = \lambda \text{tr}(X) = 0.$$

Somit muss  $\langle X, Y \rangle = \langle Y, Z \rangle = \langle Z, X \rangle = 0$  gelten. ■

**Proposition 2.2.8.** Es ist  $\mathbb{R}^{n \times n} = (\text{Sym}(n) \cap \mathfrak{sl}(n)) \oplus \mathfrak{so}(n) \oplus \mathbb{R} \cdot \mathbb{1}$ .

*Beweis.* Sei  $X \in (\text{Sym}(n) \cap \mathfrak{sl}(n)) \cap \mathfrak{so}(n)$ , dann gilt  $X^T = X$ ,  $\text{tr } X = 0$  und  $X^T = -X$ , also  $X = -X$  und damit  $X = 0$ .

Sei  $X \in \mathfrak{so}(n) \cap \mathbb{R} \cdot \mathbb{1}$ , dann gilt  $X^T = -X$  und  $X = \lambda \cdot \mathbb{1}$ , also  $X = \lambda \cdot \mathbb{1} = \lambda \mathbb{1}^T = X^T = -X$  und damit  $X = 0$ .

Sei  $X \in (\text{sym}(n) \cap \mathfrak{sl}(n)) \cap \mathbb{R} \cdot \mathbb{1}$ , also  $X^T = X$ ,  $\text{tr } X = 0$  und  $X = \lambda \cdot \mathbb{1}$ , also  $0 = \text{tr } X = n \cdot \lambda$  und damit  $\lambda = 0$ , was  $X = 0$  impliziert.

Wir haben in Proposition 2.1.7 gesehen, dass sich jedes  $X \in \mathbb{R}^{n \times n}$  aufspalten lässt in

$$X = \underbrace{\text{dev}_n \text{sym } X}_{\in \text{Sym}(n) \cap \mathfrak{sl}(n)} + \underbrace{\text{skew } X}_{\in \mathfrak{so}(n)} + \underbrace{\frac{1}{n} \text{tr}(\text{sym } X) \cdot \mathbb{1}}_{\in \mathbb{R} \cdot \mathbb{1}}, \quad (2.2.5)$$

somit gilt  $\mathbb{R}^{n \times n} = (\text{Sym}(n) \cap \mathfrak{sl}(n)) \oplus \mathfrak{so}(n) \oplus \mathbb{R} \cdot \mathbb{1}$  ■



## 2. Matrix-Analysis

**Bemerkung 2.2.9.** Wir können also jedes  $X \in \mathbb{R}^{n \times n}$  eindeutig in der Form (2.2.5) aufspalten.

**Proposition 2.2.10.** Seien  $\mathcal{U}_i \subseteq \mathbb{R}^{n \times n}$  mit  $\mathbb{R}^{n \times n} = \bigoplus_{i=1}^n \mathcal{U}_i$  gegeben. Dann gilt für alle  $X, Y \in \mathbb{R}^{n \times n}$  mit  $X = \sum_{i=1}^n X_i$ , sowie  $Y = \sum_{i=1}^n Y_i$  und  $X_i, Y_i \in \mathcal{U}_i$

$$\langle X, Y \rangle = \sum_{i=1}^n \langle X_i, Y_i \rangle. \quad (2.2.6)$$

*Beweis.* Mit der Linearität des Skalarproduktes und dem Umstand, dass  $\langle A, B \rangle = 0$  für alle  $A \in \mathcal{U}_i$  und  $B \in \mathcal{U}_j$  mit  $i \neq j$ , gilt

$$\langle X, Y \rangle = \left\langle \sum_{i=1}^n X_i, \sum_{i=1}^n Y_i \right\rangle = \sum_{i,j=1}^n \langle X_i, Y_j \rangle = \sum_{i=1}^n \langle X_i, Y_i \rangle + \sum_{i \neq j} \underbrace{\langle X_i, Y_j \rangle}_{=0} = \sum_{i=1}^n \langle X_i, Y_i \rangle. \quad \blacksquare$$

Dies gibt uns die Möglichkeit der Zerlegung des Frobenius-Skalarproduktes für quadratische Matrizen in spurfreien Anteil und Spurateil bzw. in spurfrei symmetrischen Anteil, schiefssymmetrischen Anteil und Spurateil:

**Folgerung 2.2.11.** Für alle  $X, Y \in \mathbb{R}^{n \times n}$  gilt

$$\langle X, Y \rangle = \langle \text{dev}_n X, \text{dev}_n Y \rangle + \frac{1}{n} \text{tr}(X) \text{tr}(Y) \quad (2.2.7)$$

$$= \langle \text{dev}_n \text{sym} X, \text{dev}_n \text{sym} Y \rangle + \langle \text{skew} X, \text{skew} Y \rangle + \frac{1}{n} \text{tr}(X) \text{tr}(Y). \quad (2.2.8)$$

*Beweis.* Aus

$$X = \underbrace{\text{dev}_n \text{sym} X}_{\in \text{Sym}(n) \cap \text{sl}(n)} + \underbrace{\text{skew} X}_{\in \mathfrak{so}(n)} + \underbrace{\frac{1}{n} \text{tr}(\text{sym} X) \cdot \mathbb{1}}_{\in \mathbb{R} \cdot \mathbb{1}},$$

folgt

$$\begin{aligned} \langle X, Y \rangle &= \langle \text{dev}_n X, \text{dev}_n Y \rangle + \left\langle \frac{1}{n} \text{tr}(X) \cdot \mathbb{1}, \frac{1}{n} \text{tr}(Y) \cdot \mathbb{1} \right\rangle \\ &= \langle \text{dev}_n \text{sym} X, \text{dev}_n \text{sym} Y \rangle + \langle \text{skew} X, \text{skew} Y \rangle + \left\langle \frac{1}{n} \text{tr}(X) \cdot \mathbb{1}, \frac{1}{n} \text{tr}(Y) \cdot \mathbb{1} \right\rangle. \end{aligned}$$

Und mit der Linearität des Skalarproduktes gilt

$$\left\langle \frac{1}{n} \text{tr}(X) \cdot \mathbb{1}, \frac{1}{n} \text{tr}(Y) \cdot \mathbb{1} \right\rangle = \frac{1}{n^2} \text{tr}(X) \text{tr}(Y) \cdot \langle \mathbb{1}, \mathbb{1} \rangle = \frac{1}{n} \text{tr}(X) \text{tr}(Y). \quad \blacksquare$$

Die Zerlegung in den spurfreien symmetrischen Anteil, den schiefssymmetrischen Anteil und den Spurateil

$$\langle X, Y \rangle = \langle \text{dev}_n \text{sym} X, \text{dev}_n \text{sym} Y \rangle + \langle \text{skew} X, \text{skew} Y \rangle + \frac{1}{n} \text{tr}(X) \text{tr}(Y)$$

gibt uns noch die Möglichkeit der Gewichtung:

**Proposition 2.2.12.** Die Abbildung  $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa} : \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  mit  $\mu, \mu_c, \kappa \in \mathbb{R}_+$  und

$$\langle A, B \rangle_{\mu, \mu_c, \kappa} := \mu \langle \text{dev}_n \text{sym} A, \text{dev}_n \text{sym} B \rangle + \mu_c \langle \text{skew} A, \text{skew} B \rangle + \frac{\kappa}{2} \text{tr}(A) \text{tr}(B). \quad (2.2.9)$$

bildet ein Skalarprodukt auf  $\mathbb{R}^{n \times n}$ . Wir nennen sie das gewichtete isotrope Skalarprodukt.

*Beweis.* Wir zeigen,  $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa}$  ist symmetrisch, linear und positiv definit.

## 2. Matrix-Analysis

i) Die *Symmetrie* folgt direkt aus der Symmetrie des Frobenius-Skalarproduktes und der Kommutativität der Multiplikation:

$$\begin{aligned}\langle Y, X \rangle_{\mu, \mu_c, \kappa} &= \mu \langle \text{dev}_n \text{sym } Y, \text{dev}_n \text{sym } X \rangle + \mu_c \langle \text{skew } Y, \text{skew } X \rangle + \frac{\kappa}{2} \text{tr}(Y) \cdot \text{tr}(X) \\ &= \mu \langle \text{dev}_n \text{sym } X, \text{dev}_n \text{sym } Y \rangle + \mu_c \langle \text{skew } X, \text{skew } Y \rangle + \frac{\kappa}{2} \text{tr}(X) \cdot \text{tr}(Y) = \langle X, Y \rangle_{\mu, \mu_c, \kappa}\end{aligned}$$

ii) Aufgrund der Symmetrie reicht es, die *Linearität* nur im ersten Argument zu zeigen. Sie folgt direkt aus der Linearität von  $\text{dev}$ ,  $\text{sym}$ ,  $\text{skew}$ ,  $\text{tr}$  und des Frobenius-Skalarproduktes:

$$\begin{aligned}\langle \alpha X + \beta Z, Y \rangle_{\mu, \mu_c, \kappa} &= \mu \langle \text{dev}_n \text{sym}(\alpha X + \beta Z), \text{dev}_n \text{sym}(Y) \rangle + \mu_c \langle \text{skew}(\alpha X + \beta Z), \text{skew}(Y) \rangle + \frac{\kappa}{2} \text{tr}(\alpha X + \beta Z) \cdot \text{tr}(Y) \\ &= \mu(\alpha \langle \text{dev}_n \text{sym } X, \text{dev}_n \text{sym } Y \rangle + \beta \langle \text{dev}_n \text{sym } Z, \text{dev}_n \text{sym } Y \rangle) \\ &\quad + \mu_c(\alpha \langle \text{skew } X, \text{skew } Y \rangle + \beta \langle \text{skew } Z, \text{skew } Y \rangle) \\ &\quad + \frac{\kappa}{2}(\alpha \text{tr}(X) \cdot \text{tr}(Y) + \beta \text{tr}(Z) \cdot \text{tr}(Y)) \\ &= \alpha \langle X, Y \rangle_{\mu, \mu_c, \kappa} + \beta \langle Z, Y \rangle_{\mu, \mu_c, \kappa}\end{aligned}$$

iii) Zur *positiven Definitheit*: Es ist

$$\langle X, X \rangle_{\mu, \mu_c, \kappa} = \underbrace{\mu \langle \text{dev}_n \text{sym } X, \text{dev}_n \text{sym } X \rangle}_{\geq 0} + \underbrace{\mu_c \langle \text{skew } X, \text{skew } X \rangle}_{\geq 0} + \underbrace{\frac{\kappa}{2} \text{tr}(X)^2}_{\geq 0},$$

somit  $\langle X, X \rangle_{\mu, \mu_c, \kappa} \geq 0$  für alle  $X \in \mathbb{R}^{n \times n}$ . Zudem folgt aus  $X = 0$  offenkundig  $\langle X, X \rangle_{\mu, \mu_c, \kappa} = 0$ . Ist andererseits  $\langle X, X \rangle_{\mu, \mu_c, \kappa} = 0$ , so muss aufgrund der positiven Definitheit des Frobenius-Skalarproduktes  $\text{dev}_n \text{sym } X = 0$ ,  $\text{skew } X = 0$  und  $\text{tr}(X) = 0$  gelten. Nun folgt  $\text{sym}(X) = \text{dev}_n \text{sym}(X) + \frac{1}{n} \text{tr}(X) \cdot \mathbb{1} = 0$ , also  $X = \text{sym}(X) + \text{skew}(X) = 0$ . ■

**Proposition 2.2.13.** Für alle  $A, B \in \mathbb{R}^{n \times n}$  und  $Q \in O(n)$  gilt

$$\langle Q^T A Q, Q^T B Q \rangle_{\mu, \mu_c, \kappa} = \langle A, B \rangle_{\mu, \mu_c, \kappa}. \quad (2.2.10)$$

*Beweis.* Wir nutzen die Definition des gewichteten, isotropen Skalarproduktes, verwenden die Isotropie von  $\text{sym}$ ,  $\text{dev}_n$ ,  $\text{skew}$  und  $\text{tr}$ . Schließlich stellen wir die Identität mit Folgerung 2.2.6 her:

$$\begin{aligned}\langle Q^T X Q, Q^T Y Q \rangle_{\mu, \mu_c, \kappa} &= \mu \langle \text{dev}_n \text{sym}(Q^T X Q), \text{dev}_n \text{sym}(Q^T Y Q) \rangle + \mu_c \langle \text{skew}(Q^T X Q), \text{skew}(Q^T Y Q) \rangle + \frac{\kappa}{2} \text{tr}(Q^T X Q) \text{tr}(Q^T Y Q) \\ &= \mu \langle Q^T \text{dev}_n \text{sym}(X) Q, Q^T \text{dev}_n \text{sym}(Y) Q \rangle + \mu_c \langle Q^T \text{skew}(X) Q, Q^T \text{skew}(Y) Q \rangle + \frac{\kappa}{2} \text{tr}(X) \text{tr}(Y) \\ &= \mu \langle \text{dev}_n \text{sym } X, \text{dev}_n \text{sym } X \rangle + \mu_c \langle \text{skew } X, \text{skew } X \rangle + \frac{\kappa}{2} \text{tr}(X) \text{tr}(Y) \\ &= \langle X, Y \rangle_{\mu, \mu_c, \kappa}.\end{aligned}$$
 ■

**Proposition 2.2.14.** Seien  $X, Y \in \mathbb{R}^{n \times n}$  und  $\mathbb{C}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  eine lineare Abbildung mit

$$\mathbb{C}.X := \mu \text{dev}_n \text{sym } X + \mu_c \text{skew } X + \frac{\kappa}{2} \text{tr}(X) \cdot \mathbb{1}. \quad (2.2.11)$$

Dann ist  $\langle X, Y \rangle_{\mu, \mu_c, \kappa} = \langle \mathbb{C}.X, Y \rangle$ .

## 2. Matrix-Analysis

*Beweis.* Es gilt sowohl

$$\langle \text{dev}_n \text{sym } X, Y \rangle = \underbrace{\langle \underbrace{\text{dev}_n \text{sym } X}_{\in \text{Sym}(n) \cap \mathfrak{sl}(n)}, \underbrace{\frac{1}{n} \text{tr}(\text{sym } Y) \cdot \mathbb{1}}_{\in \mathbb{R} \cdot \mathbb{1}} \rangle}_{=0} + \underbrace{\langle \underbrace{\text{dev}_n \text{sym } X}_{\in \text{Sym}(n) \cap \mathfrak{sl}(n)}, \underbrace{\text{skew } Y}_{\in \mathfrak{so}(n)} \rangle}_{=0}$$

als auch

$$\langle \text{skew } X, Y \rangle = \langle \text{skew } X, \text{skew } Y \rangle + \underbrace{\langle \underbrace{\text{skew } X}_{\in \mathfrak{so}(n)}, \underbrace{\text{sym } Y}_{\in \text{Sym}(n)} \rangle}_{=0}$$

und damit

$$\begin{aligned} \langle \mathbb{C}.X, Y \rangle &= \langle \mu \text{dev}_n \text{sym } X + \mu_c \text{skew } X + \frac{\kappa}{2} \text{tr}(x) \cdot \mathbb{1}, Y \rangle \\ &= \mu \langle \text{dev}_n \text{sym } X, Y \rangle + \mu_c \langle \text{skew } X, Y \rangle + \frac{\kappa}{2} \text{tr}(X) \cdot \langle \mathbb{1}, Y \rangle \\ &= \mu \langle \text{dev}_n \text{sym } X, \text{dev}_n \text{sym } Y \rangle + \mu_c \langle \text{skew } X, \text{skew } Y \rangle + \frac{\kappa}{2} \text{tr}(X) \cdot \text{tr}(Y) \\ &= \langle X, Y \rangle_{\mu, \mu_c, \kappa}. \end{aligned}$$

**Proposition 2.2.15.** Seien  $X, Y \in \mathbb{R}^{n \times n}$ . Dann gilt für  $\frac{\lambda}{2} = \frac{\kappa}{2} - \frac{\mu}{n}$ :

$$\langle X, Y \rangle_{\mu, \mu_c, \kappa} = \mu \langle \text{sym } X, \text{sym } Y \rangle + \mu_c \langle \text{skew } X, \text{skew } Y \rangle + \frac{\lambda}{2} \text{tr}(X) \cdot \text{tr}(Y). \quad (2.2.12)$$

*Beweis.* Es ist

$$\begin{aligned} \langle \text{sym } X, \text{sym } Y \rangle &= \langle \text{dev}_n \text{sym } X + \frac{1}{n} \text{tr}(\text{sym } X) \cdot \mathbb{1}, \text{dev}_n \text{sym } Y + \frac{1}{n} \text{tr}(\text{sym } Y) \cdot \mathbb{1} \rangle \\ &= \langle \text{dev}_n \text{sym } X, \text{dev}_n \text{sym } Y \rangle + \frac{1}{n^2} \text{tr}(\text{sym } X) \text{tr}(\text{sym } Y) \underbrace{\langle \mathbb{1}, \mathbb{1} \rangle}_{=n} \\ &\quad + \frac{1}{n} \text{tr}(\text{sym } X) \underbrace{\langle \mathbb{1}, \text{dev}_n \text{sym } Y \rangle}_{=\text{tr}(\text{dev}_n \text{sym } Y)=0} + \frac{1}{n} \text{tr}(\text{sym } Y) \underbrace{\langle \text{dev}_n \text{sym } X, \mathbb{1} \rangle}_{=\text{tr}(\text{dev}_n \text{sym } X)=0} \\ &= \langle \text{dev}_n \text{sym } X, \text{dev}_n \text{sym } Y \rangle + \frac{1}{n} \text{tr}(X) \text{tr}(Y) \end{aligned}$$

also

$$\begin{aligned} \langle X, Y \rangle_{\mu, \mu_c, \kappa} &= \mu \langle \text{dev}_n \text{sym } X, \text{dev}_n \text{sym } Y \rangle + \mu_c \langle \text{skew } X, \text{skew } Y \rangle + \frac{\kappa}{2} \text{tr}(X) \cdot \text{tr}(Y) \\ &= \mu \langle \text{sym } X, \text{sym } Y \rangle - \frac{\mu}{n} \text{tr}(X) \text{tr}(Y) + \mu_c \langle \text{skew } X, \text{skew } Y \rangle + \frac{\kappa}{2} \text{tr}(X) \cdot \text{tr}(Y) \\ &= \mu \langle \text{sym } X, \text{sym } Y \rangle + \mu_c \langle \text{skew } X, \text{skew } Y \rangle + \frac{\lambda}{2} \text{tr}(X) \cdot \text{tr}(Y) \end{aligned}$$

mit  $\frac{\lambda}{2} = \frac{\kappa}{2} - \frac{\mu}{n}$ . ■

**Definition 2.2.16.** Das Frobenius-Skalarprodukt induziert die *Frobeniusnorm* für quadratische Matrizen  $\|\cdot\|: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  mit  $\|A\| := \sqrt{\text{tr } A^T A}$ .

**Bemerkung 2.2.17.** Somit entspricht das Quadrat der Frobeniusnorm einer Matrix der Summe der quadrierten Elemente der Matrix, d.h.

$$\|A\|^2 = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2. \quad (2.2.13)$$

## 2. Matrix-Analysis

Wir erinnern uns an die Zerlegung des Skalarproduktes in

$$\begin{aligned}\langle X, Y \rangle &= \langle \operatorname{dev}_n X, \operatorname{dev}_n Y \rangle + \frac{1}{n} \operatorname{tr}(X) \operatorname{tr}(Y) \\ &= \langle \operatorname{dev}_n \operatorname{sym} X, \operatorname{dev}_n \operatorname{sym} Y \rangle + \langle \operatorname{skew} X, \operatorname{skew} Y \rangle + \frac{1}{n} \operatorname{tr}(X) \operatorname{tr}(Y).\end{aligned}$$

Durch eine Gewichtung erhielten wir das gewichtete, isotrope Skalarprodukt. Aber auch ungewichtet ergibt diese Zerlegung schon eine wichtige Gleichheit:

**Proposition 2.2.18.** *Für alle  $X \in \mathbb{R}^{n \times n}$  gilt*

$$\|X\|^2 = \|\operatorname{dev}_n X\|^2 + \frac{1}{n} |\operatorname{tr} X|^2 = \|\operatorname{dev}_n \operatorname{sym} X\|^2 + \|\operatorname{skew} X\|^2 + \frac{1}{n} |\operatorname{tr} X|^2.$$

Gewichten wir nun, erhalten wir:

**Definition 2.2.19.** Das gewichtete, isotrope Skalarprodukt  $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa} : \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  induziert die *gewichtete Frobeniusnorm* für quadratische Matrizen  $\|\cdot\|_{\mu, \mu_c, \kappa} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  mit

$$\|A\|_{\mu, \mu_c, \kappa}^2 := \mu \|\operatorname{dev}_n \operatorname{sym} A\|^2 + \mu_c \|\operatorname{skew} A\|^2 + \frac{\kappa}{2} \cdot \operatorname{tr}(A)^2.$$

**Bemerkung 2.2.20.** Mit  $\mu = \mu_c = 1$  und  $\kappa = \frac{2}{n}$  erhalten wir  $\|A\|_{\mu, \mu_c, \kappa}^2 = \|A\|^2$ .

**Proposition 2.2.21.** *Die Frobeniusnorm ist submultiplikativ, d.h. für alle  $A, B \in \mathbb{R}^{n \times n}$  gilt*

$$\|AB\| \leq \|A\| \cdot \|B\|. \quad (2.2.14)$$

*Beweis.* Bezeichnen wir mit  $u := (a_{i1}, \dots, a_{in})$  die  $i$ -te Zeile von  $A$  und mit  $v := (b_{1j}, \dots, b_{nj})$  die  $j$ -te Spalte von  $B$ , so gewährt uns die Cauchy-Schwarzsche Ungleichung auf  $\mathbb{R}^n$  genau  $|\langle u, v \rangle| \leq \|u\| \cdot \|v\|$  und damit

$$\left( \sum_{k=1}^n a_{ik} b_{kj} \right)^2 \leq \left( \sum_{k=1}^n a_{ik}^2 \right) \left( \sum_{k=1}^n b_{kj}^2 \right).$$

Nun können wir abschätzen:

$$\begin{aligned}\|AB\|^2 &= \left\| \left( \sum_{k=1}^n a_{ik} b_{kj} \right)_{ij} \right\|^2 \\ &= \sum_{i=1}^n \sum_{j=1}^n \left( \sum_{k=1}^n a_{ik} b_{kj} \right)^2 \leq \sum_{i=1}^n \sum_{j=1}^n \left( \sum_{k=1}^n a_{ik}^2 \right) \left( \sum_{k=1}^n b_{kj}^2 \right) \\ &= \left( \sum_{i=1}^n \sum_{k=1}^n a_{ik}^2 \right) \left( \sum_{k=1}^n \sum_{j=1}^n b_{kj}^2 \right) = \|A\|^2 \cdot \|B\|^2.\end{aligned}$$

Mit der Positivität der Norm folgt die Aussage. ■

**Proposition 2.2.22.** *Sowohl die Frobeniusnorm, als auch die gewichtete, isotrope Frobeniusnorm sind isotrop<sup>5</sup>, d.h. für alle  $X \in \mathbb{R}^{n \times n}$  und  $Q \in O(n)$  gilt*

$$\|Q^T X Q\| = \|X\| \quad \text{und} \quad \|Q^T X Q\|_{\mu, \mu_c, \kappa} = \|X\|_{\mu, \mu_c, \kappa}. \quad (2.2.15)$$

*Beweis.* Die Frobeniusnorm wird durch das Frobenius-Skalarprodukt induziert, so gilt unter Ausnutzung von Folgerung 2.2.6

$$\|Q^T X Q\|^2 = \langle Q^T X Q, Q^T X Q \rangle = \langle X, X \rangle = \|X\|^2.$$

Völlig analog verläuft die Rechnung bei der isotropen, gewichteten Norm, die durch das isotrope, gewichtete Skalarprodukt induziert wird. Hier verwenden wir Proposition 2.2.13. ■

<sup>5</sup>siehe auch Definition 2.5.1

## 2. Matrix-Analysis

**Lemma 2.2.23.** Sei  $A \in \text{Sym}(n)$  eine reelle, symmetrische  $n \times n$ -Matrix. Dann gilt:

- i)  $A$  ist selbstadjugiert. D.h. für alle  $u, v \in \mathbb{R}^n$  gilt  $\langle Au, v \rangle = \langle u, Av \rangle$ .
- ii) Jeder Eigenwert  $\lambda$  von  $A$  ist reell. D.h. für jedes  $\lambda \in \mathbb{C}$  für das es ein  $v \in \mathbb{R}^n$  gibt mit  $Av = \lambda v$ , gilt  $\lambda \in \mathbb{R}$ . Insbesondere sind auch die Eigenvektoren reell.
- iii) Sind  $u, v \in \mathbb{R}^n$  Eigenvektoren von  $A$  zu unterschiedlichen Eigenwerten, so sind sie orthogonal zueinander. D.h. für  $u, v \in \mathbb{R}^n$  und  $\lambda, \mu \in \mathbb{R}$  mit  $Au = \lambda u$ ,  $Av = \mu v$  und  $\lambda \neq \mu$  gilt  $\langle u, v \rangle = 0$ .
- iv) Die Matrix  $A$  ist orthogonal diagonalisierbar. D.h. es existieren (nicht notwendigerweise verschiedene) Eigenwerte  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  und ein  $Q \in O(n)$  sodass gilt:

$$A = Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^T.$$

Die Spalten von  $Q$  sind die Eigenvektoren zu den Eigenwerten  $\lambda_1, \dots, \lambda_n$ . Sie werden auch Hauptachsen von  $A$  genannt, die Transformation  $A \mapsto Q^T A Q$  in die Diagonalmatrix  $\text{diag}(\lambda_1, \dots, \lambda_n)$  auch Hauptachsentransformation.

- v) Für die Spur von  $A$  gilt:

$$\text{tr}(A) = \lambda_1 + \dots + \lambda_n.$$

- vi) Für die Determinante von  $A$  gilt:

$$\det A = \lambda_1 \cdot \dots \cdot \lambda_n.$$

- vii) Für die Norm von  $A$  gilt:

$$\|A\|^2 = \lambda_1^2 + \dots + \lambda_n^2.$$

*Beweis.* i): Seien  $u, v \in \mathbb{R}^n$ . Mit der Symmetrie von  $A$  erhalten wir  $\langle Au, v \rangle = \langle u, A^T v \rangle = \langle u, Av \rangle$ .

ii): Sei  $v \in \mathbb{C}^n$  Eigenvektor zum Eigenwert  $\lambda \in \mathbb{C}$ . Mit der Selbstadjungiertheit gilt

$$\lambda \langle v, v \rangle = \langle Av, v \rangle = \langle v, Av \rangle = \bar{\lambda} \langle v, v \rangle$$

und damit  $\lambda = \bar{\lambda}$ . Da  $v$  Lösung des reellen Gleichungssystems  $(A - \lambda \cdot \mathbb{1})v = 0$  ist, ist  $v \in \mathbb{R}^n$ . Die Eigenvektoren einer symmetrischen Matrix sind somit ebenfalls reell.

iii): Sei  $u$  Eigenvektor zum Eigenwert  $\lambda$  und  $v$  Eigenvektor zum Eigenwert  $\mu$  mit  $\lambda \neq \mu$ . Mit ii) ist  $\lambda, \mu \in \mathbb{R}$  und  $u, v \in \mathbb{R}^n$  und mit dem reellen Skalarprodukt gilt

$$\lambda \langle u, v \rangle = \langle Au, v \rangle = \langle u, Av \rangle = \mu \langle u, v \rangle,$$

also  $(\lambda - \mu) \langle u, v \rangle = 0$ . Mit  $\lambda \neq \mu$  erhalten wir  $\langle u, v \rangle = 0$ .

iv): Wir beweisen die Aussage per vollständiger Induktion über  $n$ .

Für  $n = 1$  ist  $A$  bereits in Diagonalgestalt.

Nehmen wir nun an, dass es zu jeder symmetrischen  $n \times n$ -Matrix Eigenwerte  $\lambda_1, \dots, \lambda_n$  und ein  $Q \in O(n)$  gibt, dessen Spalten Eigenvektoren von  $A$  zu  $\lambda_1, \dots, \lambda_n$  sind, sodass  $A = Q^T \text{diag}(\lambda_1, \dots, \lambda_n) Q$  gilt. Sei  $A$  eine symmetrische  $(n+1) \times (n+1)$ -Matrix. Nach dem Fundamentalsatz der Algebra hat  $A$  über  $\mathbb{C}$  mindestens einen Eigenwert  $\lambda$ , mit i) ist dieser reell. Sei  $v_1$  ein normierter Eigenvektor zu  $\lambda$ . Wir ergänzen mit  $v_2, \dots, v_{n+1}$  zu einer Orthonormalbasis  $\hat{Q} \in O(n+1)$  von  $\mathbb{R}^{n+1}$ . Dann ist

$$\langle v_1, Av_1 \rangle = \langle Av_1, v_1 \rangle = \lambda \langle v_1, v_1 \rangle = \lambda$$

und für  $i, j \in \{2, \dots, n+1\}$

$$\langle v_1, Av_i \rangle = \langle Av_1, v_i \rangle = \lambda \langle v_1, v_i \rangle = 0,$$

$$\langle v_i, Av_1 \rangle = \bar{\lambda} \langle v_1, v_i \rangle = 0,$$

$$\langle v_i, Av_j \rangle = \langle Av_i, v_j \rangle = \overline{\langle v_j, Av_i \rangle} = \langle v_j, Av_i \rangle.$$

Somit erhalten wir

$$\widehat{Q}A\widehat{Q}^T = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & & & \\ \vdots & \tilde{A} & & \\ 0 & & & \end{pmatrix}$$

mit  $\tilde{A} \in \text{Sym}(n)$ . Nach Induktionsvoraussetzung gibt es zu  $\tilde{A}$  eine Orthonormalbasis  $\tilde{Q} = (\tilde{v}_1 | \dots | \tilde{v}_n)$ , sodass  $D = \tilde{Q}\tilde{A}\tilde{Q}^T$  Diagonalgestalt hat. Wählen wir also  $Q = (v_1 | \tilde{v}_1 | \dots | \tilde{v}_n)$ , so gilt

$$QAQ^T = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & & & \\ \vdots & D & & \\ 0 & & & \end{pmatrix} \in \text{Sym}(n+1).$$

v)-vii) folgen direkt aus der Isotropie von Spur, Determinante und Norm. ■

### 2.3. Tangentialräume

**Lemma 2.3.1.** *Der Raum der schiefsymmetrischen Matrizen  $\mathfrak{so}(n)$  ist der Tangentialraum des Raums der Drehmatrizen  $SO(n)$  im Punkt  $\mathbb{1}$ . Oder gleichbedeutend: Jeder differenzierbare Weg, welcher gänzlich in  $SO(n)$  und durch  $\mathbb{1}$  verläuft, hat in der Identität seine Tangente in  $\mathfrak{so}(n)$ . Das heißt: Sei  $\gamma: (-1, 1) \rightarrow SO(n)$  eine differenzierbare Funktion mit  $\gamma(0) = \mathbb{1}$ , dann ist  $\gamma'(0) \in \mathfrak{so}(n)$ .*

*Beweis.* Sei  $f: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  mit  $f(X) = X^T X$ . Dann ist für  $H \in \mathbb{R}^{n \times n}$

$$f(\mathbb{1} + H) = \mathbb{1}^T \mathbb{1} + H^T \mathbb{1} + \mathbb{1}^T H + H^T H \quad \text{mit} \quad \left\| \frac{H^T H}{\|H\|} \right\| \leq \|H^T\| \rightarrow 0.$$

also  $Df[\mathbb{1}] \cdot H = H^T + H$ . Zudem gilt  $(f \circ \gamma)(t) = \mathbb{1}$  für alle  $t \in (-1, 1)$ . Mittels der Kettenregel erhalten wir nun

$$0 = \frac{d}{dt}(f \circ \gamma)(t) = Df[\gamma(0)] \cdot \gamma'(0) = \gamma'(0)^T + \gamma'(0),$$

somit  $\gamma'(0)^T = -\gamma'(0)$ . ■

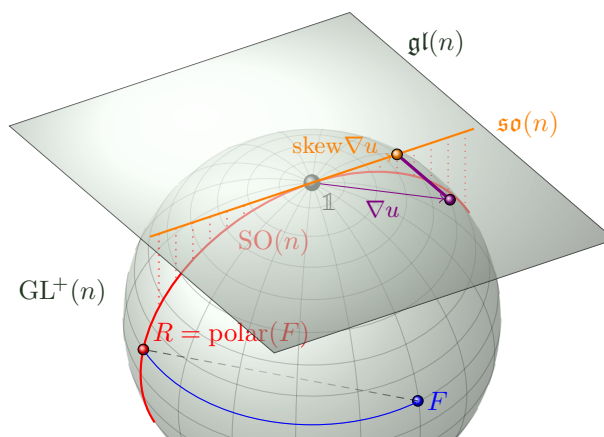


Abbildung 2.1.:  $\mathfrak{so}(n)$  als Tangentialraum von  $SO(n)$

## 2. Matrix-Analysis

**Folgerung 2.3.2.** Jedes  $Q \in \text{SO}(n)$  lässt sich darstellen als

$$Q = \mathbb{1} + W + \mathcal{O}(\|Q - \mathbb{1}\|^2) \quad \text{mit } W \in \mathfrak{so}(n). \quad (2.3.1)$$

**Lemma 2.3.3.** Der Raum der spurfreien Matrizen  $\mathfrak{sl}(n)$  ist der Tangentialraum des Raums der Matrizen mit Determinante eins  $\text{SL}(n)$  im Punkt  $\mathbb{1}$ . Oder gleichbedeutend: Jeder differenzierbare Weg, welcher gänzlich in  $\text{SL}(n)$  und durch  $\mathbb{1}$  verläuft, hat in der Identität seine Tangente in  $\mathfrak{sl}(n)$ . Das heißt: Sei  $\gamma: (-1, 1) \rightarrow \text{SL}(n)$  eine differenzierbare Funktion mit  $\gamma(0) = \mathbb{1}$ , dann ist  $\gamma'(0) \in \mathfrak{sl}(n)$ .

*Beweis.* Für  $x, y \in \mathbb{R}^n$  definieren wir das dyadische Produkt als  $x \otimes y := xy^T$  (hierbei fassen wir  $x$  bzw.  $y^T$  als einspaltige bzw. einzeilige Matrizen auf).

Für  $i \neq j$  ist  $\det(\mathbb{1} + h(e_i \otimes e_j)) = 1 = \det(\mathbb{1}) + 0 \cdot h + 0$ , somit gilt für die Ableitung an der Stelle  $\mathbb{1}$  in  $e_i \otimes e_j$ -Richtung  $D \det(\mathbb{1}).(e_i \otimes e_j) = 0$ .

Für  $i = j$  ist  $\det(\mathbb{1} + h(e_i \otimes e_j)) = 1 + h = \det(\mathbb{1}) + 1 \cdot h + 0$ , somit gilt für die Ableitung an der Stelle  $\mathbb{1}$  in  $e_i \otimes e_j$ -Richtung  $D \det(\mathbb{1}).(e_i \otimes e_j) = 1$ .

Die Determinantenabbildung  $\det: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  ist als Polynomfunktion differenzierbar. Wir erhalten somit

$$D \det[\mathbb{1}].H = D \det[\mathbb{1}]. \sum_{i,j=1}^n h_{ij}(e_i \otimes e_j) = \sum_{i,j=1}^n h_{ij} D \det[\mathbb{1}].(e_i \otimes e_j) = \sum_{i=1}^n h_{ii} \cdot 1 = \text{tr}(H) = \langle \mathbb{1}, H \rangle.$$

Sei nun  $f: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  mit  $f(x) = \det(X)$ . Es gilt somit  $(f \circ g)(t) = 1$  für alle  $t \in (-1, 1)$ . Mittels der Kettenregel erhalten wir nun

$$0 = \frac{d}{dt}(f \circ \gamma)(t) = Df[\gamma(0)]. D\gamma[0] = \text{tr}(\gamma'(0)). \quad \blacksquare$$

**Folgerung 2.3.4.** Jedes  $X \in \text{SL}(n)$  lässt sich darstellen als

$$X = \mathbb{1} + A + \mathcal{O}(\|X - \mathbb{1}\|^2) \quad \text{mit } A \in \mathfrak{sl}(n). \quad (2.3.2)$$

## 2.4. Polarzerlegung

**Definition 2.4.1.** Sei  $F \in \text{GL}^+(n)$ . Wir bezeichnen als *linke Polarzerlegung* von  $F$  die multiplikative Zerlegung

$$F = R \cdot U \quad \text{mit } R \in \text{SO}(n) \text{ und } U \in \text{Sym}^+(n), \quad (2.4.1)$$

und als *rechte Polarzerlegung* von  $F$  die multiplikative Zerlegung

$$F = V \cdot R \quad \text{mit } R \in \text{SO}(n) \text{ und } V \in \text{Sym}^+(n). \quad (2.4.2)$$

Die Eigenwerte  $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$  von  $U$  und  $V$  nennen wir die Singulärwerte von  $F$ .

**Satz 2.4.2** (Existenz und Eindeutigkeit der Polarzerlegung). *Zu jedem  $F \in \text{GL}^+(n)$  existiert eindeutig die linke und die rechte Polarzerlegung. Der orthogonale Faktor  $R \in \text{SO}(n)$  in beiden ist identisch, zudem gilt  $U = \sqrt{F^T F}$  und  $V = \sqrt{F F^T}$ . Auch haben  $U$  und  $V$  dieselben Eigenwerte und sind durch den Basiswechsel  $U = R^T V R$  ineinander überführbar.*

*Beweis.* Zur Existenz der Zerlegung: Seien  $C := F^T F$  und  $B := F F^T$ . Dann ist  $\det C = \det B = (\det F)^2 > 0$  und  $C^T = (F^T F)^T = F^T F = C$ , bzw.  $B^T = (F F^T)^T = F F^T = B$ , also  $C, B \in \text{Sym}^+(n)$  und wir können  $U, V \in \text{Sym}^+(n)$  definieren durch  $U := \sqrt{C}$  und  $V := \sqrt{B}$  und erhalten zudem  $\det U = \det V = \det F$ . Setzen wir zudem  $R := F U^{-1}$ , dann ist  $\det R = \frac{\det F}{\det U} = 1$  und  $R^T R = (F U^{-1})^T (F U^{-1}) = U^{-1} F^T F U^{-1} = U^{-1} U^2 U^{-1} = \mathbb{1}$ , also  $R \in \text{SO}(n)$ . Aufgrund von  $R U = V R$  erhalten wir  $U = R^T V R$  und damit haben  $U$  und  $V$  dieselben Eigenwerte.

Zur Eindeutigkeit der Zerlegung: Seien nun zudem  $\tilde{R} \in \text{SO}(n)$ ,  $\tilde{U} \in \text{Sym}^+(n)$  mit  $F = \tilde{R} \cdot \tilde{U}$  gegeben, dann gilt  $U^2 = F^T F = \tilde{U}^2$  und damit  $U = \tilde{U}$ . Da  $\tilde{R}$  eindeutig über  $\tilde{U}$  festgelegt ist, gilt  $\tilde{R} = R$ . Der Beweis zur Eindeutigkeit von  $V$  erfolgt analog.  $\blacksquare$

## 2.5. Isotropie und Koaxialität

**Definition 2.5.1.** Wir bezeichnen eine Menge  $\mathcal{D} \subseteq \mathbb{R}^n$  als *isotrop*, wenn  $Qx \in \mathcal{D}$  für alle  $x \in \mathcal{D}$  und alle  $Q \in O(n)$  gilt.

Wir bezeichnen eine Menge  $\mathfrak{M} \subseteq \mathbb{R}^{n \times n}$  als *isotrop*, wenn  $Q^T X Q \in \mathfrak{M}$  für alle  $X \in \mathfrak{M}$  und alle  $Q \in O(n)$  gilt, d.h. wenn  $\mathfrak{M}$  gegenüber orthogonalen Basistransformationen abgeschlossen ist.

Sei  $\mathfrak{M} \subseteq \mathbb{R}^{n \times n}$  isotrop. Wir bezeichnen eine Funktion  $\psi: \mathfrak{M} \rightarrow \mathbb{R}$  als *isotrop*, wenn  $\psi(Q^T X Q) = \psi(X)$  für alle  $X \in \mathfrak{M}$  und alle  $Q \in O(n)$  gilt, d.h. wenn  $\psi$  invariant gegenüber orthogonalen Basistransformationen ist.

Sei  $\mathfrak{M} \subseteq \mathbb{R}^{n \times n}$  isotrop. Wir bezeichnen eine Tensorfunktion  $\Phi: \mathfrak{M} \rightarrow \mathbb{R}^{n \times n}$  als *isotrop*, wenn  $\Phi(Q^T X Q) = Q^T \Phi(X) Q$  für alle  $X \in \mathfrak{M}$  und alle  $Q \in O(n)$  gilt.

Sei  $\mathcal{D} \subseteq \mathbb{R}^n$  isotrop. Wir bezeichnen eine Funktion  $\varphi: \mathcal{D} \rightarrow \mathbb{R}^n$  als *isotrop*, wenn für alle  $x \in \mathcal{D}$  mit  $(y_1, \dots, y_n) = \varphi(x_1, \dots, x_n)$  und jede Permutation  $(i_1, \dots, i_n)$  von  $(1, \dots, n)$  auch  $(y_{i_1}, \dots, y_{i_n}) = \varphi(x_{i_1}, \dots, x_{i_n})$  gilt, d.h. wenn jede Vertauschung von Komponenten des Arguments zu den gleichen Vertauschungen in den Komponenten des Bildes unter  $\varphi$  führt.

**Proposition 2.5.2.** Sei  $\mathcal{D} \subseteq \mathbb{R}^n$  isotrop. Eine Funktion  $\varphi: \mathcal{D} \rightarrow \mathbb{R}^n$  ist genau dann isotrop, wenn es eine isotrope Tensorfunktion  $\Phi: \text{Sym}(n) \rightarrow \text{Sym}(n)$  gibt, sodass für alle  $(x_1, \dots, x_n) \in \mathbb{R}^n$  mit  $(y_1, \dots, y_n) = \varphi(x_1, \dots, x_n)$  gilt

$$\Phi(\text{diag}(x_1, \dots, x_n)) = \text{diag}(y_1, \dots, y_n). \quad (2.5.1)$$

*Beweis.* Wir bezeichnen mit

$$\widetilde{SO}(n) := \{(e_{i_1} | \dots | e_{i_n})^T \mid (i_1, \dots, i_n) \text{ ist Permutation von } (1, \dots, n)\} \subseteq O(n)$$

die Menge aller rechtwinkligen Drehungen auf  $\mathbb{R}^{n \times n}$  und mit

$$\text{Diag}(n) := \{\text{diag}(x_1, \dots, x_n) \mid (x_1, \dots, x_n) \in \mathbb{R}^n\}$$

die Menge aller  $n \times n$  Diagonalmatrizen.

Sei  $(i_1, \dots, i_n)$  eine beliebige Permutation von  $(1, \dots, n)$ . Zudem sei  $(x_1, \dots, x_n) \in \mathbb{R}^n$  und  $(y_1, \dots, y_n) := \varphi(x_1, \dots, x_n)$ . Dann gilt für  $Q := (e_{i_1} | \dots | e_{i_n})^T \in \widetilde{SO}(n)$

$$\begin{aligned} \Phi(Q \text{diag}(x_1, \dots, x_n) Q^T) &= \Phi(\text{diag}(x_{i_1}, \dots, x_{i_n})) \\ &= \text{diag}(\varphi_1(x_{i_1}, \dots, x_{i_n}), \dots, \varphi_n(x_{i_1}, \dots, x_{i_n})) \end{aligned}$$

und

$$\begin{aligned} Q \Phi(\text{diag}(x_1, \dots, x_n)) Q^T &= Q \text{diag}(\varphi_1(x_1, \dots, x_n), \dots, \varphi_n(x_1, \dots, x_n)) Q^T \\ &= \text{diag}(\varphi_{i_1}(x_1, \dots, x_n), \dots, \varphi_{i_n}(x_1, \dots, x_n)) = \text{diag}(y_{i_1}, \dots, y_{i_n}) \end{aligned}$$

Angenommen, es sei  $\Phi: \text{Sym}(n) \rightarrow \text{Sym}(n)$  eine isotrope Tensorfunktion mit

$$\Phi(\text{diag}(x_1, \dots, x_n)) = \text{diag}(y_1, \dots, y_n)$$

für alle  $(x_1, \dots, x_n) \in \mathbb{R}^n$  und  $(y_1, \dots, y_n) := \varphi(x_1, \dots, x_n)$ . Dann folgt für jede Permutation  $(i_1, \dots, i_n)$  von  $(1, \dots, n)$

$$\text{diag}(\varphi_1(x_{i_1}, \dots, x_{i_n}), \dots, \varphi_n(x_{i_1}, \dots, x_{i_n})) = \text{diag}(y_{i_1}, \dots, y_{i_n}),$$

und damit  $\varphi(x_{i_1}, \dots, x_{i_n}) = (y_{i_1}, \dots, y_{i_n})$  für alle  $(x_1, \dots, x_n) \in \mathbb{R}^n$  und  $(y_1, \dots, y_n) := \varphi(x_1, \dots, x_n)$ .

Andersherum angenommen, es gelte

$$\varphi(x_{i_1}, \dots, x_{i_n}) = (y_{i_1}, \dots, y_{i_n})$$



## 2. Matrix-Analysis

für alle  $(x_1, \dots, x_n) \in \mathbb{R}^n$  und  $(y_1, \dots, y_n) := \varphi(x_1, \dots, x_n)$  und für jede Permutation  $(i_1, \dots, i_n)$  von  $(1, \dots, n)$ . Dann gilt für jedes  $Q \in \widetilde{SO}(n)$  und alle  $(x_1, \dots, x_n) \in \mathbb{R}^n$

$$\Phi(Q \operatorname{diag}(x_1, \dots, x_n) Q^T) = Q \Phi(\operatorname{diag}(x_1, \dots, x_n)) Q^T,$$

also für alle  $D \in \operatorname{Diag}(n)$  und alle  $Q \in \widetilde{SO}(n)$

$$\Phi(Q D Q^T) = Q \Phi(D) Q^T,$$

Damit ist die Abbildungsvorschrift von  $\Phi$  genau auf allen Diagonalmatrizen festgelegt. Wir erweitern diese auf Argumente aus  $\operatorname{Sym}(n)$  und alle  $Q \in O(n)$  zu einer isotropen Funktion.  $\blacksquare$

**Definition 2.5.3.** Gegeben seien zwei symmetrische Matrizen  $A, B \in \operatorname{Sym}(n)$ . Wir nennen  $A$  und  $B$  *simultan* oder auch *gemeinsam diagonalisierbar*, wenn es ein  $Q \in O(n)$  gibt, sodass  $Q$  Eigenvektorbasis von  $A$  und  $B$  ist, d.h.

$$A = Q^T \operatorname{diag}(\lambda_1, \dots, \lambda_n) Q \quad \text{und} \quad B = Q^T \operatorname{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_n) Q. \quad (2.5.2)$$

Hierbei sind  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Eigenwerte von  $A$  und  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n \in \mathbb{R}$  die Eigenwerte von  $B$ .

**Definition 2.5.4.** Gegeben seien zwei symmetrische Matrizen  $A, B \in \operatorname{Sym}(n)$ . Ist jeder Eigenvektor von  $A$  auch ein Eigenvektor von  $B$ , so nennen wir  $A$  *koaxial* zu  $B$ . Ist zudem auch  $B$  koaxial zu  $A$ , d.h. beide haben genau die gleichen Eigenvektoren, so nennen wir  $A$  und  $B$  *bi-koaxial* zueinander.

**Bemerkung 2.5.5.** Die Matrix  $A$  ist genau dann koaxial zu  $B$ , wenn jede orthogonale Eigenvektorbasis  $Q \in O(n)$  von  $A$  auch eine Eigenvektorbasis von  $B$  ist.

**Definition 2.5.6.** Sei  $\mathcal{U} \subseteq \operatorname{Sym}(n)$ . Wir bezeichnen eine Funktion  $\Phi: \mathcal{U} \rightarrow \operatorname{Sym}(n)$  als *koaxial*, wenn für jedes  $X \in \mathcal{U}$  gilt, dass  $X$  koaxial zu  $\Phi(X)$  ist und als *bi-koaxial*, wenn für jedes  $X \in \mathcal{U}$  gilt, dass  $X$  und  $\Phi(X)$  bi-koaxial sind.

**Definition 2.5.7.** Sei  $X \in \operatorname{Sym}(n)$ . Wir bezeichnen mit  $I_k(X) := e_k(\lambda_1, \dots, \lambda_n)$  die *k-te Invariante* von  $X$ . Hierbei sind  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Eigenwerte von  $X$  und  $e_k$  das *k-te elementarsymmetrische Polynom*.

**Bemerkung 2.5.8.** Die erste Invariante einer Matrix ist stets deren Spur, d.h.  $I_1(X) = \operatorname{tr}(X)$ , die  $n-1$ -te Invariante ist stets die Spur ihres Kofaktors, d.h.  $I_{n-1}(X) = \operatorname{tr}(\operatorname{Cof} X)$ , die  $n$ -te Invariante ist stets deren Determinante, d.h.  $I_n(X) = \det X$ . Speziell für  $n = 3$  erhalten wir

$$I_1 = \operatorname{tr}(X), \quad I_2 = \operatorname{tr}(\operatorname{Cof} X) = \frac{1}{2} \operatorname{tr}(X)^2 - \frac{1}{2} \operatorname{tr}(X^2), \quad I_3 = \det X. \quad (2.5.3)$$

**Bemerkung 2.5.9.** Die Invarianten erben von den elementarsymmetrischen Polynomen ihre Differenzierbarkeit: Sie sind differenzierbare Funktionen in den Eigenwerten von  $X$  mit partiellen Ableitungen

$$\frac{\partial I_k}{\partial \lambda_i}(z) = e_{k-1}(\lambda_1, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_n). \quad (2.5.4)$$

**Proposition 2.5.10.** Sei  $\mathfrak{S} \subseteq \operatorname{Sym}(n)$  und  $P: \mathfrak{S} \rightarrow \operatorname{Sym}(n)$  ein Polynom. Dann ist  $P$  koaxial.

*Beweis.* Sei  $v$  ein Eigenvektor von  $X$  zum Eigenwert  $a$ . Leicht lässt sich zeigen, dass dann auch  $X^n v = a^n v$  gilt.

Wir beweisen die Aussage per Induktion über den Polynomgrad. Für  $n = 0$  gilt  $P_0(X)v = \lambda \cdot \mathbb{1}v = \lambda v$ . Angenommen, es gibt ein Polynom  $P_n: \mathbb{R} \rightarrow \mathbb{R}$  mit  $P_n(X)v = P_n(a)v$ , dann ist

$$(\lambda X^{n+1} + P_n(X))v = \lambda X^{n+1}v + P_n(X)v = \lambda a^{n+1}v + P_n(a)v = P_{n+1}(a)v,$$

mit  $P_{n+1}(a) := \lambda a^{n+1} + P_n(a)$ .  $\blacksquare$

**Proposition 2.5.11.** Sei  $\mathfrak{S} \subseteq \operatorname{Sym}(n)$  eine isotrope Menge und  $\Phi: \mathfrak{S} \rightarrow \operatorname{Sym}(n)$  eine koaxiale Funktion. Dann ist für jede Diagonalmatrix  $D \in \mathfrak{S}$  auch  $\Phi(D)$  eine Diagonalmatrix.

## 2. Matrix-Analysis

*Beweis.* Ist  $D$  Diagonalmatrix, so sind die Einheitsnormalenvektoren  $e_1, \dots, e_n$  jeweils ein Eigenvektor von  $D$ , somit gibt es, da  $\Phi$  koaxial,  $b_1, \dots, b_n \in \mathbb{R}$  mit  $\Phi(D)e_i = b_i e_i$  für alle  $i \in \{1, \dots, n\}$ , woraus  $\Phi(D) = \text{diag}(b_1, \dots, b_n)$  folgt. ■

Das Hauptresultat der folgenden Proposition ist, dass **jede isotrope Funktion** auch **koaxial** ist. Die Umkehrung gilt offenkundig nicht, wie das Gegenbeispiel  $\Phi: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  mit  $\Phi(X) = x_{11} \cdot \mathbb{1}$  zeigt.<sup>6</sup>

**Proposition 2.5.12.** *Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge und  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  eine isotrope Funktion. Dann ist  $\Phi$  auch koaxial, d.h. sei  $X = Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^T$ , dann ist  $\Phi(X) = Q \text{diag}(\mu_1, \dots, \mu_n) Q^T$ . Hierbei sind  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Eigenwerte von  $X$  und  $\mu_1, \dots, \mu_n \in \mathbb{R}$  die Eigenwerte von  $\Phi(X)$ . Weiterhin ist  $Q \in O(n)$ , bestehend aus den Eigenvektoren jeweils zu  $\lambda_1, \dots, \lambda_n$  und  $\mu_1, \dots, \mu_n$  als Spalten.*

*Insbesondere hängen die  $\mu_1, \dots, \mu_n$  nur von  $\lambda_1, \dots, \lambda_n$  ab, d.h. es existiert eine isotrope Funktion  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$  mit  $(\lambda_1, \dots, \lambda_n) \mapsto (\mu_1, \dots, \mu_n)$  mit*

$$\Phi(Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^T) = Q \text{diag}(\varphi(\lambda_1, \dots, \lambda_n)) Q^T = Q \text{diag}(\mu_1, \dots, \mu_n) Q^T. \quad (2.5.5)$$

*Beweis.* Zur Koaxialität:

Sei  $v$  ein Eigenvektor von  $X$  und  $Q \in O(n)$  sei die Spiegelung an der (Hyper-)ebene, auf der  $v$  senkrecht steht, d.h.

$$Q.v = -v \quad \text{und} \quad Q.x = x \quad \text{für alle } x \in \mathbb{R}^n \text{ mit } \langle v, x \rangle = 0.$$

Dann ist  $QXQ^T = X$  und wir erhalten mit der Isotropie von  $\Phi$

$$Q\Phi(X)Q^T = \Phi(QXQ^T) = \Phi(X),$$

also

$$Q\Phi(X) = \Phi(X)Q$$

und damit

$$Q\Phi(X).v = \Phi(X)Q.v = -\Phi(X).v.$$

Somit steht  $\Phi(X).v$  ebenso wie  $v$  senkrecht auf  $Q$ . Dies kann nur sein, wenn  $\Phi(X).v$  ein Vielfaches von  $v$  ist. Also ist  $v$  ein Eigenvektor von  $\Phi(X)$ .

Siehe auch Ogden [62, Theorem 4.2.4, S. 193].

Zur Existenz von  $\varphi$ :

Sei  $X \in \mathfrak{S}$ . Als symmetrische Matrix ist  $X$  diagonalisierbar. Seien dazu  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die (nicht notwendigerweise unterschiedlichen) Eigenwerte von  $X$  und  $Q \in O(n)$ , sodass

$$X = Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^T$$

gilt. Da  $\Phi$  isotrop, ist  $\Phi(X) = Q \Phi(\text{diag}(\lambda_1, \dots, \lambda_n)) Q^T$ . Mit der aus dem ersten Beweisteil gefolgerten Koaxialität und Proposition 2.5.11 folgt, dass  $\Phi(\text{diag}(\lambda_1, \dots, \lambda_n))$  eine Diagonalmatrix ist. Bezeichnen wir die Einträge mit  $\mu_1, \dots, \mu_n$ , so erhalten wir

$$\Phi(X) = Q \Phi(\text{diag}(\lambda_1, \dots, \lambda_n)) Q^T = Q \text{diag}(\mu_1, \dots, \mu_n) Q^T.$$

Die rechte Gleichheit bleibt bei allen  $Q \in O(n)$  erhalten, was die Existenz von  $\varphi$  mit  $\varphi(\lambda_1, \dots, \lambda_n) = (\mu_1, \dots, \mu_n)$  sichert, die mit Proposition 2.5.2 isotrop ist. ■

**Folgerung 2.5.13.** *Da für jede isotrope Tensorfunktion  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  jede Basis aus Eigenvektoren  $Q \in O(n)$  von  $X$  auch eine Basis von  $\Phi(X)$  ist, lassen sich  $X$  und  $\Phi(X)$  gemeinsam mit diesem*

---

<sup>6</sup>Da die Identität  $\mathbb{1}$  ganz  $\mathbb{R}^{n \times n}$  als Eigenraum hat, ist die Koaxialität klar. Aber  $\Phi$  ist nicht isotrop, wie z.B. für  $n = 2$  zu sehen:  $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \Phi \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \Phi \left( \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) \neq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Phi \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ .

Siehe hierzu auch Bemerkung 2.5.29 auf Seite 34.

## 2. Matrix-Analysis

$Q$  diagonalisieren (und kommutieren deshalb auch). Die Umkehrung gilt allerdings nicht. Hierfür betrachten wir  $\Phi: \text{Sym}^+(2) \rightarrow \text{Sym}(n)$  mit  $X \mapsto X^2$  für  $X = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ , also  $\Phi(X) = \mathbb{1}$ . Es ist für  $Q = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in \text{O}(2)$  dann  $Q\Phi(X)Q^T = \mathbb{1}$  und  $QXQ^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ .

**Proposition 2.5.14.** Ein Vektor  $v = (v_1, \dots, v_n) \in \mathbb{R}^n \setminus \{0\}$  ist genau dann Eigenvektor einer Diagonalmatrix  $D = \text{diag}(d_1, \dots, d_n) \in \mathbb{R}^{n \times n}$  zum Eigenwert  $a \in \mathbb{R}$ , wenn für alle  $i \in \{1, \dots, n\}$  gilt:  $d_i = a$  oder  $v_i = 0$ . Oder anders ausgedrückt, wenn  $v$  nur an den Komponenten von Null verschieden ist, die auf der Diagonale von  $D$  gleich  $a$  sind.

*Beweis.* Es ist genau dann  $v$  Eigenvektor von  $D$  zum Eigenwert  $a$ , wenn  $Dv = av$  gilt. Ausmultipliziert bedeutet dies:

$$\begin{pmatrix} d_1 v_1 \\ d_2 v_2 \\ \vdots \\ d_n v_n \end{pmatrix} = \begin{pmatrix} a v_1 \\ a v_2 \\ \vdots \\ a v_n \end{pmatrix},$$

was gleichbedeutend mit  $d_i v_i = a v_i$ , also  $d_i = a$  oder  $v_i = 0$ , für alle  $i \in \{1, \dots, n\}$  ist. ■

**Lemma 2.5.15.** Seien  $A, B \in \text{Sym}(n)$  gemeinsam diagonalisierbar mit  $Q \in \text{O}(n)$

$$A = Q \text{diag}(a_1, \dots, a_n) Q^T \quad \text{und} \quad B = Q \text{diag}(b_1, \dots, b_n) Q^T. \quad (2.5.6)$$

Dann ist  $A$  genau dann koaxial zu  $B$ , wenn  $a_i = a_j \Rightarrow b_i = b_j$  für alle  $i, j \in \{1, \dots, n\}$  gilt. Hierbei sind  $a_1, \dots, a_n \in \mathbb{R}$  die Eigenwerte von  $A$  und  $b_1, \dots, b_n \in \mathbb{R}$  die Eigenwerte von  $B$ .

*Beweis.* Angenommen  $A$  ist koaxial zu  $B$  und sei  $i, j \in \{1, \dots, n\}$ . Für  $i = j$  ist die Aussage trivial, deshalb sei  $i \neq j$ . Wähle  $v$  als Summe des  $i$ -ten und  $j$ -ten Einheitsvektors, d.h.  $v = e_i + e_j$ . Nach Proposition 2.5 ist  $v$  nun genau dann ein Eigenvektor einer Diagonalmatrix  $\text{diag}(d_1, \dots, d_n)$ , wenn  $d_i = d_j$  gilt. Ist also  $a := a_i = a_j$ , dann gilt somit

$$\text{diag}(a_1, \dots, a_n) v = a v$$

und damit

$$Q \text{diag}(a_1, \dots, a_n) Q^T Q v = a Q v,$$

somit ist  $Qv$  Eigenvektor von  $A$  zum Eigenwert  $a$ . Aufgrund der Koaxialität ist  $Qv$  auch Eigenvektor von  $B$ , damit gibt es ein  $b$  mit

$$Q \text{diag}(b_1, \dots, b_n) Q^T Q v = b Q v,$$

also

$$\text{diag}(b_1, \dots, b_n) v = b v$$

und es muss  $b_i = b_j = b$  gelten.

Angenommen, es gilt  $a_i = a_j \Rightarrow b_i = b_j$  für alle  $i, j \in \{1, \dots, n\}$ . Sei  $Q \in \text{O}(n)$  so gewählt, dass  $A = Q \text{diag}(a_1, \dots, a_n) Q^T$  und sei  $v \in \mathbb{R}^n$  ein Eigenvektor von  $A$  zum Eigenwert  $a$ , d.h. es ist  $Av = av$  und damit

$$\text{diag}(a_1, \dots, a_n) Q^T v = a Q^T v.$$

Somit ist  $v' = Q^T v$  ein Eigenvektor von  $\text{diag}(a_1, \dots, a_n)$  zum Eigenwert  $a$ . Sei  $I$  die Indexmenge der Komponenten von  $v'$ , die ungleich Null sind. Ist  $I$  einelementig, so ist  $v'$  trivialerweise auch Eigenvektor von  $\text{diag}(b_1, \dots, b_n)$ . Hat  $I$  mehr Elemente, so gilt nach Proposition genau  $a_i = a$  für alle  $i \in I$ . Nach Voraussetzung gibt es ein  $b$  mit  $b_i = b$  für alle  $i \in I$  und  $v'$  ist Eigenvektor von  $\text{diag}(b_1, \dots, b_n)$ , d.h.  $\text{diag}(b_1, \dots, b_n) Q^T v = b Q^T v$ , was

$$Q \text{diag}(b_1, \dots, b_n) Q^T v = b v,$$

also  $Bv = bv$  impliziert. ■

## 2. Matrix-Analysis

**Folgerung 2.5.16.** Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge und  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  eine isotrope Funktion. Dann ist  $\Phi$  bi-koaxial genau dann, wenn für alle  $X \in \mathfrak{S}$  die folgende Implikation gilt:

$$a_i \neq a_j \Rightarrow b_i \neq b_j \quad \text{für alle } i, j \in \{1, \dots, n\}, \quad (2.5.7)$$

wobei  $a_1, \dots, a_n \in \mathbb{R}$  die Eigenwerte von  $X$  und  $b_1, \dots, b_n \in \mathbb{R}$  die Eigenwerte von  $\Phi(X)$  bezeichnen. Hierbei gilt

$$\text{diag}(b_1, \dots, b_n) = \Phi(\text{diag}(a_1, \dots, a_n)), \quad (2.5.8)$$

d.h.  $b_1, \dots, b_n$  sind die zugehörigen Eigenwerte von  $\Phi(X)$  zu  $a_1, \dots, a_n$  von  $X$ .

*Beweis.* Sei  $X \in \mathfrak{S}$ . Da  $\Phi$  isotrop ist, ist nach Proposition 2.5.12 auch  $X$  koaxial zu  $\Phi(X)$ , sodass  $X$  und  $\Phi(X)$  genau dann bi-koaxial sind, wenn  $\Phi(X)$  koaxial zu  $X$  ist. Zudem sind  $X$  und  $\Phi(X)$  gemeinsam diagonalisierbar. Mit Proposition 2.5.15 gilt für jedes  $X \in \mathfrak{S}$  die Aussage:  $\Phi(X)$  ist genau dann koaxial zu  $X$ , wenn  $b_i = b_j \Rightarrow a_i = a_j$  für alle  $i, j \in \{1, \dots, n\}$  gilt. ■

**Lemma 2.5.17.** Seien  $A, B \in \text{Sym}(n)$  und  $A$  koaxial zu  $B$ . Dann gibt es zu jedem Eigenwert  $a$  von  $A$  einen Eigenwert  $b$  von  $B$ , sodass jeder Eigenvektor  $v$  von  $A$  zum Eigenwert  $a$  auch Eigenvektor von  $B$  zum Eigenwert  $b$  ist, d.h.

$$Av = av \Rightarrow Bv = bv \quad \text{für alle } v \in \mathbb{R}^n. \quad (2.5.9)$$

*Beweis.* Angenommen, es gäbe zwei Eigenvektoren  $v_1, v_2$  von  $A$  zum Eigenwert  $a$ , die Eigenvektoren von  $B$  mit unterschiedlichen Eigenwerten  $b_1, b_2$  sind.

Es gilt  $A(v_1 + v_2) = Av_1 + Av_2 = av_1 + av_2 = a(v_1 + v_2)$ , so ist  $v_1 + v_2$  Eigenvektor von  $A$  und damit auch von  $B$ , es gibt also ein  $b \in \mathbb{R}$  mit  $B(v_1 + v_2) = b(v_1 + v_2)$ .

Im Falle der linearen Abhängigkeit von  $v_1$  und  $v_2$ , d.h. es gibt  $s \in \mathbb{R}$  mit  $sv_1 = v_2$  ist  $b_1v_1 + b_2v_2 = Bv_1 + Bv_2 = B(v_1 + v_2) = B(1 + s)v_1 = (1 + s)b_1v_1 = b_1v_1 + b_1v_2$ , also  $b_1 = b_2$ , ein Widerspruch.

Im Falle der linearen Unabhängigkeit von  $v_1$  und  $v_2$  ist  $b_1v_1 + b_2v_2 = Bv_1 + Bv_2 = B(v_1 + v_2) = b(v_1 + v_2)$ , also  $(b - b_1)v_1 + (b - b_2)v_2 = 0$  und damit  $b - b_1 = b - b_2 = 0$ , also  $b_1 = b_2$ , ein Widerspruch. ■

**Lemma 2.5.18.** Seien  $A, B \in \text{Sym}(n)$  simultan diagonalisierbare Matrizen. Dann kommutieren  $A$  und  $B$ .

*Beweis.* Seien  $A = QDQ^T$  und  $B = Q\tilde{D}Q^T$  mit Diagonalmatrizen  $D, \tilde{D}$  und  $Q \in O(n)$ , dann gilt

$$AB = QDQ^TQ\tilde{D}Q^T = QD\tilde{D}Q^T = Q\tilde{D}DQ^T = Q\tilde{D}Q^TQDQ^T = BA. \quad \blacksquare$$

**Folgerung 2.5.19.** Seien  $A, B \in \text{Sym}(n)$ . Ist  $A$  koaxial zu  $B$  oder  $B$  koaxial zu  $A$ , dann kommutieren  $A$  und  $B$ .

**Lemma 2.5.20.** Gegeben seien zwei kommutierende symmetrische Matrizen  $A, B \in \text{Sym}(n)$ . Hat  $A$  nur unterschiedliche Eigenwerte, so ist  $A$  koaxial zu  $B$ .

**Bemerkung 2.5.21.** Haben sowohl  $A$  als auch  $B$  nur unterschiedliche Eigenwerte, so ist auch  $B$  koaxial zu  $A$ , ergo sind beide bi-koaxial.

*Beweis.* Seien  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Eigenwerte von  $A$  und  $Q \in O(n)$ , sodass  $\text{diag}(\lambda_1, \dots, \lambda_n) = Q A Q^T$  gilt. Wir definieren

$$C := \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix} := Q B Q^T$$

dann gilt aufgrund der Kommutativität von  $A$  und  $B$

$$\underbrace{Q A Q^T}_{\text{diag}(\lambda_1, \dots, \lambda_n)} \underbrace{Q B Q^T}_C = Q A B Q^T = Q B A Q^T = \underbrace{Q B Q^T}_C \underbrace{Q A Q^T}_{\text{diag}(\lambda_1, \dots, \lambda_n)},$$

## 2. Matrix-Analysis

also ist

$$\begin{pmatrix} \lambda_1 c_{11} & \lambda_1 c_{12} & \dots & \lambda_1 c_{1n} \\ \lambda_2 c_{21} & \lambda_2 c_{22} & \dots & \lambda_2 c_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_n c_{n1} & \lambda_n c_{n2} & \dots & \lambda_n c_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_1 c_{11} & \lambda_2 c_{12} & \dots & \lambda_n c_{1n} \\ \lambda_1 c_{21} & \lambda_2 c_{22} & \dots & \lambda_n c_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_1 c_{n1} & \lambda_2 c_{n2} & \dots & \lambda_n c_{nn} \end{pmatrix}$$

und damit  $\lambda_i c_{ij} = \lambda_j c_{ij}$  für alle  $i, j \in \{1, \dots, n\}$ . Ist nun  $\lambda_i \neq \lambda_j$ , so folgt  $c_{ij} = 0$ , also ist auch  $C$  eine Diagonalmatrix. ■

Da Bi-Koaxialität die simultane Diagonalisierbarkeit impliziert, sind wir imstande zu folgern:

**Folgerung 2.5.22.** *Seien  $A, B \in \text{Sym}(n)$ . Sind zudem alle Eigenwerte jeweils von  $A$  und von  $B$  paarweise verschieden, so sind  $A$  und  $B$  genau dann bi-koaxial, wenn sie kommutieren.*

**Proposition 2.5.23.** *Seien  $A, B \in \text{Sym}^+(n)$  kommutierenden Matrizen und  $\mathcal{U}$  Eigenraum von  $A$  und  $u \in \mathcal{U}$ . Dann ist auch  $Bu \in \mathcal{U}$ .*

*Beweis.* Der Eigenwert zu  $u$  sei  $\lambda$ , dann gilt  $Au = \lambda u$ , aber aufgrund der Kommutativität auch

$$ABu = BAu = B\lambda u = \lambda Bu,$$

also ist auch  $Bu \in \mathcal{U}$ . ■

Wie gesehen, folgt im Allgemeinen aus der Kommutativität nicht die Bi-Koaxialität. Begnügen wir uns mit gemeinsamer Diagonalisierbarkeit, können wir feststellen:

**Lemma 2.5.24.** *Seien  $A, B \in \text{Sym}(n)$  kommutierende Matrizen. Dann lassen sich  $A$  und  $B$  gemeinsam diagonalisieren.*

*Beweis.* Der Beweis erfolgt per Induktion über  $n$ .

Für  $n = 1$  ist nichts zu zeigen.

Nehmen wir nun an, dass alle kommutierenden  $A, B \in \text{Sym}(s)$  mit  $s \leq n$  sich gemeinsam diagonalisieren lassen. Seien nun  $A, B \in \mathbb{R}^{n+1}$ . Ohne Einschränkung nehmen wir an, dass  $A$  mehr als nur einen Eigenwert besitzt, ansonsten wäre ganz  $\text{Sym}(n+1)$  Eigenraum und damit  $A = \lambda_1 \cdot \mathbb{1} = Q\lambda_1 \mathbb{1} Q^T$  für alle  $Q \in O(n+1)$  und damit auch für die Eigenvektorbasis von  $B$ . Also habe  $A$  die Eigenwerte  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$  mit zugehörigen Eigenräumen  $\mathcal{U}_1, \dots, \mathcal{U}_k$  und  $k > 1$ . Wählen wir Basen  $B_1, \dots, B_k$  für die  $\mathcal{U}_1, \dots, \mathcal{U}_k$ , so ist  $Q := (B_1, \dots, B_k) := (b_1, \dots, b_{n+1}) \in O(n+1)$  Eigenvektorbasis von  $A$ .

Wir betrachten nun ohne Einschränkung  $\mathcal{U}_i$  mit Eigenwert  $\lambda_i$  und Basis  $\mathcal{B}_i = \{b_l, b_{l+1}, \dots, b_m\}$ . Für jedes  $b \in \mathcal{B}_i$  gilt offenkundig  $Ab = \lambda_i b$  und damit  $Q^T A b = (0, \dots, \lambda_i, \dots, 0)$ , so ist  $Q^T A Q$ , wie zu erwarten, eine Diagonalmatrix. Da nach Proposition 2.5.23  $Bb$  im Eigenraum  $\mathcal{U}_i$  liegt, gibt es Koeffizienten  $\alpha_l, \dots, \alpha_m$  mit  $Bb = \alpha_l b_l + \dots + \alpha_m b_m$  und damit  $Q^T B b = (0, \dots, \alpha_l, \dots, \alpha_m, \dots, 0)$ , was dazu führt, dass  $Q^T B Q$  eine Blockmatrix ist. Jeder Block aus  $Q^T B Q$  kommutiert natürlich mit dem entsprechenden Block aus  $Q^T A Q$ , sodass sich jeweils beide gemeinsam diagonalisieren lassen. Also gibt es ein  $R \in O(n+1)$ , sodass  $R^T (Q^T B Q) R$ , wie natürlich auch  $R^T (Q^T A Q) R$  eine Diagonalmatrix ist, ergo lassen sich  $A$  und  $B$  mit  $\tilde{Q} := QR$  gemeinsam diagonalisieren. ■

**Folgerung 2.5.25.** *Seien  $A, B \in \text{Sym}(n)$ . Dann sind  $A$  und  $B$  genau dann simultan diagonalisierbar, wenn sie kommutieren.*

**Lemma 2.5.26** (Eigenschaften isotrope Funktionen). *Für eine isotrope Menge  $\mathfrak{S} \subseteq \text{Sym}(n)$  sei  $\psi: \mathfrak{S} \rightarrow \mathbb{R}$  eine reellwertige Tensorfunktion. Dann sind folgende vier Aussagen äquivalent:*

- 1)  $\psi$  ist isotrop, d.h. es ist  $\psi(Q^T X Q) = \psi(X)$  für alle  $Q \in O(n)$ .
- 2)  $\psi(X)$  hängt nur von den Eigenwerten (auch nicht von deren Reihenfolge) von  $X \in \text{Sym}(n)$  ab. D.h. wenn eine Permutation  $i_1, \dots, i_n$  mit  $\lambda_{i_k} = \tilde{\lambda}_k$  für alle  $k \in \{1, \dots, n\}$  existiert, dann folgt  $\psi(X) = \psi(Y)$ .

## 2. Matrix-Analysis

3) Es existiert eine eindeutige symmetrische Funktion  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  in den Eigenwerten, sodass  $\psi(X) = g(\lambda_1, \dots, \lambda_n)$  für alle  $X \in \mathfrak{S}$  gilt.

4) Es existiert eine eindeutige Funktion  $\Psi: T \rightarrow \mathbb{R}$  in den Invarianten, sodass  $\psi(X) = \Psi(I_1, \dots, I_n)$  für alle  $X \in \mathfrak{S}$  gilt.

Hierbei bezeichnen wir mit  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Eigenwerte von  $X$ , mit  $I_1, \dots, I_n$  die Invarianten von  $X$  und mit  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n \in \mathbb{R}$  die Eigenwerte von  $Y$ . Zudem ist  $T = \{(I_1, \dots, I_n) \mid X \in \mathfrak{S}\} \subseteq \mathbb{R}^n$  die Menge der Invariantenvektoren, die auf Matrizen aus  $\mathfrak{S}$  führen.<sup>7</sup>

*Beweis.* 1)  $\Rightarrow$  2): Seien  $X, Y \in \text{Sym}(n)$  beliebig, aber mit gemeinsamen Eigenwerten  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ . Definieren wir  $D := \text{diag}(\lambda_1, \dots, \lambda_n)$ , dann gibt es  $P, Q \in O(n)$  mit  $X = P^T D P$  und  $Y = Q^T D Q$ . Mit 1) gilt:  $\psi(X) = \psi(P^T D P) = \psi(D) = \psi(Q^T D Q) = \psi(Y)$ . Zudem ist die Reihenfolge der Eigenwerte egal, denn mit  $Q = (e_{i_1} \mid \dots \mid e_{i_n})$  gilt

$$\psi(\text{diag}(\lambda_1, \dots, \lambda_n)) = \psi(Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^T) = \psi(\text{diag}(\lambda_{i_1}, \dots, \lambda_{i_n})).$$

2)  $\Rightarrow$  3): Zur Existenz: Wir definieren  $g(\lambda_1, \dots, \lambda_n) := \psi(\text{diag}(\lambda_1, \dots, \lambda_n))$ . Wegen 2) gilt für alle  $X$  mit Eigenwerten  $\lambda_1, \dots, \lambda_n$  und jede Permutation  $i_1, \dots, i_n$  auch  $\psi(X) = g(\lambda_1, \dots, \lambda_n) = g(\lambda_{i_1}, \dots, \lambda_{i_n})$ . Zur Eindeutigkeit: Seien  $g, \tilde{g}: \mathbb{R}^n \rightarrow \mathbb{R}$  Funktionen mit  $\psi(X) = g(\lambda_1, \dots, \lambda_n)$  und  $\psi(X) = \tilde{g}(\lambda_1, \dots, \lambda_n)$  für alle  $X \in \text{Sym}(n)$ . Auf die Teilmenge aller Diagonalmatrizen  $\{\text{diag}(\lambda_1, \dots, \lambda_n) \mid \lambda_1, \dots, \lambda_n \in \mathbb{R}\} \subseteq \text{Sym}(n)$  eingeschränkt, bedeutet dies  $\psi(\text{diag}(\lambda_1, \dots, \lambda_n)) = g(\lambda_1, \dots, \lambda_n) = \tilde{g}(\lambda_1, \dots, \lambda_n)$  für alle  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ , was  $g \equiv \tilde{g}$  impliziert.

3)  $\Rightarrow$  4): Nach Folgerung 2.13.13 lässt sich jede symmetrische Funktion als Funktion in den elementarsymmetrischen Polynomen ausdrücken, die Invarianten sind definiert als die elementarsymmetrischen Polynome, ausgewertet in den Eigenwerten.

4)  $\Rightarrow$  1): Für alle  $Q \in O(n)$  hat  $X \in \mathfrak{S}$  die gleichen Invarianten<sup>8</sup> wie  $Q^T X Q$ , somit gilt  $\psi(X) = \Psi(\lambda_1, \dots, \lambda_n) = \psi(Q^T X Q)$ . ■

**Proposition 2.5.27.** Seien  $A, B \in \text{Sym}(n)$  und  $A$  koaxial zu  $B$ . Dann gibt es  $\alpha_1, \dots, \alpha_n \in \mathbb{R}$  mit  $B = \sum_{i=0}^{n-1} \alpha_i A^i$ .

*Beweis.* Seien  $a_1, \dots, a_n \in \mathbb{R}$  die (nicht notwendigerweise verschiedenen) Eigenwerte von  $A$  und  $b_1, \dots, b_n$  die (nicht notwendigerweise verschiedenen) Eigenwerte von  $B$ . Aufgrund der Koaxialität sind  $A$  und  $B$  gemeinsam diagonalisierbar. So sei  $Q \in O(n)$ , sodass

$$A = Q \text{diag}(a_1, \dots, a_n) Q^T \quad \text{und} \quad B = Q \text{diag}(b_1, \dots, b_n) Q^T$$

gilt. Nun bezeichnen wir mit  $a_{i_1}, \dots, a_{i_m}$  und  $m \leq n$  die paarweise verschiedenen Eigenwerte von  $A$ . Die Vandermonde-Matrix

$$V(a_{i_1}, \dots, a_{i_m}) := \begin{pmatrix} 1 & a_{i_1} & a_{i_1}^2 & \dots & a_{i_1}^{m-1} \\ 1 & a_{i_2} & a_{i_2}^2 & \dots & a_{i_2}^{m-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{i_m} & a_{i_m}^2 & \dots & a_{i_m}^{m-1} \end{pmatrix}$$

hat die Determinante

$$\det V(a_{i_1}, \dots, a_{i_m}) = \prod_{j < k} (a_{i_j} - a_{i_k}) \neq 0,$$

sodass das Gleichungssystem

$$\begin{pmatrix} b_{i_1} \\ \vdots \\ b_{i_m} \end{pmatrix} = V(a_{i_1}, \dots, a_{i_m}) \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_{m-1} \end{pmatrix}$$

<sup>7</sup>So hat beispielsweise die komplexe Matrix  $\text{diag}(-i, i) \notin \text{Sym}(2)$  die reellen Invarianten  $(I_1, I_2) = (0, 1) \in \mathbb{R}^2$

<sup>8</sup>Sie sind ihrem Namen nach invariant gegenüber Basistransformationen.

## 2. Matrix-Analysis

also

$$b_{i_k} = \alpha_0 + \alpha_1 a_{i_k} + \alpha_2 a_{i_k}^2 + \dots + \alpha_{m-1} a_{i_k}^{m-1} \quad \text{für alle } k \in \{1, \dots, m\}$$

eine eindeutige Lösung  $(\alpha_0, \dots, \alpha_{m-1})^T$  hat. Da  $A$  koaxial zu  $B$  gilt mit Proposition 2.5.15 für alle  $a_i = a_{i_k}$  auch  $b_i = b_{i_k}$ . So ist das Gleichungssystem äquivalent zu

$$b_i = \alpha_0 + \alpha_1 a_i + \alpha_2 a_i^2 + \dots + \alpha_{m-1} a_i^{m-1} \quad \text{für alle } i \in \{1, \dots, n\},$$

also

$$\text{diag}(b_1, \dots, b_n) = \sum_{i=0}^{n-1} \alpha_i \text{diag}(a_1, \dots, a_n)^i.$$

Beidseitiges Multiplizieren von links mit  $Q$  und von rechts mit  $Q^T$  führt dann auf  $B = \sum_{i=0}^{n-1} \alpha_i A^i$  mit eindeutiger Lösung  $\alpha_0, \dots, \alpha_{m-1} \in \mathbb{R}$  unter der Voraussetzung  $\alpha_m = \dots = \alpha_{n-1} = 0$ . Ohne diese Voraussetzung ist die Lösung womöglich nicht mehr eindeutig. ■

Kommen wir nun zu einer zentralen Eigenschaft isotroper Funktionen, die sich auch in [69] findet:

**Lemma 2.5.28** (Darstellungssatz). *Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge. Eine Funktion  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  ist genau dann isotrop, wenn es nur von den Invarianten von  $X$  abhängige Koeffizienten  $\alpha_0, \dots, \alpha_{n-1}$  gibt (d.h. sie sind Funktionen in den Invarianten), sodass gilt*

$$\Phi(X) = \sum_{i=0}^{n-1} \alpha_i X^i = \alpha_0 \cdot \mathbb{1} + \alpha_1 X + \dots + \alpha_{n-1} X^{n-1} \quad \text{für alle } X \in \mathfrak{S}. \quad (2.5.10)$$

Ist sogar  $\mathfrak{S} \subseteq \text{Sym}^+(n)$ , dann ist (äquivalent zur Darstellung (2.5.10))  $\Phi$  genau dann isotrop, wenn es nur von den Invarianten von  $X$  abhängige Koeffizienten  $\beta_{-1}, \beta_0, \dots, \beta_{n-2}$  gibt (d.h. sie sind Funktionen in den Invarianten), sodass gilt

$$\Phi(X) = \sum_{i=-1}^{n-2} \beta_i X^i = \beta_{-1} \cdot X^{-1} + \beta_0 \cdot \mathbb{1} + \dots + \beta_{n-2} X^{n-2} \quad \text{für alle } X \in \mathfrak{S}. \quad (2.5.11)$$

**Bemerkung 2.5.29.** Blume [10] bezieht sich zwar nur auf den dreidimensionalen Fall, charakterisiert durch (2.5.10), wobei  $\alpha_0, \dots, \alpha_{n-1}$  skalarwertige Funktionen in  $X$  sind, genau alle koaxialen Funktionen. Mit dieser Aussage lässt sich der Darstellungssatz auch anders formulieren: Jede koaxiale Funktion mit Darstellung (2.5.10) ist genau dann isotrop, wenn  $\alpha_0, \dots, \alpha_{n-1}: \mathfrak{S} \rightarrow \mathbb{R}$  isotrope Funktionen in  $X$  sind.<sup>9</sup>

*Beweis.* Angenommen  $\Phi$  ist isotrop. Sei  $X \in \mathfrak{S}$ . Da jede isotrope Funktion auch koaxial ist gibt es mit Proposition 2.5.27 Koeffizienten  $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{R}$ , sodass

$$\Phi(X) = \alpha_0 \cdot \mathbb{1} + \alpha_1 X + \alpha_2 X^2 + \dots + \alpha_{n-1} X^{n-1}$$

gilt. Aufgrund der Isotropie von  $\Phi$  gilt dann auch

$$\Phi(Q^T X Q) = \alpha_0 \cdot \mathbb{1} + \alpha_1 (Q^T X Q) + \alpha_2 (Q^T X Q)^2 + \dots + \alpha_{n-1} (Q^T X Q)^{n-1} \quad \text{für alle } Q \in O(n),$$

was bedeutet, dass die Koeffizienten  $\alpha_0, \dots, \alpha_{n-1}$  sich nicht ändern, wenn  $X$  durch  $Q^T X Q$  ersetzt wird. Sie sind somit Invarianten in  $A$ .

Die Äquivalenz zur Darstellung in  $\beta_{-1}, \dots, \beta_{n-2}$  offenbart sich mittels des Satzes von Cayley-Hamilton, nachdem eine Matrix  $X$  Nullstelle ihres charakteristischen Polynoms ist, es also zu jedem  $X \in \text{Sym}^+(n)$  Koeffizienten  $\gamma_{-1}, \gamma_0, \dots, \gamma_{n-2}$  gibt mit

$$0 = -X^n + \gamma_{n-2} X^{n-1} + \dots + \gamma_0 X + \gamma_{-1} \cdot \mathbb{1}.$$

<sup>9</sup>Die Äquivalenz der Darstellung in den Invarianten und der Darstellung als isotrope Funktion folgt aus Lemma 2.5.26.

## 2. Matrix-Analysis

Die Koeffizienten sind invariant gegenüber dem Ersetzen von  $X$  durch  $Q^T X Q$ , sie sind also Invarianten in  $X$ . Wir formen um:

$$X^{n-1} = \gamma_{n-2} X^{n-2} + \dots + \gamma_0 \cdot \mathbb{1} + \gamma_{-1} X^{-1}$$

und setzen ein:

$$\Phi(X) = \underbrace{(\alpha_0 + \alpha_{n-1}\gamma_0)}_{:=\beta_0} \cdot \mathbb{1} + \underbrace{(\alpha_1 + \alpha_{n-1}\gamma_1)}_{:=\beta_1} X + \dots + \underbrace{(\alpha_{n-2} + \alpha_{n-1}\gamma_{n-1})}_{:=\beta_{n-2}} X^{n-2} + \underbrace{\alpha_{n-1}\gamma_{-1}}_{:=\beta_{-1}} X^{-1}.$$

Angenommen, es gibt für jedes  $X \in \mathfrak{S}$  eine Darstellung

$$\Phi(X) = \alpha_0 \cdot \mathbb{1} + \alpha_1 X + \alpha_2 X^2 + \dots + \alpha_{n-1} X^{n-1}$$

so gilt für jede Potenz  $(QXQ^T)^k = QX^kQ^T$ , zudem sind die Koeffizienten selbst invariant gegenüber dem Ersetzen von  $X$  durch  $QXQ^T$ , somit ist  $\Phi$  offenkundig isotrop. ■

### 2.5.1. Isotrope Funktionen

**Proposition 2.5.30.** Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine offene, isotrope Menge und  $\Phi: \mathfrak{S} \rightarrow \mathbb{R}$  eine differenzierbare, isotrope Funktion, dann ist auch ihre Ableitung isotrop, d.h. es gilt

$$D\Phi(QXQ^T) = Q D\Phi(X) Q^T \quad \text{für alle } X \in \mathfrak{S} \text{ und alle } Q \in O(n). \quad (2.5.12)$$

*Beweis.* Sei  $X \in \mathfrak{S}$  und  $Q \in O(n)$ . Entwickeln wir  $W$  an der Stelle  $QXQ^T$  in Richtung  $QHQ^T$ , so erhalten wir

$$W(QXQ^T + QHQ^T) = W(QXQ^T) + \langle DW(QXQ^T), QHQ^T \rangle + \underbrace{O(\|QHQ^T\|^2)}_{= \|H\|}.$$

Entwickeln wir  $W$  an der Stelle  $X$  in Richtung  $H$ , so erhalten wir

$$W(X + H) = W(X) + \langle DW(X), H \rangle + O(\|H\|^2).$$

Die Abbildung  $W$  ist isotrop, deshalb gilt  $W(QXQ^T + QHQ^T) = W(Q(X + H)Q^T) = W(X + H)$ , was auf

$$\langle Q^T DW(QXQ^T) Q - DW(X), H \rangle = O(\|H\|^2)$$

führt. Der Ausdruck links ist linear in  $H$ , der Ausdruck rechts quadratisch in  $H$ , somit muss

$$Q^T DW(QXQ^T) Q - DW(X) = 0$$

gelten, was gleichbedeutend mit  $DW(QXQ^T) = Q DW(X) Q^T$  ist. ■

**Definition 2.5.31.** Sei  $I \subseteq \mathbb{R}$  ein reelles Intervall. Dann bezeichnen wir mit  $\text{Sym}^I(n)$  die Menge aller symmetrischen Matrizen mit Eigenwerten  $\lambda_1, \dots, \lambda_n \in I$ , also

$$\text{Sym}^I(n) := \{X \in \text{Sym}(n) \mid \lambda_1, \dots, \lambda_n \in I\}. \quad (2.5.13)$$

**Proposition 2.5.32.** Sei  $\Phi: \text{Sym}^I(n) \rightarrow \mathbb{R}$  eine differenzierbare, isotrope Funktion mit der Darstellung in den Eigenwerten  $g: I^n \rightarrow \mathbb{R}$ , d.h.  $\Phi(X) = g(\lambda_1, \dots, \lambda_n)$ . Dann ist auch  $g$  differenzierbar mit partiellen Ableitungen:

$$\frac{\partial g}{\partial \lambda_i}(\lambda_1, \dots, \lambda_n) = \frac{\partial \Phi}{\partial x_{ii}}(\text{diag}(\lambda_1, \dots, \lambda_n)). \quad (2.5.14)$$

*Beweis.* Sei  $X \in \text{Sym}^I(n)$  und  $Q \in O(n)$  derart, dass  $Q^T X Q = \text{diag}(\lambda_1, \dots, \lambda_n)$ , dann ist mit Proposition 2.5.30

$$D\Phi(\text{diag}(\lambda_1, \dots, \lambda_n)) = Q^T D\Phi[X] Q = \text{diag}\left(\frac{\partial g}{\partial \lambda_1}(\lambda_1, \dots, \lambda_n), \dots, \frac{\partial g}{\partial \lambda_n}(\lambda_1, \dots, \lambda_n)\right). \quad \blacksquare$$



## 2. Matrix-Analysis

**Proposition 2.5.33.** Sei  $I \subset \mathbb{R}$  ein Intervall und  $f: I \rightarrow \mathbb{R}$  eine reelle Funktion. Dann ist die Ableitung der primären Matrixfunktion  $T_f: \text{Sym}^1(n) \rightarrow \text{Sym}(n)$  selbstadjungiert im Frobenius-Skalarprodukt, d.h. es gilt

$$\langle DT_f[A].H, \tilde{H} \rangle = \langle H, DT_f[A].\tilde{H} \rangle \quad \text{für alle } A \in \text{Sym}^1(n) \text{ und } H, \tilde{H} \in \text{Sym}(n). \quad (2.5.15)$$

*Beweis.* Siehe Martin und Neff [47, Lemma 2.1]. ■

**Lemma 2.5.34.** Sei  $I \subset \mathbb{R}$  ein reelles Intervall,  $f: I \rightarrow \mathbb{R}$  eine reelle Funktion und  $T_f: \text{Sym}^1(n) \rightarrow \text{Sym}(n)$  die zugehörige primäre Matrixfunktion. Dann gilt

$$D \text{tr}(T_f[X]) = T_{f'}(X) \quad \text{für alle } X \in \text{Sym}^1(n). \quad (2.5.16)$$

**Bemerkung 2.5.35.** Definieren wir  $W: \text{Sym}^1 \rightarrow \text{Sym}(n)$  mit  $W = \text{tr} T_f$ , dann gilt  $DW[X] = T_{f'}(X)$ . Somit ist  $W$  ein Potential von  $T_{f'}$ . Dazu siehe auch Martin und Neff [47, Proposition 2.5].

*Beweis.* Sei  $X \in \text{Sym}^1(n)$ . Wir nutzen die Linearität und die Selbstadjungiertheit der Ableitung, um die Ableitung von  $\text{tr}(T_f)$  an der Stelle  $X$  in Richtung  $H$  umzuformen:

$$\begin{aligned} D \text{tr}(T_f)[X].H &= \text{tr}(DT_f[X].H) \\ &= \langle DT_f[X].H, \mathbb{1} \rangle = \langle DT_f[X].\mathbb{1}, H \rangle \quad \text{für alle } H \in \text{Sym}(n). \end{aligned}$$

Die Ableitung von  $T_f$  an der Stelle  $X$  in Richtung  $\mathbb{1}$  lautet mittels des Differenzenquotienten

$$\begin{aligned} DT_f[X].\mathbb{1} &= \lim_{t \rightarrow 0} \frac{f(X + t \cdot \mathbb{1}) - f(X)}{t} \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \left( Q \text{diag}(f(\lambda_1 + t), \dots, f(\lambda_n + t)) Q^T - Q \text{diag}(f(\lambda_1), \dots, f(\lambda_n)) Q^T \right) \\ &= \lim_{t \rightarrow 0} Q \text{diag} \left( \frac{f(\lambda_1 + t) - f(\lambda_1)}{t}, \dots, \frac{f(\lambda_n + t) - f(\lambda_n)}{t} \right) Q^T \\ &= Q \text{diag}(f'(\lambda_1), \dots, f'(\lambda_n)) Q^T \\ &= T_{f'}(X). \end{aligned}$$

Zusammen erhalten wir  $D \text{tr}(T_f)[X].H = T_{f'}(X).H$  für alle  $H \in \text{Sym}(n)$ , was die zu zeigende Aussage impliziert. ■

**Proposition 2.5.36.** Für alle  $X \in \text{Sym}^+(n)$  gilt

$$D \|\log X\|^2 = 2 X^{-1} \log X, \quad (2.5.17)$$

$$D \|\text{dev}_n \log X\|^2 = 2 X^{-1} \text{dev}_n \log X = 2 X^{-1} \log \frac{X}{(\det X)^{1/n}}, \quad (2.5.18)$$

$$D |\text{tr}(\log X)|^2 = 2 X^{-1} \text{tr} \log X = 2 X^{-1} \log \det(X). \quad (2.5.19)$$

*Beweis.* Es ist

$$\|\log X\|^2 = \text{tr}((\log X)^T (\log X)) = \text{tr}((\log X)^2) = T_f(X) \quad \text{mit } f(x) = (\log x)^2,$$

so folgt mit Proposition 2.5.34

$$D \|\log X\|^2 = T_{f'}(X) = 2 X^{-1} \log X.$$

Analog bestimmen wir  $D \text{tr}(\log X) = X^{-1}$ , sodass wir per Kettenregel

$$D |\text{tr}(\log X)|^2 = 2 \text{tr}(\log X) X^{-1} = 2 X^{-1} \text{tr}(\log X)$$

## 2. Matrix-Analysis

erhalten. Verwenden wir Proposition 2.2.18, also

$$\|\log X\|^2 = \|\operatorname{dev}_n \log X\|^2 + \frac{1}{n} |\operatorname{tr} \log X|^2,$$

dann können wir die dritte Ableitung als Differenz über die beiden anderen bestimmen:

$$\begin{aligned} D\|\operatorname{dev}_n \log X\|^2 &= D\|\log X\|^2 - \frac{1}{n} D|\operatorname{tr}(\log X)| \\ &= 2X^{-1}(\log X - \frac{1}{n} \operatorname{tr}(\log X) \cdot \mathbb{1}) = 2X^{-1} \operatorname{dev}_n \log X. \quad \blacksquare \end{aligned}$$

Verwenden wir die letzte Proposition mit der Diagonalisierung  $X = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ , so erhalten wir in Übereinstimmung mit Proposition 2.5.32 eine Diagonalmatrix als Ableitungsmatrix, deren Diagonaleinträge die Ableitungen nach den Eigenwerten  $\lambda_1, \dots, \lambda_n$  sind.

**Folgerung 2.5.37.** *Für alle  $X \in \operatorname{Sym}^+(n)$  gilt*

$$\frac{d}{d\lambda_i} \|\log X\|^2 = \frac{2}{\lambda_i} \log \lambda_i, \quad (2.5.20)$$

$$\frac{d}{d\lambda_i} \|\operatorname{dev}_n \log X\|^2 = \frac{2}{\lambda_i} \log \frac{\lambda_i}{(\det X)^{1/n}}, \quad (2.5.21)$$

$$\frac{d}{d\lambda_i} |\operatorname{tr}(\log X)|^2 = \frac{2}{\lambda_i} \operatorname{tr} \log X. \quad (2.5.22)$$

Hierbei sind  $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$  die Eigenwerte von  $X$ .

**Proposition 2.5.38.** *Für alle  $X \in \operatorname{Sym}^+(n)$  gilt*

$$\|\log X\|^2 = \sum_{i=1}^n (\log \lambda_i)^2, \quad (2.5.23)$$

$$\|\operatorname{dev}_n \log X\|^2 = \sum_{i=1}^n \left( \log \frac{\lambda_i}{(\lambda_1 \cdots \lambda_n)^{1/n}} \right)^2 = \frac{1}{n} \sum_{i,j=1, i < j}^n \left( \log \frac{\lambda_i}{\lambda_j} \right)^2, \quad (2.5.24)$$

$$|\operatorname{tr}(\log X)|^2 = (\log \lambda_1 \cdots \lambda_n)^2 = \left( \sum_{i=1}^n \log \lambda_i \right)^2. \quad (2.5.25)$$

Hierbei sind  $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$  die Eigenwerte von  $X$ .

*Beweis.* Wir rechnen nach. Es ist

$$\|\log X\|^2 = \|\operatorname{diag}(\log \lambda_1, \dots, \log \lambda_n)\|^2 = \sum_{i=1}^n (\log \lambda_i)^2$$

und

$$|\operatorname{tr}(\log X)|^2 = \left( \operatorname{tr}(\operatorname{diag}(\log \lambda_1, \dots, \log \lambda_n)) \right)^2 = \left( \sum_{i=1}^n \log \lambda_i \right)^2 = (\log \lambda_1 \cdots \lambda_n)^2.$$

Es gilt einerseits  $\log X = \operatorname{dev}_n \log X + \frac{1}{n} \operatorname{tr}(\log X) \cdot \mathbb{1}$  als auch nach Proposition 2.2.18 auf Seite 23 die Identität  $\|\log X\|^2 = \|\operatorname{dev}_n \log X\|^2 + \frac{1}{n} |\operatorname{tr}(\log X)|^2$ . Mit der ersten Identität erhalten wir sofort

$$\|\operatorname{dev}_n \log X\|^2 = \left\| \log X - \frac{1}{n} \log \det X \cdot \mathbb{1} \right\|^2 = \left\| \log \frac{X}{\det X^{1/n}} \right\|^2 = \sum_{i=1}^n \left( \log \frac{\lambda_i}{(\lambda_1 \cdots \lambda_n)^{1/n}} \right)^2.$$

## 2. Matrix-Analysis

Nun verwenden wir die zweite Identität, sowie zudem (2.5.23) und (2.5.24) und erhalten

$$\begin{aligned}
 \|\operatorname{dev}_n \log X\|^2 &= \|\log X\|^2 - \frac{1}{n} |\operatorname{tr}(\log X)|^2 \\
 &= \sum_{k=1}^n (\log \lambda_k)^2 - \frac{1}{n} \left( \sum_{k=1}^n \log \lambda_k \right)^2 \\
 &= \frac{1}{n} \left( \sum_{i=1}^n \sum_{j=1}^n (\log \lambda_i)^2 - \sum_{i=1}^n \log \lambda_i \sum_{j=1}^n \log \lambda_j \right) \\
 &= \frac{1}{2n} \left( \sum_{i=1}^n \sum_{j=1}^n (\log \lambda_i)^2 - 2 \sum_{i=1}^n \sum_{j=1}^n \log \lambda_i \log \lambda_j + \sum_{i=1}^n \sum_{j=1}^n (\log \lambda_j)^2 \right) \\
 &= \frac{1}{2n} \sum_{i=1}^n \sum_{j=1}^n (\log \lambda_i - \log \lambda_j)^2 \\
 &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \left( \log \frac{\lambda_i}{\lambda_j} \right)^2 \\
 &= \frac{1}{n} \sum_{i,j=1, i < j}^n \left( \log \frac{\lambda_i}{\lambda_j} \right)^2. \quad \blacksquare
 \end{aligned}$$

**Proposition 2.5.39.** Sei  $\Phi: \operatorname{Sym}^+(n) \rightarrow \mathbb{R}$  mit  $\Phi(U) = |\operatorname{tr} \log U|^2$ . Dann ist

$$\frac{\partial \Phi}{\partial I_k}(U) = 0 \quad \text{für alle } k \in \{1, \dots, n-1\} \quad \text{und} \quad \frac{\partial \Phi}{\partial I_n}(U) = \frac{2}{\det U} \operatorname{tr} \log U. \quad (2.5.26)$$

Hierbei sind  $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$  die Eigenwerte und  $I_1, \dots, I_n$  die Invarianten von  $U$ .

*Beweis.* Aufgrund der Identität  $\operatorname{tr} \log U = \log \det U$  ist  $\Phi(U) = (\log I_n)^2$ , ein nur von  $I_n$  abhängender Ausdruck. Somit ist  $\Phi$  konstant in  $I_1, \dots, I_{n-1}$ , die zugehörigen partiellen Ableitungen  $\frac{\partial \Phi}{\partial I_1} = \dots = \frac{\partial \Phi}{\partial I_{n-1}} = 0$ . Die  $n$ -te partielle Ableitung ergibt sich direkt als  $\frac{\partial \Phi}{\partial I_n} = \frac{2}{I_n} \log I_n$ .  $\blacksquare$

Ist eine isotrope Funktion sogar linear, dann existieren (feste) Koeffizienten, sodass sich für jedes Argument  $X$  der Funktionswert  $\mathbb{C} \cdot X$  gewichtet in symmetrischen Anteil, schiefsymmetrischen Anteil und Spuranteil zerlegen lässt, die Abbildung sich also als gewichtete Zerlegung darstellen lässt. Diesen Umstand werden wir in Abschnitt 4.3.13 auf Seite 123 ff. bei der Einführung linearen Elastizitätsgesetzes und in dessen Anwendung auf Shear-Deformationen in Abschnitt 7.1 auf Seite 157 ff. verwenden.

**Lemma 2.5.40.** Sei  $\mathbb{C}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  eine isotrope lineare Funktion. Dann existieren konstante Koeffizienten  $\mu, \mu_c, \Lambda \in \mathbb{R}$ , sodass sich  $\mathbb{C} \cdot X$  als mit ihnen gewichtete Summe des symmetrischen, des schiefsymmetrischen und des Spuranteils von  $X$  darstellen lässt, d.h.

$$\mathbb{C} \cdot X = \mu \operatorname{sym} X + \mu_c \operatorname{skew} X + \frac{\Lambda}{2} \operatorname{tr}(X) \cdot \mathbb{1} \quad \text{für alle } X \in \mathbb{R}^{n \times n}.$$

*Beweis.* Siehe Boor [11].  $\blacksquare$

### 2.5.2. Semi-Invertierbarkeit, Baker-Ericksen und empirischen Ungleichungen

Wie wir in Lemma 2.5.28 (Darstellungssatz) auf Seite 34 gesehen haben, können isotrope Funktionen genau über die Darstellbarkeit durch

$$\Phi(X) = \alpha_0 \cdot \mathbb{1} + \alpha_1 X + \dots + \alpha_{n-1} X^{n-1} = \beta_{-1} \cdot B^{-1} + \beta_0 \cdot \mathbb{1} + \dots + \beta_{n-2} X^{n-2}$$

## 2. Matrix-Analysis

für alle  $X \in \mathfrak{S}$  charakterisiert werden. Ausdrücklich sei hier nochmal erwähnt, dass die hier die  $\alpha_0, \dots, \alpha_{n-1}$  und  $\beta_{-1}, \beta_0, \dots, \beta_{n-2}$  **selbst Funktionen**  $\mathbb{R}^n \rightarrow \mathbb{R}$  **sind**, welche nur von den Invarianten des Arguments  $X$  abhängen dürfen.

Recht ähnlich dieser Darstellung folgen wir einer Definition von C. Truesdell und definieren eine abgeschwächte Variante der Invertierbarkeit von  $\Phi$  unter Kenntnis gewisser Informationen des Arguments:

**Definition 2.5.41** (Semi-Invertierbarkeit). Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge. Eine Abbildung  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  heißt *semi-invertierbar*, wenn es nur von den Invarianten von  $X$  abhängige Koeffizienten  $\psi_0, \dots, \psi_{n-1}$  gibt (d.h. sie sind Funktionen in den Invarianten), sodass gilt

$$X = \sum_{k=0}^{n-1} \psi_k \Phi(X)^k = \psi_0 \cdot \mathbb{1} + \psi_1 \Phi(X) + \dots + \psi_{n-1} \Phi(X)^{n-1} \quad \text{für alle } X \in \mathfrak{S}. \quad (2.5.27)$$

Wir können  $\Phi$  somit „invertieren“, wenn zumindest die Invarianten des Urbildes gegeben sind.

Tatsächlich müssen semi-invertierbare Funktion nicht notwendigerweise invertierbar sein: Als ein Gegenbeispiel schlagen Truesdell und Moon 1975 in einem Artikel [78, S. 185] vor:

**Beispiel 2.5.42.** Sei  $\Phi: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  mit  $\Phi(X) = \text{dev}_3 X$ , also

$$\Phi(X) = X - \frac{1}{3} \text{tr} X \cdot \mathbb{1} = \sum_{i=0}^2 \alpha_i X^i \quad (2.5.28)$$

mit  $\alpha_0(a, b, c) = -\frac{1}{3}a$ ,  $\alpha_1 \equiv 1$ ,  $\alpha_2 \equiv 0$ . Wir sehen, offenkundig ist  $\Phi$  nicht invertierbar, denn es gilt  $\Phi(\mathbb{1}) = \Phi(0) = 0$ . Aber  $\Phi$  ist semi-invertierbar, denn es gilt

$$X = \frac{1}{3} \text{tr} X \cdot \mathbb{1} + \Phi(X) = \psi_0 \cdot \mathbb{1} + \psi_1 \Phi(X) + \psi_2 \Phi(X)^2 \quad (2.5.29)$$

mit  $\psi_0(a, b, c) = \frac{a}{3}$ ,  $\psi_1 \equiv 1$ ,  $\psi_2 \equiv 0$ .

Als weiteres Beispiel sei das von der quadratischen Hencky-Energie aus Beispiel 6.1.20 auf Seite 144 induzierte Elastizitätsgesetz genannt. Dies ist ebenso semi-invertierbar, aber nicht invertierbar ist (siehe auch Bemerkung 6.1.21).

Wir erinnern uns: Für  $\mathfrak{S} \subseteq \text{Sym}(n)$  ist eine Funktion  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  genau dann bi-koaxial, wenn  $X$  und  $\Phi(X)$  für jedes  $X \in \mathfrak{S}$  bi-koaxial ist, d.h. für jedes  $Q \in O(n)$  gilt

$$Q^T X Q \text{ ist in Diagonalgestalt genau dann, wenn } Q^T \Phi(X) Q \text{ in Diagonalgestalt ist.}$$

Im folgenden Lemma, das sich auch in Blume [10, S. 277] findet, sehen wir, dass genau die bi-koaxialen Funktionen semi-invertierbar sind.

**Lemma 2.5.43.** *Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge. Eine isotrope Abbildung  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  ist genau dann semi-invertierbar, wenn sie bi-koaxial ist.*

*Beweis.* Angenommen,  $\Phi$  ist semi-invertierbar. Sei  $X \in \mathfrak{S}$  beliebig aber festgehalten. Dann ist  $X$  aufgrund der Semi-Invertierbarkeit ein Polynom in  $\Phi(X)$  und mit Proposition 2.5.10 ist jeder Eigenvektor von  $\Phi(X)$  auch Eigenvektor von  $X$ . Aufgrund der Isotropie ist auch jeder Eigenvektor von  $X$  auch Eigenvektor von  $\Phi(X)$ . Also ist  $\Phi$  bi-koaxial.

Angenommen,  $\Phi$  ist bi-koaxial. Sei  $X \in \mathfrak{S}$ , dann ist  $\Phi(X)$  koaxial zu  $X$ . Mit Proposition 2.5.27 gibt es daher Koeffizienten  $\psi_0, \dots, \psi_{n-1} \in \mathbb{R}$ , sodass

$$X = \psi_0 \cdot \mathbb{1} + \psi_1 \Phi(X) + \psi_2 \Phi(X)^2 + \dots + \psi_{n-1} \Phi(X)^{n-1}$$

gilt. Aufgrund der Isotropie von  $\Phi$  gilt dann auch

$$Q^T X Q = \psi_0 \cdot \mathbb{1} + \alpha_1 \Phi(Q^T X Q) + \psi_2 \Phi(Q^T X Q)^2 + \dots + \psi_{n-1} \Phi(Q^T X Q)^{n-1} \quad \text{für alle } Q \in O(n),$$

was bedeutet, dass die Koeffizienten  $\psi_0, \dots, \psi_{n-1}$  sich nicht ändern, wenn  $X$  durch  $Q^T X Q$  ersetzt wird. Mit Lemma 2.5.26 sind sie somit Funktionen der Invarianten von  $X$ . Also ist  $\Phi$  semi-invertierbar. ■

## 2. Matrix-Analysis

Lemma 2.5.43 macht die Semi-Invertierbarkeit einer isotropen Funktion zu einer begehrten Eigenschaft.

Wir haben in Folgerung 2.5.16 gesehen, dass wir die Koaxialität einer Funktion  $\Phi$  auch mittels  $a_i \neq a_j \Rightarrow b_i \neq b_j$  für alle  $i, j \in \{1, \dots, n\}$  charakterisieren können. Hierbei sind  $a_i$  die Eigenwerte von  $X$  und  $b_i$  die zugehörigen Eigenwerte von  $\Phi(X)$ .<sup>10</sup> Diese Eigenschaft haben Baker und Ericksen [1] verschärft.

**Definition 2.5.44** (strikte Baker-Ericksen-Ungleichungen ( $\text{BE}^+$ )). Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge und  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  eine isotrope Funktion. Dann erfüllt  $\Phi$  die *strikten Baker-Ericksen-Ungleichungen*, wenn für alle  $X \in \mathfrak{S}$  die folgende Implikation gilt:

$$a_i > a_j \quad \Rightarrow \quad b_i > b_j \quad \text{für alle } i, j \in \{1, \dots, n\}, \quad (2.5.30)$$

wobei  $a_1, \dots, a_n \in \mathbb{R}$  die Eigenwerte von  $X$  und  $b_1, \dots, b_n \in \mathbb{R}$  die Eigenwerte von  $\Phi(X)$  bezeichnen. Hierbei gilt

$$\text{diag}(b_1, \dots, b_n) = \Phi(\text{diag}(a_1, \dots, a_n)), \quad (2.5.31)$$

d.h.  $b_1, \dots, b_n$  sind die zugehörigen Eigenwerte von  $\Phi(X)$  zu  $a_1, \dots, a_n$  von  $X$ .

**Definition 2.5.45** (Baker-Ericksen-Ungleichungen (BE)). Sei  $\Phi$  wie in der vorigen Definition gegeben. Dann erfüllt  $\Phi$  die *Baker-Ericksen-Ungleichungen*, wenn für alle  $X \in \mathfrak{S}$  die folgende Implikation gilt:

$$a_i \geq a_j \quad \Rightarrow \quad b_i \geq b_j \quad \text{für alle } i, j \in \{1, \dots, n\}. \quad (2.5.32)$$

**Bemerkung 2.5.46.** Da aufgrund der Isotropie von  $\Phi$  schon die Implikation  $a_i = a_j \Rightarrow b_i = b_j$  gilt, implizieren die strikten Baker-Ericksen-Ungleichungen die Baker-Ericksen-Ungleichungen. Aber nicht jede isotrope Funktion, welche die Baker-Ericksen-Ungleichungen erfüllt, erfüllt auch die strikten Baker-Ericksen-Ungleichungen. Als Gegenbeispiel hierzu betrachten wir die konstante Nullfunktion  $\Phi \equiv 0$ .

**Bemerkung 2.5.47.** Als Verschärfung der Bi-Koaxialitätseigenschaft  $a_i \neq a_j \Rightarrow b_i \neq b_j$  ist natürlich jede Funktion, welche die Baker-Ericksen-Ungleichungen erfüllt, auch bi-koaxial.

Eine weitere Verschärfung der Bi-Koaxialitätseigenschaft von isotropen Funktionen wollen wir nun beschreiben. Truesdell nannte sie 1975 [78] zuerst die *empirischen Ungleichungen* (empirical inequalities). Hierbei beziehen sich die besagten Ungleichungen auf Koeffizientenfunktionen in der Darstellung einer isotropen Funktion laut Darstellungssatz (Lemma 2.5.28).

**Definition 2.5.48** (empirische Ungleichungen (E-TSS)). Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge und  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  eine isotrope Funktion mit der Darstellung

$$\Phi(X) = \sum_{i=-1}^{n-2} \beta_i X^i = \beta_{-1} X^{-1} + \beta_0 \cdot \mathbb{1} + \dots + \beta_{n-2} X^{n-2} \quad \text{für alle } X \in \mathfrak{S}. \quad (2.5.33)$$

Die sogenannten *empirischen Ungleichungen* (empirical inequalities) stellen folgende Bedingungen an die Abbildungen in den Invarianten  $\beta_{-1}, \beta_0, \dots, \beta_{n-2}: \mathbb{R}^n \rightarrow \mathbb{R}$ . Es soll für alle  $X \in \mathfrak{S}$  gelten:

$$\begin{array}{l} \beta_{-1} \leq 0, \\ \beta_0 \leq 0 \end{array} \quad \text{und} \quad \begin{array}{l} \beta_k \geq 0 \\ \beta_k > 0 \end{array} \quad \text{für alle } k \in \{1, \dots, n-2\}, \quad (2.5.34)$$

**Bemerkung 2.5.49.** Im Fall  $n = 3$ , also  $\Phi(X) = \beta_{-1} X^{-1} + \beta_0 \cdot \mathbb{1} + \beta_1 X$  lauten die *empirischen Ungleichungen*  $\beta_{-1} \leq 0, \beta_0 \leq 0$  und  $\beta_1 > 0$  für alle  $X \in \mathfrak{S}$ .

Truesdell fordert in seiner Definition der empirischen Ungleichungen, dass für **ein**  $k \in \{1, \dots, n-2\}$  die strikte Ungleichung  $\beta_k > 0$  für **alle**  $X \in \mathfrak{S}$  gilt. Dies ist für die in diesem Abschnitt herausgearbeitete Implikationskette, an dessen Ende die Semi-Invertierbarkeit steht, nicht notwendig. Es reicht, wenn für jedes  $X \in \mathfrak{S}$  jeweils eine der Ungleichungen strikt ist. Zudem stellt sich die Forderung an  $\beta_0$  für die Implikationskette als unnützlich heraus. Mehr noch, in Bemerkung 6.1.6 aus Seite 138 stellen wir ein praktisch relevantes Beispiel vor, das alle Bedingungen der (WE-TSS) erfüllt, bis auf die an  $\beta_0$ , welche verletzt wird. Um diese unnötige Strenge bereinigt, definieren wir die schwachen empirischen Ungleichungen.

<sup>10</sup>d.h.  $\Phi(\text{diag}(a_1, \dots, a_n)) = (b_1, \dots, b_n)$ .

## 2. Matrix-Analysis

**Definition 2.5.50** (schwache empirische Ungleichungen (WE-TSS)). Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge und  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  eine isotrope Funktion mit der Darstellung

$$\Phi(X) = \sum_{i=-1}^{n-2} \beta_i X^i = \beta_{-1} X^{-1} + \beta_0 \cdot \mathbb{1} + \dots + \beta_{n-2} X^{n-2} \quad \text{für alle } X \in \mathfrak{S}. \quad (2.5.35)$$

Die sogenannten *schwachen empirischen Ungleichungen* (weak empirical inequalities) stellen folgende Bedingungen an die Abbildungen in den Invarianten  $\beta_{-1}, \beta_1, \dots, \beta_{n-2}: \mathbb{R}^n \rightarrow \mathbb{R}$ . Es soll für alle  $X \in \mathfrak{S}$  gelten:

$$\beta_{-1} \leq 0 \quad \text{und} \quad \beta_k \geq 0 \quad \text{für alle } k \in \{1, \dots, n-2\},$$

Zudem muss es zu jedem  $X \in \mathfrak{S}$  ein  $k \in \{1, \dots, n-2\}$  geben, sodass die Ungleichung  $\beta_k > 0$  strikt ist, insbesondere ist es nicht erforderlich, dass eine der Ungleichungen  $\beta_k \geq 0$  für alle  $X \in \mathfrak{S}$  strikt ist.

**Bemerkung 2.5.51.** Im Fall  $n = 3$ , also  $\Phi(X) = \beta_{-1} X^{-1} + \beta_0 \cdot \mathbb{1} + \beta_1 X$  lauten die *schwachen empirischen Ungleichungen*  $\beta_{-1} \leq 0$  und  $\beta_1 \geq 0$  für alle  $X \in \mathfrak{S}$ , wobei für jedes  $X$  eine der Ungleichungen jeweils strikt ist.

Die empirischen Ungleichungen sind definitionsgemäß eine Verschärfung der schwachen empirischen Ungleichungen, somit erfüllt jede Funktion, welche die (E-TSS) erfüllt auch die (WE-TSS). Letztere verschärfen zudem die strict Baker-Ericksen-Ungleichungen ( $\text{BE}^+$ ), wie das folgende Lemma zeigt. Und da die ( $\text{BE}^+$ ), wie gezeigt, eine Verschärfung der Bi-Koaxialität bedeuten, erhalten wir die gewünschte Eigenschaft der Verschärfung von (bi-coax) durch die empirischen Ungleichungen.

**Lemma 2.5.52.** Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge und  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  erfülle die schwachen empirischen Ungleichungen. Dann erfüllt  $\Phi$  auch die strikten Baker-Ericksen-Ungleichungen.

*Beweis.* Sei  $X \in \text{Sym}(n)$ . Da  $\Phi$  die schwachen empirischen Ungleichungen erfüllt, gibt es  $\beta_{-1} \leq 0$  und  $\beta_1 \geq 0, \dots, \beta_{n-2} \geq 0$ , wobei eine der Ungleichungen sogar strikt ist, sodass  $\Phi(X) = \sum_{l=-1}^{n-2} \beta_l X^l$  gilt. Seien  $a_1, \dots, a_n \in \mathbb{R}$  die Eigenwerte von  $X$  und  $b_1, \dots, b_n \in \mathbb{R}$  die Eigenwerte von  $\Phi(X)$  mit  $\text{diag}(b_1, \dots, b_n) = \Phi(\text{diag}(a_1, \dots, a_n))$ . Aufgrund der Isotropie von  $\Phi$  und da die  $\beta_l$  Invarianten von  $X$  sind, folgt

$$\text{diag}(b_1, \dots, b_n) = \sum_{l=-1}^{n-2} \beta_l \text{diag}(a_1, \dots, a_n)^l$$

und damit

$$b_k = \sum_{l=-1}^{n-2} \beta_l a_k^l \quad \text{für alle } k \in \{1, \dots, n\}.$$

Gelte nun  $i, j \in \{1, \dots, n\}$  mit  $a_i > a_j$ , dann ist  $\beta_{-1} a_i^{-1} \geq \beta_{-1} a_j^{-1}$  und  $\beta_1 a_i \geq \beta_1 a_j, \dots, \beta_{n-2} a_i^{n-2} \geq \beta_{n-2} a_j^{n-2}$ , wobei eine Ungleichung strikt ist. So erhalten wir

$$b_i = \beta_{-1} a_i^{-1} + \beta_0 + \sum_{l=1}^{n-2} \beta_l a_i^l > \beta_{-1} a_j^{-1} + \beta_0 + \sum_{l=1}^{n-2} \beta_l a_j^l = b_j,$$

also erfüllt  $\Phi$  die Baker-Ericksen-Ungleichungen. ■

Abschließend beziehen wir Satz 2.5.56 noch auf den Begriff der Invertierbarkeit.

**Definition 2.5.53** (Invertierbarkeit). Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  und  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$ . Wir nennen  $\Phi$  invertierbar, wenn es eine Umkehrfunktion  $\Phi^{-1}: \text{Sym}(n) \rightarrow \mathfrak{S}$  gibt, für die

$$X = \Phi^{-1}(\Phi(X)) \quad \text{für alle } X \in \mathfrak{S} \quad (2.5.36)$$

und

$$Y = \Phi(\Phi^{-1}(Y)) \quad \text{für alle } Y \in \text{Sym}(n) \quad (2.5.37)$$

gilt.

## 2. Matrix-Analysis

**Proposition 2.5.54.** *Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge und  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  eine invertierbare isotrope Funktion. Dann ist  $\Phi$  auch semi-invertierbar.*

*Beweis.* Sei  $X \in \mathfrak{S}$  und  $Y := \Phi(X)$ . Aufgrund der Isotropie gilt  $\Phi(Q^T X Q) = Q^T Y Q$  für alle  $Q \in O(n)$  und damit für die Umkehrfunktion  $\Phi^{-1}(Q^T Y Q) = Q^T X Q$ , d.h.  $\Phi^{-1}$  ist ebenso isotrop. Mit dem Darstellungssatz 2.5.28 gibt es Funktionen  $\alpha_0, \dots, \alpha_{n-1}$  in den Invarianten von  $Y$  mit

$$X = \alpha_0 \cdot \mathbb{1} + \alpha_1 Y + \dots + \alpha_{n-1} Y^{n-1}.$$

Nach Lemma 2.5.26 gibt es isotrope Funktionen  $\tilde{\alpha}_0, \dots, \tilde{\alpha}_{n-1}$  mit  $\alpha_i(I_1(Y), \dots, I_n(Y)) = \tilde{\alpha}_i(Y)$ . Mit der Isotropie von  $\Phi$  ist auch  $X \mapsto \tilde{\alpha}_i(\Phi(X))$  isotrop, was die Existenz von Funktionen  $\psi_0, \dots, \psi_{n-1}$  in den Invarianten von  $X$  mit  $\alpha_i(I_1(Y), \dots, I_n(Y)) = \psi_i(I_1(X), \dots, I_n(X))$  und damit

$$X = \psi_0 \cdot \mathbb{1} + \psi_1 \Phi(X) + \dots + \psi_{n-1} \Phi(X)^{n-1}$$

impliziert. ■

**Bemerkung 2.5.55.** Auch wenn für isotrope Funktionen aus Invertierbarkeit immer Semi-Invertierbarkeit folgt, können wir diese Implikation nicht auf die empirischen Ungleichungen oder die Baker-Ericksen-Ungleichungen verschärfen. Als Gegenbeispiel hierzu betrachten wir die isotrope Funktion  $\Phi: \text{Sym}(n) \rightarrow \text{Sym}(n)$  mit  $\Phi(X) = \mathbb{1} - X$ . Sie ist invertierbar, erfüllt aber trotz ihrer Semi-Invertierbarkeit nicht die Baker-Ericksen-Ungleichungen (und aufgrund der Implikationskette aus Satz 2.5.56 erfüllt sie nicht (E-TSS), (WE-TSS), (BE<sup>+</sup>)).

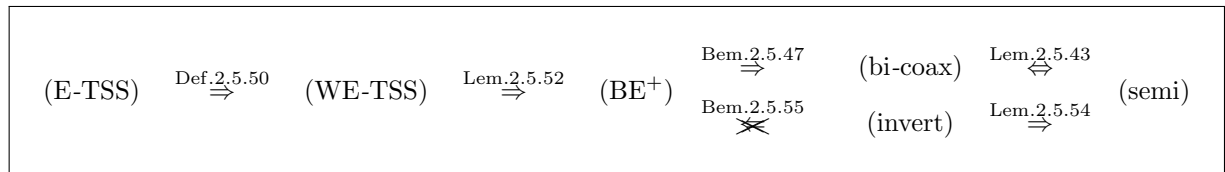
Wir fassen unsere Ergebnisse zusammen:

**Satz 2.5.56.** *Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge und  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  eine isotrope Funktion, d.h. es gilt*

$$\Phi(X) = \sum_{i=-1}^{n-2} \beta_i X^i = \beta_{-1} X^{-1} + \beta_0 \cdot \mathbb{1} + \beta_1 X \quad \text{für alle } X \in \mathfrak{S}. \quad (2.5.38)$$

Hierbei sind die  $\beta_{-1}, \dots, \beta_{n-2}$  reellwertige Funktionen in den Invarianten von  $X$ . Zu jedem  $X$  bezeichnen wir die Eigenwerte von  $X$  mit  $a_1, \dots, a_n$  und die Eigenwerte von  $\Phi(X)$  mit  $b_1, \dots, b_n$ , wobei  $\Phi(\text{diag}(a_1, \dots, a_n)) = \text{diag}(b_1, \dots, b_n)$  gilt.

Dann gelten folgende Implikationen:



Hierbei bezeichnen:<sup>11</sup>

(E-TSS):  $\Phi$  erfüllt die empirischen Ungleichungen.

D.h. es gilt  $\beta_{-1} \leq 0$ ,  $\beta_0 \leq 0$  und  $\beta_0 \geq 1$ ,  $\dots$ ,  $\beta_{n-2} \geq 0$  für alle  $X$  und es gibt ein  $k \in \{1, \dots, n-2\}$  für das  $\beta_k > 0$  für alle  $X$  gilt.

(WE-TSS):  $\Phi$  erfüllt die schwachen empirischen Ungleichungen.

D.h. es gilt  $\beta_{-1} \leq 0$ ,  $\beta_1 \geq 0$ ,  $\dots$ ,  $\beta_{n-2} \geq 0$  für alle  $X$ , wobei für jedes  $X$  eine der Ungleichungen jeweils strikt ist.

<sup>11</sup>Die Schreibweise speziell im Fall  $n = 3$  finden wir auf Seite 135 vor. Dort wird der Satz auf Elastizitätsgesetze angewandt.

## 2. Matrix-Analysis

(BE<sup>+</sup>):  $\Phi$  erfüllt die strikten Baker-Ericksen-Ungleichungen.

D.h. es gilt  $a_i > a_j \Rightarrow b_i > b_j$  für alle  $i, j \in \{1, \dots, n\}$

(bi-coax):  $\Phi$  ist bi-koaxial.

D.h. jeder Eigenvektor von  $X$  zu  $a_i$  ist Eigenvektor von  $\Phi(B)$  zu  $b_i$  und umgekehrt.

(semi):  $\Phi$  ist semi-invertierbar.

D.h. es gilt  $X = \sum_{i=0}^{n-1} \psi_i \Phi(X)^i$  mit reellen Funktionen  $\psi_0, \dots, \psi_{n-1}$  in den Invarianten von  $X$ .

(invert):  $\Phi$  ist invertierbar.

D.h. es gibt eine Umkehrfunktion  $\Phi^{-1}: \text{Sym}(n) \rightarrow \mathfrak{S}$  mit  $X = \Phi^{-1}(\Phi(X))$  für alle  $X \in \mathfrak{S}$  und  $Y = \Phi(\Phi^{-1}(Y))$  für alle  $Y \in \text{Sym}(n)$ .

**Bemerkung 2.5.57.** Die Implikation (E-TSS)  $\Rightarrow$  (semi) wird bereits von Truesdell in [78] erwähnt. Mit ihr haben wir ein einfaches Kriterium für die Semi-Invertierbarkeit:

Sei  $\mathfrak{S} \subseteq \text{Sym}(n)$  eine isotrope Menge und  $\Phi: \mathfrak{S} \rightarrow \text{Sym}(n)$  eine isotrope Funktion mit der Darstellung

$$\Phi(X) = \sum_{i=-1}^{n-2} \beta_i X^i \quad \text{für alle } X \in \mathfrak{S}.$$

und es gilt  $\beta_{-1} \leq 0$  und  $\beta_1 \geq 0, \dots, \beta_{n-2} \geq 0$ , wobei für jedes  $X$  wenigstens eine Ungleichung strikt ist, dann ist  $\Phi$  semi-invertierbar.<sup>12</sup>

## 2.6. Matrixexponential und Matrixlogarithmus

### 2.6.1. Exponential

**Proposition 2.6.1.** Sei  $f_n: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  mit  $f_n(X) = \sum_{k=0}^n \frac{X^k}{k!}$ . Dann konvergiert  $f_n$  punktweise.

*Beweis.* Wir zeigen, dass  $(f_n(X))_{n \in \mathbb{N}}$  eine Cauchyfolge ist. Dazu definieren wir  $g_n: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  mit  $g_n(X) = \sum_{k=0}^n \frac{\|X\|^k}{k!}$ , dann ist für jedes  $x \in \mathbb{R}^{n \times n}$  und alle  $n, m \in \mathbb{N}$  mit  $n \leq m$

$$\|f_m(X) - f_n(X)\| = \left\| \sum_{k=n+1}^m \frac{X^k}{k!} \right\| \leq \sum_{k=n+1}^m \frac{\|X\|^k}{k!} = |g_m(X) - g_n(X)|. \quad (2.6.1)$$

Da für jedes  $X \in \mathbb{R}^{n \times n}$  die Abbildung  $n \mapsto g_n(X) = \sum_{k=0}^n \frac{\|X\|^k}{k!}$  beschränkt und monoton wachsend ist, ist  $g_n$  punktweise konvergent und damit  $(g_n(X))_{n \in \mathbb{N}}$  für jedes  $X \in \mathbb{R}^{n \times n}$  eine Cauchyfolge. Nun folgt mit Abschätzung (2.6.1) direkt das zu Zeigende. ■

Die Wohldefiniertheit der folgenden Definition ergibt sich überhaupt erst durch Proposition 2.6.1.

<sup>12</sup>Allerdings betont Truesdell in [78], dass die Implikation (BE<sup>+</sup>)  $\Rightarrow$  (semi) nur unter der Voraussetzung nichtmehrfacher Nullstellen gilt und – abweichend zu Bemerkung 2.5.47 – im Allgemeinen *nicht* gilt. Sie sprechen von einfachen Gegenbeispielen, die sie trotz ihrer Einfachheit nicht angeben (können?).



## 2. Matrix-Analysis

**Definition 2.6.2.** Wir definieren das *Matrixexponential* auf den quadratischen Matrizen  $\exp: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  über die Potenzreihe

$$\exp(X) := \sum_{k=0}^{\infty} \frac{X^k}{k!}. \quad (2.6.2)$$

Für  $n = 1$  entspricht das Exponential der Exponentialfunktion auf den reellen Zahlen.

**Definition 2.6.3.** Sei  $f: \mathbb{R} \rightarrow \mathbb{R}$ . Die zugehörige *primäre Matrixfunktion*  $T_f: \text{Sym}(n) \rightarrow \text{Sym}(n)$  ist definiert als

$$T_f(X) = Q \text{diag}(f(\lambda_1), \dots, f(\lambda_n)) Q^T = \sum_{i=1}^n f(\lambda_i) \cdot v_i \otimes v_i, \quad (2.6.3)$$

wobei  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Eigenwerte mit zugehörigen Eigenvektoren  $v_1, \dots, v_n$  von  $X$  sind. Somit ist auch  $X = Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^T$  mit  $Q := (v_1 | \dots | v_n)$ .

Für weitere Informationen über Matrixfunktionen siehe auch Richter [68].

**Bemerkung 2.6.4.** Ist die induzierende Funktion  $f$  nur auf den positiven Zahlen definiert, erhalten wir durch analoge Definition  $T_f: \text{Sym}^+(n) \rightarrow \text{Sym}(n)$ .

**Folgerung 2.6.5.** Sei  $\Phi: \text{Sym}(n) \rightarrow \text{Sym}(n)$  eine primäre Matrixfunktion. Dann gilt  $X\Phi(X) = \Phi(X)X$  für alle  $X \in \text{Sym}(n)$ .

*Beweis.* Die Aussage folgt, da  $X$  und  $\Phi(X)$  nach Definition offenkundig gemeinsam diagonalisierbar sind. ■

**Proposition 2.6.6.** Sei  $f_n: \text{Sym}(n) \rightarrow \text{Sym}(n)$  mit  $f_n(X) = \sum_{k=0}^n \frac{X^k}{k!}$ . Dann konvergiert  $f_n$  punktweise gegen die primäre Matrixexponentialfunktion  $T_{\exp}: \text{Sym}(n) \rightarrow \text{Sym}(n)$ .

*Beweis.* Es ist mit  $\tilde{f}_n: \mathbb{R} \rightarrow \mathbb{R}$  und  $\tilde{f}_n(x) = \sum_{k=0}^n \frac{x^k}{k!}$

$$\begin{aligned} f_n(X) &= \sum_{k=0}^n \frac{1}{k!} (Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^T)^k \\ &= \sum_{k=0}^n \frac{1}{k!} Q \text{diag}(\lambda_1, \dots, \lambda_n)^k Q^T \\ &= Q \sum_{k=0}^n \text{diag}\left(\frac{1}{k!} \lambda_1^k, \dots, \frac{1}{k!} \lambda_n^k\right) Q^T \\ &= Q \text{diag}\left(\sum_{k=0}^n \frac{\lambda_1^k}{k!}, \dots, \sum_{k=0}^n \frac{\lambda_n^k}{k!}\right) Q^T = T_{\tilde{f}_n}(X), \end{aligned}$$

wobei  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Eigenwerte von  $X$  sind.

Da nun  $\tilde{f}_n$  gegen die Exponentialfunktion konvergiert ist  $\lim_{n \rightarrow \infty} f_n = T_{\exp}(X)$ . ■

**Folgerung 2.6.7.** Auf die Gruppe der symmetrischen Matrizen eingeschränkt ist das Exponential die primäre Matrixexponentialfunktion, d.h.  $\exp|_{\text{Sym}} = T_{\exp}$ .

**Proposition 2.6.8.** Für  $X, Y \in \mathbb{R}^{n \times n}$  mit  $XY = YX$  gilt

$$\exp(X + Y) = \exp(X) \cdot \exp(Y). \quad (2.6.4)$$

## 2. Matrix-Analysis

*Beweis.* Mit dem Cauchyprodukt

$$\sum_{k=0}^{\infty} a_k \cdot \sum_{k=0}^{\infty} b_k = \sum_{n=0}^{\infty} \sum_{k=0}^n a_k b_{n-k}$$

erhalten wir

$$\begin{aligned} \exp(X) \cdot \exp(Y) &= \sum_{k=0}^n \frac{(X)^k}{k!} \cdot \sum_{k=0}^n \frac{Y^k}{k!} = \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{X^k Y^{n-k}}{k!(n-k)!} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} X^k Y^{n-k} = \sum_{n=0}^{\infty} \frac{(X+Y)^n}{n!} = \exp(X+Y). \end{aligned}$$

Hierbei ist  $\sum_{k=0}^n \binom{n}{k} X^k Y^{n-k} = (X+Y)^n$  genau dann, wenn  $X$  und  $Y$  kommutieren. ■

**Folgerung 2.6.9.** Für  $X = A \cdot N$  mit  $A \in \text{Sym}(n)$  und  $N \in \mathbb{R}^{n \times n}$  nilpotent mit  $N^s = 0$  gilt

$$\exp(X) = \exp(A) \cdot \exp(N) = Q \text{diag}(\exp(\lambda_1), \dots, \exp(\lambda_n)) Q^T \cdot \sum_{k=0}^{s-1} \frac{N^k}{k!}. \quad (2.6.5)$$

Hierbei sind  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Eigenwerte von  $A$  mit Diagonalisierung  $A = Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^T$  mit  $Q \in O(n)$ .

Sicherlich ist  $aXbX = bXaX$ , so erhalten wir:

**Folgerung 2.6.10.** Für alle  $X \in \mathbb{R}^{n \times n}$  und  $a, b \in \mathbb{R}$  gilt

$$\exp((a+b)X) = \exp(aX) \cdot \exp(bX). \quad (2.6.6)$$

Setzen wir  $a = b = 1$ , erhalten wir

**Folgerung 2.6.11.** Für alle  $X \in \mathbb{R}^{n \times n}$  gilt

$$\exp(2X) = (\exp(X))^2. \quad (2.6.7)$$

Setzen wir jetzt noch  $a = 1$  und  $b = -1$ , so ist  $\mathbb{1} = \exp(X) \cdot \exp(-X)$  und damit:

**Folgerung 2.6.12.** Für alle  $X \in \mathbb{R}^{n \times n}$  gilt

$$\exp(-X) = (\exp(X))^{-1}. \quad (2.6.8)$$

**Proposition 2.6.13.** Für  $X \in \mathbb{R}^{n \times n}$  und  $S \in \text{GL}(n)$  gilt

$$\exp(SXS^{-1}) = S \exp(X) S^{-1}. \quad (2.6.9)$$

*Beweis.* Für alle  $k \in \mathbb{N}_0$  ist  $(SXS^{-1})^k = SX^kS^{-1}$ . ■

Auf symmetrische Matrizen  $X$  und orthogonale Matrizen  $Q$  eingeschränkt, erhalten wir die Isotropie des Exponentials:

**Folgerung 2.6.14** (Isotropie). Für  $X \in \text{Sym}(n)$  und  $Q \in O(n)$  gilt

$$\exp(QXQ^T) = Q \exp(X) Q^T. \quad (2.6.10)$$

Direkt aus der Definition der Potenzreihe folgt

## 2. Matrix-Analysis

**Proposition 2.6.15.** Für alle  $X \in \mathbb{R}^{n \times n}$  gilt

$$\exp(X^T) = (\exp(X))^T. \quad (2.6.11)$$

**Folgerung 2.6.16.** Auf die Menge der symmetrischen Matrizen eingeschränkt ist das Exponential eine primäre Matrixfunktion mit  $\exp: \text{Sym}(n) \rightarrow \text{Sym}^+(n)$ . Es gilt

$$\exp(X) = Q \text{diag}(\exp(\lambda_1), \dots, \exp(\lambda_n)) Q^T = \sum_{i=1}^n \exp(\lambda_i) \cdot v_i \otimes v_i, \quad (2.6.12)$$

wobei  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Eigenwerte von  $X$  sind mit zugehörigen, paarweise orthogonales Eigenvektoren  $v_1, \dots, v_n$  von  $X$  und  $Q = (v_1 \mid \dots \mid v_n)$ .

*Beweis.* Für  $X \in \mathbb{R}^{n \times n}$  mit  $X = X^T$ , also  $X \in \text{Sym}(n)$  gilt  $\exp(X) = \exp(X^T) = (\exp(X))^T$ , also  $\exp(X) \in \text{Sym}(n)$ . Mit Proposition 2.6.13 erhalten wir

$$\begin{aligned} \exp(X) &= \exp(Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^T) = Q \exp(\text{diag}(\lambda_1, \dots, \lambda_n)) Q^T \\ &= Q \text{diag}(\exp(\lambda_1), \dots, \exp(\lambda_n)) Q^T. \end{aligned} \quad \blacksquare$$

### 2.6.2. Logarithmus

**Definition 2.6.17.** Wir definieren den *Matrixlogarithmus* auf den quadratischen Matrizen  $\exp: \mathbb{R}^{n \times n} \rightarrow \mathcal{P}(\mathbb{R}^{n \times n})$  mit  $X \mapsto \text{Log}(X) := \{ Y \in \mathbb{R}^{n \times n} \mid \exp(Y) = X \}$ , als die Menge aller Lösungen  $\exp(\text{Log}(X)) = X$ .

**Proposition 2.6.18.** Für alle  $X \in \mathbb{R}^{n \times n}$  gilt

$$2 \text{Log}(X) \subseteq \text{Log}(X^2). \quad (2.6.13)$$

*Beweis.* Mit Folgerung 2.6.11 ist

$$\exp(2 \text{Log}(X)) = (\exp(\text{Log}(X)))^2 = X^2,$$

also  $2 \text{Log}(X) \subseteq \text{Log}(X^2)$ . ■

**Proposition 2.6.19.** Für  $X \in \mathbb{R}^{n \times n}$  und  $S \in \text{GL}(n)$  gilt

$$\text{Log}(SXS^{-1}) = S \text{Log}(X) S^{-1}. \quad (2.6.14)$$

*Beweis.* Mit Proposition 2.6.13 ist  $\exp(S \text{Log} X S^{-1}) = S X S^{-1}$ , also  $\text{Log}(SXS^{-1}) = S \text{Log}(X) S^{-1}$ . ■

Auf symmetrische Matrizen  $X$  und orthogonale Matrizen  $Q$  eingeschränkt, erhalten wir die Isotropie des Logarithmus:

**Folgerung 2.6.20** (Isotropie). Für  $X \in \text{Sym}(n)$  und  $Q \in \text{O}(n)$  gilt

$$\text{Log}(QXQ^T) = Q \text{Log}(X) Q^T. \quad (2.6.15)$$

**Folgerung 2.6.21.** Auf die Menge der symmetrischen, positiv definiten Matrizen eingeschränkt ist der Logarithmus eine primäre Matrixfunktion, die wir aufgrund ihrer Ähnlichkeit zum reellen Logarithmus mit  $\log: \text{Sym}^+(n) \rightarrow \text{Sym}(n)$  bezeichnen. Für für alle  $X \in \text{Sym}^+(n)$  ist

$$\log(X) = Q \text{diag}(\log(\lambda_1), \dots, \log(\lambda_n)) Q^T = \sum_{i=1}^n \log(\lambda_i) \cdot v_i \otimes v_i, \quad (2.6.16)$$

wobei  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Eigenwerte von  $X$  sind mit zugehörigen, paarweise orthogonalen Eigenvektoren  $v_1, \dots, v_n$  von  $X$  und  $Q = (v_1 \mid \dots \mid v_n)$ .

*Beweis.* Mit  $X = Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^T$  und mit der Umkehrfunktion<sup>13</sup>  $\log: \mathbb{R}_+ \rightarrow \mathbb{R}$  von der reellen

<sup>13</sup>Als Umkehrfunktion ist sie eindeutig.

Exponentialfunktion  $\exp: \mathbb{R} \rightarrow \mathbb{R}_+$  folgt

$$\operatorname{Log}(X) = Q \operatorname{Log}(\operatorname{diag}(\lambda_1, \dots, \lambda_n)) Q^T = Q \operatorname{diag}(\log \lambda_1, \dots, \log \lambda_n) Q^T = T_{\log}(X).$$

Siehe auch Richter [67]. ■

**Bemerkung 2.6.22.** Für  $n = 1$  entspricht der Logarithmus der Logarithmusfunktion auf den reellen Zahlen.

## 2.7. Spurgleichungen

Für nichtkommutierende Matrizen  $X, Y \in \mathbb{R}^{n \times n}$  gilt zwar nicht  $\exp(X+Y) = \exp(X) \cdot \exp(Y)$ , aber in die Spurabbildung eingeschlossen und nur für symmetrische Matrizen gilt zumindest die folgende Ungleichung:

**Proposition 2.7.1** (Golden-Thomson-Ungleichung). *Für alle  $X, Y \in \operatorname{Sym}(n)$  gilt*

$$\operatorname{tr}(\exp(X+Y)) \leq \operatorname{tr}(\exp X \cdot \exp Y). \quad (2.7.1)$$

Ebenfalls in die Spurabbildung eingeschlossen ist die folgende Ungleichung über das Exponential:

**Proposition 2.7.2** (Bernstein-Spurungleichung). *Für  $X \in \mathbb{R}^{n \times n}$  gilt*

$$\operatorname{tr}(\exp(X+X^T)) \geq \operatorname{tr}(\exp X \cdot \exp X^T). \quad (2.7.2)$$

**Bemerkung 2.7.3.** Man beachte das zur Golden-Thomson-Ungleichung umgedrehte Ungleichheitszeichen.

*Beweis.* Siehe Bernstein [6, S. 156]. ■

**Proposition 2.7.4.** *Für  $X \in \mathbb{R}^{n \times n}$  gilt*

$$\operatorname{tr}(\exp X) \leq \operatorname{tr}(\exp(\operatorname{sym} X)). \quad (2.7.3)$$

*Beweis.* Formel (2.3) aus Bernstein [6, Theorem 2.1, S. 156] besagt: Für alle  $A, B \in \mathbb{R}^{n \times n}$  gilt

$$\frac{1}{2} \operatorname{tr}(\exp(A+A^T) + \exp(B+B^T)) \geq \frac{1}{2} \operatorname{tr}(\exp(2A) + \exp(2B)).$$

Setzen wir  $A := B := \frac{1}{2} X$ , dann gilt

$$\begin{aligned} \operatorname{tr}\left(\exp\left(\frac{1}{2}(X+X^T)\right)\right) &= \frac{1}{2} \operatorname{tr}\left(\exp\left(\frac{1}{2}(X+X^T)\right) + \exp\left(\frac{1}{2}(X+X^T)\right)\right) \\ &\geq \frac{1}{2} \operatorname{tr}(\exp(X) + \exp(X)) = \operatorname{tr}(\exp(X)). \end{aligned} \quad \blacksquare$$

**Proposition 2.7.5.** *Für  $X \in \mathbb{R}^{n \times n}$  gilt  $\operatorname{tr}(\operatorname{sym} X) = \operatorname{tr}(X)$ .*

*Beweis.* Es gilt mit der Linearität der Spur und mit ihrer Invarianz gegenüber Transponieren:

$$\operatorname{tr}(\operatorname{sym} X) = \operatorname{tr}\left(\frac{1}{2}(X+X^T)\right) = \frac{1}{2} \operatorname{tr}(X) + \frac{1}{2} \operatorname{tr}(X^T) = \operatorname{tr}(X). \quad \blacksquare$$

**Folgerung 2.7.6.** *Für  $X \in \mathbb{R}^{n \times n}$  gilt*

$$\det(\exp(\operatorname{sym} X)) = \det(\exp(X)). \quad (2.7.4)$$

*Beweis.* Aus Proposition 2.7.5 folgt direkt  $e^{\operatorname{tr}(\operatorname{sym} X)} = e^{\operatorname{tr}(X)}$ , was aufgrund der Identität  $e^{\operatorname{tr}(X)} = \det(\exp(X))$  gleichbedeutend mit dem zu Zeigenden ist. ■

## 2.8. Heron-Verfahren

**Lemma 2.8.1.** Sei  $A \in \text{Sym}^+(n)$ . Wir implementieren das Heron-Verfahren auf  $\text{Sym}^+(n)$  und definieren dazu die Folge  $(X_n)_{n \in \mathbb{N}}$  rekursiv durch  $X_0 = A$  und  $X_{n+1} = \frac{1}{2}(X_n + X_n^{-1}A)$ . Dann konvergiert (analog zum Heron-Verfahren auf  $\mathbb{R}$ )  $(X_n)_{n \in \mathbb{N}}$  mit  $B := \lim_{n \rightarrow \infty} X_n$  und  $B^2 = A$ .

*Beweis.* Sei  $Q \in \text{O}(n)$ ,  $D_0 = \text{diag}(\lambda(A))$ , sodass  $A = QD_0Q^T$  gilt. Zudem sei für ein  $k \in \mathbb{N}$  auch  $X_k \in \text{Sym}^+(n)$  und es gilt  $X_k = QD_kQ^T$  mit  $D_k = \text{diag}(\lambda(X_k))$ . Dann ist

$$\begin{aligned} X_{k+1} &= \frac{1}{2}(X_k + X_k^{-1}A) \\ &= \frac{1}{2}(QD_kQ^T + (QD_kQ^T)^{-1}QD_0Q^T) \\ &= \frac{1}{2}(QD_kQ^T + QD_k^{-1}Q^TQD_0Q^T) \\ &= Q \left( \frac{1}{2}(D_k + D_k^{-1}D_0) \right) Q^T \\ &= Q \left( \text{diag} \left( \frac{1}{2}(\lambda_1(X_k) + \lambda_1(X_k)^{-1}\lambda_1(A)), \dots, \frac{1}{2}(\lambda_n(X_k) + \lambda_n(X_k)^{-1}\lambda_n(A)) \right) \right) Q^T \\ &= Q(f_1(\lambda_1(X_k)), \dots, f_n(\lambda_n(X_k)))Q^T \quad \text{mit } f_i(x) = \frac{1}{2} \left( x + \frac{1}{x} \cdot \lambda_i(A) \right) > 0. \end{aligned}$$

Somit ist  $X_{k+1} \in \text{Sym}^+(n)$  und es gilt  $\lambda_i(X_{k+1}) = f_i(\lambda_i(X_k))$  und  $\lambda_i(X_1) = \lambda_i(A)$ . Somit konvergiert jeder Eigenwert von  $X_k$  nach dem Heronverfahren auf  $\mathbb{R}$ , d.h.  $\lim_{n \rightarrow \infty} \lambda_i(X_n) = \sqrt{\lambda_i(A)}$ . Also gilt  $B := \lim_{n \rightarrow \infty} X_n = Q\sqrt{D_0}Q^T$  und so  $B^2 = A$ . ■

## 2.9. Monotonie und die Tschebyscheff-Summenungleichung

Zunächst geben wir eine Übersicht über einige Definitionen und Basisungleichung, die wir im Folgenden benutzen. Für tiefere Betrachtung siehe Marshall, Olkin und Arnold [45] und Hardy, Littlewood und Pólya [34].

**Definition 2.9.1** (Komponentenweise Monotonie). Seien  $x, y \in \mathbb{R}^n$ . Dann schreiben wir  $x \leq y$  (und  $y \geq x$ ), falls gilt

$$x_i \leq y_i \quad \text{für alle } i \in \{1, \dots, n\}.$$

So nennen wir eine Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}$

- i) *monoton wachsend*, falls  $f(x) \geq f(y)$  für alle  $x, y \in \mathbb{R}^n$  mit  $x \leq y$ ;
- ii) *monoton fallend*, falls  $f(x) \leq f(y)$  für alle  $x, y \in \mathbb{R}^n$  mit  $x \leq y$ .

**Definition 2.9.2** (Geordneter Vektor). Let  $x \in \mathbb{R}^n$ .

- i) Wir bezeichnen mit  $x^\downarrow$  die Umordnung von  $x$  in *absteigender Sortierung*, das bedeutet eine Permutation der Komponenten von  $x$ , sodass gilt

$$x_1^\downarrow \geq x_2^\downarrow \geq \dots \geq x_n^\downarrow.$$

- ii) Wir bezeichnen mit  $x^\uparrow$  die Umordnung von  $x$  in *aufsteigender Sortierung*, das bedeutet eine Permutation der Komponenten von  $x$ , sodass gilt

$$x_1^\uparrow \leq x_2^\uparrow \leq \dots \leq x_n^\uparrow.$$

**Lemma 2.9.3** (Umordnungsungleichung). Seien  $a_1 \leq a_2 \leq \dots \leq a_n$  und  $b_1 \leq b_2 \leq \dots \leq b_n$  zwei monoton wachsende Folgen reeller Zahlen, dann gilt für alle Permutationen  $\pi$  von  $\{1, 2, \dots, n\}$

$$\sum_{i=1}^n a_i b_i \geq \sum_{i=1}^n a_{\pi(i)} b_i \geq \sum_{i=1}^n a_i b_{n+1-i}. \quad (2.9.1)$$

## 2. Matrix-Analysis

*Beweis.* Wir zeigen die linke Ungleichung von (2.9.1) mittels vollständiger Induktion. Der Beweis der rechten Ungleichung erfolgt analog.

Für  $n = 2$  existieren nur die Identität und  $\pi(1) = 2, \pi(2) = 1$  als Permutation. Der erste Fall ist trivial, für den zweiten sehen wir

$$a_1 b_1 + a_2 b_2 \geq a_2 b_1 + a_1 b_2 \Leftrightarrow \underbrace{(a_1 - a_2)}_{\geq 0} \underbrace{(b_1 - b_2)}_{\geq 0} \geq 0.$$

Als Induktionsvoraussetzung (IV) sei (2.9.1) erfüllt für ein  $n \in \mathbb{N}$  und für  $n = 2$ , dann ist  $\pi$  eine vorgegebenen Permutation von  $\{1, \dots, n+1\}$ . Wir definieren  $\tilde{\pi}$  als Permutation von  $\{1, \dots, n\}$  mit  $\tilde{\pi}(i) = \pi(i)$  für  $i \in \{1, \dots, r-1, r+1, \dots, n\}$  und  $\tilde{\pi}(r) = s$ . Nun können wir abschätzen:

$$\begin{aligned} \sum_{i=1}^{n+1} a_i b_i &= \sum_{i=1}^n a_i b_i + a_{n+1} b_{n+1} \stackrel{\text{IV } n}{\geq} \sum_{i=1}^n a_{\tilde{\pi}(i)} b_i + a_{n+1} b_{n+1} \\ &= \sum_{i=1, i \neq r}^n a_{\tilde{\pi}(i)} b_i + a_{\tilde{\pi}(r)} b_r + a_{n+1} b_{n+1} = \sum_{i=1, i \neq r}^n a_{\pi(i)} b_i + a_s b_r + a_{n+1} b_{n+1} \\ &\stackrel{\text{IV } 2}{\geq} \sum_{i=1, i \neq r}^n a_{\pi(i)} b_i + a_{n+1} b_r + a_s b_{n+1} = \sum_{i=1, i \neq r}^n a_{\pi(i)} b_i + a_{\pi(r)} b_r + a_{\pi(n+1)} b_{n+1} = \sum_{i=1}^{n+1} a_{\pi(i)} b_i, \end{aligned} \tag{2.9.2}$$

somit ist (2.9.1) für  $n+1$  erfüllt. ■

**Lemma 2.9.4** (Tschebyscheff-Summenungleichung). *Seien  $a_1 \leq a_2 \leq \dots \leq a_n$  und  $b_1 \leq b_2 \leq \dots \leq b_n$  zwei monoton wachsende Folgen reeller Zahlen. Dann gilt*

$$\sum_{i=1}^n a_i b_i \geq \frac{1}{n} \left( \sum_{i=1}^n a_i \right) \left( \sum_{i=1}^n b_i \right) \geq \sum_{i=1}^n a_i b_{n+1-i}. \tag{2.9.3}$$

**Bemerkung 2.9.5.** Gleichheit gilt in (2.9.3) genau dann, wenn zumindest  $a_1 = a_2 = \dots = a_n$  oder  $b_1 = b_2 = \dots = b_n$ . Siehe auch Mitrinović, Pečarić, Fink [50, S. 240].

*Beweis.* Die Ungleichung folgt direkt aus der Umordnungsungleichung. Wir können die Summanden des Produktes  $(\sum a_i)(\sum b_i)$  in  $n$  Summen über  $a_{\pi_k(i)} b_i$  gruppieren.<sup>14</sup>

$$\begin{aligned} \left( \sum_{i=1}^n a_i \right) \left( \sum_{i=1}^n b_i \right) &= a_1 b_1 + a_2 b_2 + a_3 b_3 + \dots + a_{n-1} b_{n-1} + a_n b_n \\ &\quad + a_2 b_1 + a_3 b_2 + a_4 b_3 + \dots + a_n b_{n-1} + a_1 b_n \\ &\quad + a_3 b_1 + a_4 b_2 + a_5 b_3 + \dots + a_1 b_{n-1} + a_2 b_n \\ &\quad \vdots \\ &\quad + a_n b_1 + a_1 b_2 + a_2 b_3 + \dots + a_{n-2} b_{n-1} + a_{n-1} b_n = \sum_{k=1}^n \sum_{i=1}^n a_{i+k-1} b_i, \end{aligned}$$

demnach  $(\sum_{i=1}^n a_i)(\sum_{i=1}^n b_i) = \sum_{k=1}^n \sum_{i=1}^n a_{\pi_k(i)} b_i$  mit  $\pi_k$  als eine Permutation von  $\{1, \dots, n\}$ . Jede Summe  $\sum_{i=1}^n a_{\pi_k(i)} b_i$  kann durch (2.9.1) abgeschätzt werden, sodass wir

$$n \sum_{i=1}^n a_i b_i \geq \left( \sum_{i=1}^n a_i \right) \left( \sum_{i=1}^n b_i \right) \geq n \sum_{i=1}^n a_i b_{n+1-i}$$

erhalten, was äquivalent zur Aussage ist. ■

<sup>14</sup>wobei der Index  $i+k-1$  modulo  $n$  gemeint ist

### 2.9.1. Partialbruchzerlegung

**Lemma 2.9.6.** *Seien  $z_1, \dots, z_n$  paarweise verschiedene komplexe Zahlen und  $k \in \{0, \dots, n-1\}$ . Dann gilt*

$$\sum_{j=1}^n \frac{z_j^k}{(t-z_j) \prod_{i=1, i \neq j}^n (z_j - z_i)} = \frac{t^k}{\prod_{i=1}^n (t-z_i)} \quad \text{für alle } t \in \mathbb{C} \setminus \{z_1, \dots, z_n\}. \quad (2.9.4)$$

Z.B. erhalten wir für  $n=3$  und  $k=2$ :

$$\frac{a^2}{(a-b)(a-c)(t-a)} + \frac{b^2}{(b-a)(b-c)(t-b)} + \frac{c^2}{(c-a)(c-b)(t-c)} = \frac{t^2}{(t-a)(t-b)(t-c)}.$$

*Beweis.* Sei  $k \in \{0, \dots, n-1\}$ . Dann existieren nach dem Satz über die Partialbruchzerlegung komplexe Zahlen  $a_1, \dots, a_n$ , sodass

$$\frac{t^k}{\prod_{i=1}^n (t-z_i)} = \sum_{i=1}^n \frac{a_i}{t-z_i} \quad \text{für alle } t \in \mathbb{C} \setminus \{z_1, \dots, z_n\}.$$

Sei nun  $j \in \{1, \dots, n\}$  beliebig. Beidseitiges Multiplizieren mit  $t-z_j$  ergibt

$$\frac{t^k}{\prod_{i=1, i \neq j}^n (t-z_i)} = \sum_{i=1}^n \frac{t-z_j}{t-z_i} a_i \quad \text{für alle } t \in \mathbb{C} \setminus \{z_1, \dots, z_n\}.$$

Die Gleichung ist stetig ergänzbar, Grenzübergang  $t \rightarrow z_j$  ergibt

$$\frac{z_j^k}{\prod_{i=1, i \neq j}^n (z_j - z_i)} = a_j. \quad \blacksquare$$

**Folgerung 2.9.7.** *Seien  $z_1, \dots, z_n \in \mathbb{C}$  paarweise verschiedene komplexe Zahlen. Dann gilt*

$$\sum_{j=1}^n \frac{z_j^k}{\prod_{i=1, i \neq j}^n (z_j - z_i)} = 0 \quad \text{für alle } k \in \{0, \dots, n-2\}. \quad (2.9.5)$$

Z.B. erhalten wir für  $n=4$  und  $k=2$ :

$$\frac{a^2}{(a-b)(a-c)(a-d)} + \frac{b^2}{(b-a)(b-c)(b-d)} + \frac{c^2}{(c-a)(c-b)(c-d)} + \frac{d^2}{(d-a)(d-b)(d-c)} = 0.$$

*Beweis.* Sei  $k \in \{0, \dots, n-2\}$ . Nach (2.9.4) gilt

$$\sum_{j=1}^{n-1} \frac{z_j^k}{(t-z_j) \prod_{i=1, i \neq j}^{n-1} (z_j - z_i)} = \frac{t^k}{\prod_{i=1}^{n-1} (t-z_i)} \quad \text{für alle } t \in \mathbb{C} \setminus \{z_1, \dots, z_{n-1}\}. \quad (2.9.6)$$

Setzen wir nun  $t := z_n$ , führt uns Umstellen direkt auf das Resultat. \blacksquare

## 2.10. Differenzierbarkeit

**Lemma 2.10.1** (Partielle Integration für Vektorfelder). *Sei  $\Omega \subseteq \mathbb{R}^n$  eine offene Menge und  $A \subset \Omega$  kompakte Menge mit glattem Rand. Dazu sei  $u: \Omega \rightarrow \mathbb{R}$  ein stetig differenzierbares Skalarfeld und  $f: \Omega \rightarrow V$  ein stetig differenzierbares Vektorfeld. Dann gilt*

$$\int_A u(x) \operatorname{div} f(x) \, dx = \int_{\partial A} \langle u(x) f(x), n(x) \rangle \, dS(x) - \int_A \langle f(x), \nabla u(x) \rangle \, dx. \quad (2.10.1)$$

Hierbei ist  $n(x)$  die nach außen orientierte Normale zur Tangentialebene von  $\Omega$  im Randpunkt  $x \in \partial\Omega$ .

## 2. Matrix-Analysis

*Beweis.* Mit Proposition 1.2.10 gilt

$$\int_A \operatorname{div}(u(x) f(x)) \, dx = \int_A \langle \nabla u(x), f(x) \rangle + u(x) \operatorname{div} f(x) \, dx.$$

Mit Lemma 1.2.13 gilt

$$\int_A \operatorname{div}(u(x) f(x)) \, dx = \int_{\partial A} \langle u(x) f(x), n(x) \rangle \, dS(x).$$

Gleichsetzen und Umstellen unter Ausnutzung der Linearität des Integrals liefert die Aussage. ■

**Lemma 2.10.2** (Partielle Integration für Matrixfelder). *Sei  $\Omega \subseteq \mathbb{R}^n$  eine offene Menge und  $A \subseteq \Omega$  eine kompakte Teilmenge mit glattem Rand. Dazu sei  $\varphi: \Omega \rightarrow \mathbb{R}^n$  ein stetig differenzierbares Vektorfeld und  $\Phi: \Omega \rightarrow \mathbb{R}^{n \times n}$  ein stetig differenzierbares Matrixfeld. Dann gilt*

$$\int_A \langle \varphi(x), \operatorname{div} \Phi(x) \rangle \, dx = \int_{\partial A} \left\langle \sum_{i=1}^n \varphi_i(x) \Phi_i(x)^T, n(x) \right\rangle \, dS(x) - \int_A \langle \Phi(x), D\varphi(x) \rangle \, dx. \quad (2.10.2)$$

Hierbei ist  $n(x)$  die nach außen orientierte Normale zur Tangentialebene von  $\Omega$  im Randpunkt  $x \in \partial\Omega$  und  $\Phi_i(x)$  bezeichnet die  $i$ -te Zeile von  $\Phi(x)$ .

*Beweis.* Wir ziehen die Aussage auf die partielle Integration für Vektorfelder zurück.

Wir drücken  $\Phi$  zeilenweise durch  $\Phi_1, \dots, \Phi_n$  mit  $\Phi_i = (\Phi_{i1} \ \dots \ \Phi_{in})$  aus, dann ist

$$\varphi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix} \quad \text{und} \quad \operatorname{div} \Phi = \begin{pmatrix} \operatorname{div} \Phi_1 \\ \vdots \\ \operatorname{div} \Phi_n \end{pmatrix}$$

also

$$\langle \varphi(x), \operatorname{div} \Phi(x) \rangle = \sum_{i=1}^n \langle \varphi_i(x), \operatorname{div} \Phi_i(x) \rangle.$$

Für die Jacobi-Matrix von  $\varphi(x)$  gilt

$$D\varphi(x) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1}(x) & \dots & \frac{\partial \varphi_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial \varphi_n}{\partial x_1}(x) & \dots & \frac{\partial \varphi_n}{\partial x_n}(x) \end{pmatrix} = \begin{pmatrix} \nabla \varphi_1(x)^T \\ \vdots \\ \nabla \varphi_n(x)^T \end{pmatrix},$$

also

$$\langle \Phi(x), D\varphi(x) \rangle = \sum_{i=1}^n \langle \Phi_i(x)^T, \nabla \varphi_i(x) \rangle.$$

Aufgrund der Linearität des Skalarproduktes gilt offensichtlich

$$\left\langle \sum_{i=1}^n \varphi_i(x) \Phi_i(x)^T, n(x) \right\rangle = \sum_{i=1}^n \langle \varphi_i(x) \Phi_i(x)^T, n(x) \rangle.$$

Wir können nun die Argumente der drei Integrale der Aussage als Summen über Ausdrücke in den Komponentenfunktionen ausdrücken. Mit Lemma 2.10.1 gilt

$$\int_A \langle \varphi_i(x), \operatorname{div} \Phi_i(x)^T \rangle \, dx = \int_{\partial A} \langle \varphi_i(x) \Phi_i(x)^T, n(x) \rangle \, dS(x) - \int_A \langle \Phi_i(x)^T, \nabla \varphi_i(x) \rangle \, dx.$$

Ein Summieren über alle  $i \in \{1, \dots, n\}$  liefert mit der Linearität des Integrals die Aussage. ■



## 2. Matrix-Analysis

Ist  $\varphi$  eine Testfunktion auf  $A$ , dann gilt sogar

**Folgerung 2.10.3.** Sei  $\Omega \subseteq \mathbb{R}^n$  eine offene Menge und  $A \subseteq \Omega$  eine kompakte Teilmenge mit glattem Rand. Dazu sei  $\varphi: \Omega \rightarrow \mathbb{R}^n$  eine Testfunktion auf  $A$  und  $\Phi: \Omega \rightarrow \mathbb{R}^{n \times n}$  ein stetig differenzierbares Matrixfeld. Dann gilt

$$\int_A \langle \varphi(x), \operatorname{div} \Phi(x) \rangle dx = - \int_A \langle D\varphi(x), \Phi(x) \rangle dx. \quad (2.10.3)$$

**Definition 2.10.4** (Lebesgue-Raum). Sei  $\Omega \subseteq \mathbb{R}^n$  offen. Für  $p \geq 1$  bezeichnen wir die Menge

$$L^p(\Omega) := \left\{ f: \Omega \rightarrow \mathbb{R} \mid \int_{\Omega} |f(x)|^p dx < \infty \right\} \quad (2.10.4)$$

mit der Norm  $\|f\|_{L^p(\Omega)}^p := \int_{\Omega} |f(x)|^p dx$  als *Lebesgue-Raum*<sup>15</sup>  $L^p(\Omega)$ .

Der Raum  $L^p(\Omega, \mathbb{R}^n)$  besteht aus allen Funktionen  $f: \Omega \rightarrow \mathbb{R}^n$ , deren Komponentenfunktionen  $f_i$  in  $L^p(\Omega)$  sind, für die Norm ergibt sich entsprechend  $\|f\|_{L^p(\Omega), \mathbb{R}^n}^p := \sum_{i=1}^n \int_{\Omega} |f_i(x)|^p dx$ .

Mit

$$L^p_{\text{loc}}(\Omega) := \left\{ f: \Omega \rightarrow \mathbb{R} \mid \int_K |f(x)|^p dx < \infty \text{ für alle kompakten Teilmengen } K \subset \Omega \right\} \quad (2.10.5)$$

bezeichnen wir die Menge der *lokal  $p$ -integrierbaren Funktionen*.<sup>16</sup> Sind für eine Funktion  $f: \Omega \rightarrow \mathbb{R}^n$  alle Komponentenfunktionen lokal  $p$ -integrierbar, so ist  $f$  selbst lokal  $p$ -integrierbar.

**Bemerkung 2.10.5.** Im Falle von  $p = 1$  ist vorige Definition äquivalent zu: Eine Funktion  $f: \Omega \rightarrow \mathbb{R}$  heißt lokal integrierbar, wenn für alle Testfunktionen  $\psi: \Omega \rightarrow \mathbb{R}$  gilt

$$\int_{\Omega} |f(x)\psi(x)| dx < \infty. \quad (2.10.6)$$

**Bemerkung 2.10.6.** Jede Funktion  $f \in L^p(\Omega)$  ist lokal  $p$ -integrierbar, d.h.  $L^p_{\text{loc}}(\Omega) \subset L^p(\Omega)$ .

**Proposition 2.10.7.** Für jede offene Menge  $\Omega \subset \mathbb{R}^n$  ist der Lebesgue-Raum  $L^p(\Omega)$  mit der Lebesgue-Norm  $\|\cdot\|_{L^p(\Omega)}$  vollständig, also ein Banachraum.

**Lemma 2.10.8** (Fundamentallema der Variationsrechnung). Sei  $\Omega \subset \mathbb{R}^n$  eine offene Menge und  $f \in L^2(\Omega, \mathbb{R}^n)$  eine reellwertige Funktion. Zudem gelte

$$\int_{\Omega} \langle f(x), \psi(x) \rangle dx = 0 \quad \text{für alle Testfunktionen } \psi \in C_0^\infty(\Omega, \mathbb{R}^n). \quad (2.10.7)$$

Dann ist  $f(x) = 0$  fast überall.

*Beweis.* Siehe Evans [26]. ■

**Bemerkung 2.10.9.** Ist  $f: \Omega \rightarrow \mathbb{R}^n$  sogar stetig, so ist  $f$  die Nullfunktion.

Um das Konzept der *schwachen Differenzierbarkeit* zu beleuchten, stellen wir zunächst fest, dass für eine beliebige differenzierbare Funktion  $f \in C^1(\Omega)$  mit offener Menge  $\Omega \in \mathbb{R}^n$  und eine beliebige Testfunktion  $\psi \in C_0^\infty(\Omega)$  die partielle Integrationsformel

$$\int_{\Omega} f(x) \frac{\partial \psi}{\partial x_i}(x) dx = - \int_{\Omega} \frac{\partial f}{\partial x_i} \psi(x) dx$$

<sup>15</sup>Hierbei sind  $f, g \in L^p$  gleich, falls sie fast überall gleich sind. Streng genommen sind die Elemente aus  $L^p$  also keine Funktionen, sondern Äquivalenzklassen von Funktionen. Deren Repräsentanten bezeichnen wir aber weiterhin als die Elemente des Lebesgue-Raums  $L^p$ .

<sup>16</sup>Im Falle von  $p = 1$  auch lokal integrierbare Funktionen genannt.

## 2. Matrix-Analysis

mit  $\frac{\partial f}{\partial x_i} \in C^0(\Omega)$  für alle  $i \in \{1, \dots, n\}$  gilt. Ist  $f$  lediglich in  $L^1_{\text{loc}}(\Omega)$ , so sind beide Seiten der Integrationsformel noch wohldefiniert, wenn wir  $\frac{\partial f}{\partial x_i}$  durch die Abbildung  $g_i \in L^1_{\text{loc}}(\Omega)$  ersetzen. Gilt zudem Gleichheit, so bezeichnen wir  $g_i$  als die  $i$ -te schwache partielle Ableitung aus der sich dann die schwache Ableitung zusammensetzt. Wir verwenden im Falle schwacher Differenzierbarkeit die gleichen Bezeichner  $Df$ ,  $\nabla f$ ,  $\frac{\partial f}{\partial x_i}$  wie bei klassischer Differenzierbarkeit.

Verallgemeinern wir das Vorgehen auf (gemischte) höhere Ableitungen, erhalten wir

**Definition 2.10.10.** Sei  $\Omega \in \mathbb{R}^n$  offen,  $f \in L^1_{\text{loc}}(\Omega) \subset L^1(\Omega)$  und  $\alpha \in \mathbb{N}_0^n$  ein Multiindex. Zudem existiere eine Funktion  $g_\alpha \in L^1_{\text{loc}}(\Omega)$  mit

$$\int_{\Omega} f(x) D^\alpha \psi(x) dx = (-1)^{|\alpha|} \int_{\Omega} g_\alpha(x) \psi(x) dx \quad \text{für alle } \psi \in C_0^\infty(\Omega). \quad (2.10.8)$$

Dann bezeichnen wir  $D^\alpha f := g_\alpha$  als (gemischte partielle) *schwache Ableitung* von  $f$ . Existieren die Ableitungen  $D^\alpha f$  für alle Multiindizes  $\alpha \in \mathbb{N}_0^k$  mit  $|\alpha| \leq k \in \mathbb{N}$ , so nennen wir  $f$  insgesamt  *$k$ -mal schwach differenzierbar* bzw. *schwach differenzierbar bis Ordnung  $k$* .

**Bemerkung 2.10.11.** Aus dem Fundamentallemma der Variationsrechnung folgt direkt<sup>17</sup>, dass sämtliche schwachen Ableitungen  $D^\alpha f$  eindeutig sind.

Ist  $f \in C^k(\Omega)$ , so sind die schwachen Ableitungen mit den klassischen identisch.

**Bemerkung 2.10.12.** Jede schwach differenzierbare Funktion ist fast überall differenzierbar. Die Umkehrung gilt allerdings nicht, als Gegenbeispiel betrachten wir die Sprungfunktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = 0$  für  $x < 0$  und  $f(x) = 1$  für  $x \geq 0$ .

**Beispiel 2.10.13.** Sei  $\Omega = (-1, 1)$  und  $f: \Omega \rightarrow \mathbb{R}$  die Betragsfunktion mit  $f(x) = |x|$ . Dann ist  $f$  nicht differenzierbar, aber fast überall differenzierbar mit  $f'(x) = \text{sgn}(x)$  für alle  $x \in \Omega \setminus \{0\}$ . Die Signumsfunktion ist dennoch die schwache Ableitung der Betragsfunktion, denn es gilt für alle  $\psi \in C_0^\infty(\Omega)$

$$\begin{aligned} \int_{-1}^1 f(x) \psi'(x) dx &= \int_{-1}^0 (-x) \psi'(x) dx + \int_0^1 x \psi'(x) dx \\ &= - \int_{-1}^0 -\psi(x) dx - \int_0^1 \psi(x) dx = - \int_{-1}^1 \text{sgn}(x) \psi(x) dx. \end{aligned} \quad (2.10.9)$$

**Definition 2.10.14.** Für  $k \in \mathbb{N}$  und  $p \geq 1$  bezeichnen wir die Menge

$$W^{k,p}(\Omega) := \{ f \in L^p(\Omega) \mid \text{gemischte partielle schwache Ableitungen } D^\alpha f \in L^p(\Omega) \text{ für alle } |\alpha| \leq k \},$$

mit der Norm

$$\|f\|_{W^{k,p}(\Omega)}^p = \sum_{|\alpha| \leq k} \|D^\alpha f\|_{L^p(\Omega)}^p \quad (2.10.10)$$

als *Sobolev-Raum*  $W^{k,p}(\Omega)$ .

Analog definieren wir die Menge der schwach differenzierbaren Funktionen mit lediglich lokal integrierbaren Ableitungen:

$$W^{k,p}_{\text{loc}}(\Omega) := \{ f \in L^p(\Omega) \mid \text{gemischte partielle schwache Ableitungen } D^\alpha f \in L^p_{\text{loc}}(\Omega) \text{ für alle } |\alpha| \leq k \}.$$

**Bemerkung 2.10.15.** Jede Funktion  $f \in W^{k,p}(\Omega)$  ist, da alle gemischten partiellen Ableitungen bis zur  $k$ -ten Ordnung auch lokal integrierbar sind, ebenfalls in  $W^{k,p}_{\text{loc}}(\Omega)$ , d.h.  $W^{k,p}(\Omega) \subset W^{k,p}_{\text{loc}}(\Omega)$ .

**Proposition 2.10.16.** Für jede offene Menge  $\Omega \subset \mathbb{R}^n$ ,  $k \in \mathbb{N}$  und  $p \geq 1$  ist der Sobolev-Raum  $W^{k,p}(\Omega)$  mit der Sobolev-Norm  $\|\cdot\|_{W^{k,p}(\Omega)}$  vollständig, also ein Banachraum.

<sup>17</sup>Für zwei Kandidaten  $g, \mathfrak{g}$  gilt  $\int_{\Omega} g(x) \psi(x) dx = \int_{\Omega} \mathfrak{g}(x) \psi(x) dx$  für alle  $\psi \in C_0^\infty(\Omega)$  und damit  $g - \mathfrak{g} = 0$  fast überall.

## 2. Matrix-Analysis

**Bemerkung 2.10.17.** Betrachten wir  $x \mapsto D^k f[x]$  als Abbildung von  $x \in \Omega$  auf den Tensor  $k$ -ter Stufe bestehend aus den gemischten partiellen Ableitungen  $k$ -ter Ordnung, so ergibt sich

$$\|D^k f\|_{L^p(\Omega)}^p = \sum_{|\alpha|=k} \|D^\alpha f\|_{L^p(\Omega)}^p \quad (2.10.11)$$

und wir können schreiben

$$\|f\|_{W^{k,p}(\Omega)}^p = \|f\|_{L^p(\Omega)}^p + \|Df\|_{L^p(\Omega)}^p + \dots + \|D^k f\|_{L^p(\Omega)}^p. \quad (2.10.12)$$

**Lemma 2.10.18.** Gegeben sei ein Gebiet<sup>18</sup>  $\Omega \subseteq \mathbb{R}^n$  mit  $C^2$ -Rand und eine Funktion  $g: \partial\Omega \rightarrow \mathbb{R}^n$  auf dem Rand von  $\Omega$ . Wir bezeichnen mit

$$M := \{ \vartheta \in C^1(\Omega) \mid \vartheta|_{\partial\Omega} = g \} \quad (2.10.13)$$

die Menge aller stetig differenzierbaren Funktionen  $\vartheta: \Omega \rightarrow \mathbb{R}^n$ , die der Randbedingung gehorchen, dass ihre stetige Fortsetzung<sup>19</sup> auf dem Rand von  $\Omega$  identisch mit  $g$  ist.

Zudem sei eine zweimal stetig partiell differenzierbare Funktion  $L: \Omega \times \mathbb{R}^n \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  mit  $(x, y, p) \mapsto L(x, y, p)$  gegeben.

Besitzt die Funktion

$$I: C^1(\Omega) \rightarrow \mathbb{R} \quad \text{mit} \quad I(\vartheta) = \int_{\Omega} L(x, \vartheta(x), D\vartheta(x)) \, dx, \quad (2.10.14)$$

ein Minimum  $\vartheta_0 \in M$ , dann erfüllt dieses die schwache Form der Euler-Lagrange-Gleichung:

$$\begin{aligned} & \int_{\Omega} \langle \nabla_y L(x, \vartheta_0(x), D\vartheta_0[x]), \psi(x) \rangle \\ & + \langle \nabla_p L(x, \vartheta_0(x), D\vartheta_0[x]), D\psi(x) \rangle \, dx = 0 \quad \text{für alle } \psi \in C_0^\infty(\Omega). \end{aligned} \quad (2.10.15)$$

Ist das Minimum  $\vartheta_0 \in M$  sogar zweimal stetig differenzierbar, so erfüllt es die Euler-Lagrange-Gleichung:

$$\nabla_y L(x, \vartheta_0(x), D\vartheta_0[x]) - \operatorname{div} \nabla_p L(x, \vartheta_0(x), D\vartheta_0[x]) = 0 \quad \text{für alle } x \in \Omega. \quad (2.10.16)$$

**Bemerkung 2.10.19.** Wir bezeichnen ein  $\vartheta$ , das die schwache Form einer Euler-Lagrange-Gleichung erfüllt, als schwache Lösung der Euler-Lagrange-Gleichung. Erfüllt  $\vartheta$  sogar die Euler-Lagrange-Gleichung selbst, so nennen wir  $\vartheta$  Lösung der Gleichung.

*Beweis.* Wir zeigen nun (2.10.15) für  $\vartheta_0$ . Hierfür sei  $\psi \in C_0^\infty(\Omega)$  eine beliebige Testfunktion auf  $\Omega$ . Für hinreichend kleines  $\varepsilon > 0$  definieren wir die Funktion  $f: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$  mit  $\Upsilon(t) = I(\vartheta_0 + t \cdot \psi)$ . Mit  $\Upsilon(0) = I(\vartheta_0)$  hat  $\Upsilon$  ihr Minimum bei  $t = 0$ . Aus der Differenzierbarkeit von  $I$  folgt auch die Differenzierbarkeit der Verkettung mit  $t \mapsto \vartheta_0 + t \cdot \psi$ , also die Differenzierbarkeit von  $\Upsilon$ , somit muss für die Minimalstelle  $\Upsilon'(0) = 0$  gelten.

Nun formen wir um:

$$\begin{aligned} 0 &= \Upsilon'(0) = \left. \frac{dI}{dt}(\vartheta_0 + t\psi) \right|_{t=0} \\ &= \left. \frac{d}{dt} \int_{\Omega} L(x, \vartheta_0(x) + t\psi(x), D\vartheta_0(x) + tD\psi(x)) \, dx \right|_{t=0} \\ &= \int_{\Omega} \left. \frac{dL}{dt}(x, \vartheta_0(x) + t\psi(x), D\vartheta_0(x) + tD\psi(x)) \right|_{t=0} \, dx \\ &= \int_{\Omega} \langle \nabla_y L(x, \vartheta_0(x), D\vartheta_0[x]), \psi(x) \rangle + \langle \nabla_p L(x, \vartheta_0(x), D\vartheta_0[x]), D\psi(x) \rangle \, dx. \end{aligned} \quad (2.10.17)$$

<sup>18</sup>D.h. eine offene, nichtleere und zusammenhängende Menge

<sup>19</sup>Wir verwenden die Einschränkung auf den Rand als Einschränkung der Fortsetzung, d.h.  $\lim_{x \rightarrow x_0} f(x) = g(x_0)$  für alle  $x_0 \in \partial\Omega$ .

## 2. Matrix-Analysis

Ist  $\vartheta_0$  sogar zweimal differenzierbar, dann existiert die Ableitung  $\operatorname{div} \nabla_p L(x, \vartheta_0(x), D\vartheta_0[x])$  und wir können mit partieller Integration für Matrixfelder (Proposition 2.10.2) weiter umformen:

$$\begin{aligned} &= \int_{\Omega} \langle \nabla_y L(x, \vartheta_0(x), D\vartheta_0[x]), \psi(x) \rangle - \langle \operatorname{div} \nabla_p L(x, \vartheta_0(x), D\vartheta_0[x]), \psi(x) \rangle dx \\ &= \int_{\Omega} \langle \nabla_y L(x, \vartheta_0(x), D\vartheta_0[x]) - \operatorname{div} \nabla_p L(x, \vartheta_0(x), D\vartheta_0[x]), \psi(x) \rangle dx. \end{aligned}$$

Da  $\psi \in C_0^\infty(\Omega)$  beliebig gewählt war, liefert uns Lemma 2.10.8 nun

$$\nabla_y L(x, \vartheta_0(x), D\vartheta_0[x]) - \operatorname{div} \nabla_p L(x, \vartheta_0(x), D\vartheta_0[x]) = 0, \quad (2.10.18)$$

was wir zeigen wollten. ■

**Bemerkung 2.10.20.** Eine Funktion  $\vartheta_0$  erfüllt die Euler-Lagrange-Gleichung, wenn sie ein lokales Minimum, ein lokales Maximum oder einen Sattelpunkt der Abbildung  $\vartheta \mapsto I(\vartheta)$  darstellt. Somit ist das Erfüllen der Euler-Lagrange-Gleichung nur ein notwendiges, aber kein hinreichendes Kriterium für eine Funktion, Minimierer von  $I$  zu sein. Nicht jedes  $\vartheta_0$ , das die Euler-Lagrange-Gleichung erfüllt, ist ein Minimierer von  $I$ .

### Wohldefiniertheit von Energiefunktionalen

**Proposition 2.10.21.** Sei  $\mathcal{D} \subseteq \mathbb{R}^{n \times n}$  und  $W: \mathcal{D} \rightarrow \mathbb{R}$  mit

$$|W(F)| \leq \|F\|^p \quad \text{für alle } F \in \mathcal{D} \quad (2.10.19)$$

und ein Gebiet  $\Omega \subseteq \mathbb{R}^n$  gegeben, dann ist für jedes  $\varphi \in W^{k,p}(\Omega)$  mit  $D\varphi(\Omega) \subseteq \mathcal{D}$  das Funktional

$$I(\varphi) := \int_{\Omega} W(D\varphi(x)) dx \quad (2.10.20)$$

wohldefiniert.

*Beweis.* Da  $\varphi \in W^{k,p}(\Omega)$ , sind sowohl  $\varphi \in L^p(\Omega)$  als auch  $D\varphi \in L^p(\Omega, \mathbb{R}^n)$ . Mit Letzterem erhalten wir

$$\int_{\Omega} |W(D\varphi(x))| dx \leq \int_{\Omega} \|D\varphi(x)\|^p dx = \|D\varphi\|_{L^p(\Omega, \mathbb{R}^n)}^p < \infty. \quad \blacksquare$$

**Folgerung 2.10.22.** Sei  $\mathcal{D} \subseteq \mathbb{R}^{n \times n}$  und  $W: \mathcal{D} \rightarrow \mathbb{R}$  mit

$$|W(F)| \leq c_1 + c_2 \|F\|^p \quad \text{für alle } F \in \mathcal{D} \quad (2.10.21)$$

mit Konstanten  $c_1 \in \mathbb{R}$  und  $c_2 \in \mathbb{R}_+$  gegeben. Sei  $\Omega \subset \mathbb{R}^n$  ein beschränktes Gebiet<sup>20</sup>, dann ist für jedes  $\varphi \in W^{k,p}(\Omega)$  mit  $D\varphi(\Omega) \subseteq \mathcal{D}$  das Funktional

$$I(\varphi) := \int_{\Omega} W(D\varphi(x)) dx \quad (2.10.22)$$

wohldefiniert.

*Beweis.* Mit Proposition 2.10.21 folgt  $\int_{\Omega} |W(D\varphi(x))| dx \leq c_1 |\Omega| + c_2 \|D\varphi\|_{L^p(\Omega, \mathbb{R}^n)}^p < \infty$ . ■

---

<sup>20</sup>D.h.  $|\Omega| := \int_{\Omega} 1 dx < \infty$ .

### 2.10.1. Differentiation von $L: \text{GL}^+(n) \rightarrow \mathbb{R}$ mit $L(QX) = L(X)$

In Abschnitt 4.2 auf Seite 101 ff untersuchen wir im Kontext der Modellierung von Elastizitätsgesetzen die Invarianz von Tensorfunktionen  $L: \text{GL}^+(n) \rightarrow \mathbb{R}$  gegenüber Rotationen von links, d.h.  $L(QX) = L(X)$  für alle  $X \in \text{GL}^+(n)$  und  $Q \in \text{SO}(n)$  und ihre Darstellung über symmetrische Funktionen  $\tilde{L}: \text{Sym}^+(n) \rightarrow \mathbb{R}$  mit  $\tilde{L}(X^T X) = L(X)$ . Die Grundlage wollen wir in diesem Abschnitt fernab des Anwendungsbezugs legen.

**Proposition 2.10.23.** Sei  $f: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  mit  $f(X) = X^T X$ . Dann ist

$$Df[X].H = X^T H + H^T X \quad \text{für alle } X, H \in \mathbb{R}^{n \times n}. \quad (2.10.23)$$

*Beweis.*

$$f(X+H) = (X+H)^T(X+H) = X^T X + \underbrace{X^T H + H^T X}_{=Df[X].H} + \underbrace{H^T H}_{=R(H)} \quad \text{mit } \lim_{H \rightarrow 0} \frac{R(H)}{\|H\|} = 0. \quad \blacksquare$$

**Proposition 2.10.24.** Seien  $A, X, H \in \mathbb{R}^{n \times n}$ . Dann gilt

$$\langle A, X^T H + H^T X \rangle = \langle 2X \text{ sym } A, H \rangle. \quad (2.10.24)$$

*Beweis.* Mit der Linearität des Skalarproduktes folgt:

$$\begin{aligned} \langle A, X^T H + H^T X \rangle &= \langle A, H^T X \rangle + \langle A, X^T H \rangle \\ &= \langle AX^T, H^T \rangle + \langle XA, H \rangle \\ &= \langle XA^T, H \rangle + \langle XA, H \rangle \\ &= \langle 2X \text{ sym } A, H \rangle. \quad \blacksquare \end{aligned}$$

**Proposition 2.10.25.** Sei  $L: \text{GL}^+(n) \rightarrow \mathbb{R}$  mit  $L(QX) = L(X)$  für alle  $X \in \text{GL}^+(n)$  und alle  $Q \in \text{SO}(n)$ . Dann existiert eine Abbildung  $\tilde{L}: \text{Sym}^+(n) \rightarrow \mathbb{R}$  mit  $\tilde{L}(X^T X) = L(X)$  für alle  $X \in \text{GL}^+(n)$ . Für diese Abbildung und für alle  $X, H \in \text{GL}^+(n)$  gilt

$$DL[X].H = D\tilde{L}[X^T X].(X^T H + H^T X), \quad (2.10.25)$$

bzw.

$$\langle \nabla L(X), H \rangle = \langle \nabla \tilde{L}(X^T X), X^T H + H^T X \rangle \quad (2.10.26)$$

und

$$\nabla \tilde{L}(X^T X) = \frac{1}{2} X^{-1} \nabla L(X) \in \text{Sym}(n). \quad (2.10.27)$$

*Beweis.* Sei  $X \in \text{GL}^+(n)$  und  $X = RU$  mit  $R \in \text{SO}(n)$  und  $U \in \text{Sym}^+(n)$ . Für beliebiges  $Y \in \text{Sym}^+(n)$  setzen wir  $\tilde{L}(Y) := L(\sqrt{Y})$ , dann ist

$$\tilde{L}(X^T X) = L(\sqrt{X^T X}) = L(U) = L(RU) = L(X).$$

Setzen wir nun  $f: \text{GL}^+(n) \rightarrow \text{Sym}(n)$  mit  $F(X) = X^T X$ , dann ist  $\tilde{L}(f(X)) = L(X)$  und Ableitung mittels der Kettenregel und Anwendung von Proposition 2.10.23 liefert

$$DL[X].H = (D\tilde{L}[X^T X] \cdot Df[X]) \cdot H = D\tilde{L}[X^T X].(X^T H + H^T X),$$

was wir im Skalarprodukt auch als

$$\langle \nabla L(X), H \rangle = \langle \nabla \tilde{L}(X^T X), X^T H + H^T X \rangle$$

schreiben können. Mit Proposition 2.10.24 wird dies zu

$$\langle \nabla L(X), H \rangle = \langle 2X \text{ sym } \nabla \tilde{L}[X^T X], H \rangle,$$

also  $\nabla L[X] = 2X \nabla \tilde{L}[X^T X]$ , da die Ableitung eines Skalarfeldes auf einem symmetrischen Argument selbst symmetrisch ist. Umstellen liefert schließlich die letzte Gleichung.  $\blacksquare$

**Folgerung 2.10.26.** Sei  $L: \text{GL}^+(n) \rightarrow \mathbb{R}$  mit  $L(QX) = L(X)$  für alle  $X \in \text{GL}^+(n)$  und alle  $Q \in \text{SO}(n)$ . Dann ist

$$(X^{-1}DL[X])^T = X^{-1}DL[X] \quad \text{für alle } X \in \text{GL}^+(n). \quad (2.10.28)$$

## 2.11. Konvexität

**Definition 2.11.1.** Sei  $\mathcal{M}$  eine Teilmenge eines reellen Vektorraums. Wir bezeichnen  $\mathcal{M}$  als *konvex*, wenn

$$tx + (1-t)y \in \mathcal{M} \quad \text{für alle } x, y \in \mathcal{M} \text{ und alle } t \in [0, 1] \quad (2.11.1)$$

gilt, also wenn die Verbindungsstrecke zweier Punkte aus  $\mathcal{M}$  selbst vollständig in  $\mathcal{M}$  enthalten ist.

**Definition 2.11.2.** Sei  $\mathcal{M}$  eine konvexe Teilmenge eines reellen Vektorraums. Wir bezeichnen eine Funktion  $f: \mathcal{M} \rightarrow \mathbb{R}$  als *konvex*, wenn

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \quad \text{für alle } x, y \in \mathcal{M} \text{ und alle } t \in [0, 1] \quad (2.11.2)$$

gilt, also wenn für die Verbindungsstrecke zwischen beliebigen Stellen  $x, y \in \mathcal{M}$  der zugehörige Graph unterhalb der Verbindung  $(x|f(x))$  mit  $(y|f(y))$  liegt.

**Proposition 2.11.3** (Konvexität stetig differenzierbarer Funktionen). Sei  $\mathcal{M}$  eine offene, konvexe Teilmenge eines reellen Vektorraums und  $f: \mathcal{M} \rightarrow \mathbb{R}$  eine stetig differenzierbare Funktion. Dann ist äquivalent:

i)  $f$  ist konvex.

ii) Für jeden Punkt  $y \in \mathcal{M}$  liegt die Tangentialebene an  $f$  in  $y$  unterhalb des Graphen, d.h. es gilt

$$f(x) - f(y) \geq \langle \nabla f(y), x - y \rangle \quad \text{für alle } x, y \in \mathcal{M}. \quad (2.11.3)$$

iii) Der Gradient von  $f$  ist monoton wachsend, d.h. es gilt

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0 \quad \text{für alle } x, y \in \mathcal{M}. \quad (2.11.4)$$

Im Falle von  $f \in C^2(\mathcal{M})$  sind i) – iii) zudem äquivalent zu:

iv) Die zweite Ableitung ist positiv semidefinit, d.h. es gilt

$$D^2f[x] \cdot (h, h) \geq 0 \quad \text{für alle } x \in \mathcal{M} \text{ und alle } h \in \text{span}(\mathcal{M}). \quad (2.11.5)$$

*Beweis.* Siehe Evans [26]. ■

Direkt aus der Linearität der Ableitung folgt:

**Proposition 2.11.4.** Sei  $\mathcal{M}$  eine offene, konvexe Teilmenge eines reellen Vektorraums und  $f, g: \mathcal{M} \rightarrow \mathbb{R}$  zwei konvexe Funktionen. Dann ist auch die Summe  $f + g$  konvex.

Ist die äußere Funktion monoton wachsend, so ist auch die Verkettung zweier konvexer Funktionen konvex:

**Proposition 2.11.5.** Für  $\mathcal{D} \subseteq \mathbb{R}$  sei  $f: \mathcal{D} \rightarrow \mathbb{R}$  eine konvexe und monoton wachsende Funktion,  $\mathcal{M}$  eine offene, konvexe Teilmenge eines reellen Vektorraums und  $g: \mathcal{M} \rightarrow \mathcal{D}$  eine konvexe Funktion. Dann ist die Verkettung  $f \circ g: \mathcal{M} \rightarrow \mathbb{R}$  konvex.

**Satz 2.11.6.** Sei  $f \in C^2(\mathbb{R}_+)$ . Dann ist die Funktion  $\text{Sym}^+(n) \rightarrow \mathbb{R}$  mit  $X \mapsto f(\det X)$  genau dann konvex, wenn

$$f''(s) + \frac{n-1}{ns} f'(s) \geq 0 \quad \text{und} \quad f'(s) \leq 0 \quad \text{für alle } s \in \mathbb{R}_+ \quad (2.11.6)$$

erfüllt ist.

## 2. Matrix-Analysis

*Beweis.* Siehe Lehmich, Neff und Lankeit [44]. ■

**Folgerung 2.11.7.** Die Funktion  $\widehat{W}: \text{Sym}^+(n) \rightarrow \mathbb{R}$  mit  $\widehat{W}(C) = -\text{tr}(\log C)$  ist konvex.

*Beweis.* Mit  $\text{tr} \log = \log \det$  müssen wir nur noch Bedingung (5.0.1) für  $f(s) = -\log(s)$  prüfen. Es ist

$$f'(s) = -\frac{1}{s} \leq 0 \quad \text{für alle } s \in \mathbb{R}_+$$

und

$$f''(s) = \frac{1}{s^2} \quad \text{und damit} \quad f''(s) + \frac{n-1}{ns} f'(s) = \frac{1}{ns^2} \geq 0 \quad \text{für alle } s \in \mathbb{R}_+. \quad \blacksquare$$

**Lemma 2.11.8.** Sei  $\alpha \in \mathbb{R}_+$  und  $\tilde{w}: \mathbb{R}_+ \rightarrow \mathbb{R}$  eine zweimal stetig differenzierbare Funktion. Dann ist die Funktion  $\text{Sym}^+(n) \rightarrow \mathbb{R}$  mit

$$X \mapsto \alpha \text{tr}(T_{\tilde{w}}(X)) \tag{2.11.7}$$

genau dann konvex, wenn  $\tilde{w}$  konvex ist.

Bildet  $\tilde{w}$  ausschließlich in die positiven Zahlen ab und ist konvex, ist sogar

$$X \mapsto \alpha \left( \text{tr}(T_{\tilde{w}}(X)) \right)^2 \tag{2.11.8}$$

konvex.

Ist  $\tilde{w}$  linear, dann ist  $X \mapsto \alpha \text{tr}(T_{\tilde{w}}(X))$  sogar für alle  $\alpha \in \mathbb{R}$  konvex.

**Bemerkung 2.11.9.** Aus diesem Lemma lässt sich ebenso wie mit Satz 2.11.6 die Konvexität von  $\widehat{W}: \text{Sym}^+(n) \rightarrow \mathbb{R}$  mit  $\widehat{W}(C) = -\text{tr}(\log C) = \text{tr}(T_{-\log}(C))$  belegen.

*Beweis.* Mit Proposition 2.5.34 ist  $D \alpha \text{tr}(T_{\tilde{w}}[X]) = \alpha T_{\tilde{w}'}(X) = T_{\alpha \tilde{w}'}(X)$ . Also ist  $X \mapsto D \text{tr}(T_{\tilde{w}}[X])$  genau dann monoton wachsend, wenn  $x \mapsto \alpha \tilde{w}'(x)$  monoton wachsend ist. Letzteres ist für positives  $\alpha$  äquivalent zum monotonen Wachstum von  $x \mapsto \tilde{w}'(x)$ . Ergo ist  $X \mapsto \alpha \text{tr}(T_{\tilde{w}}(X))$  genau dann konvex, wenn  $x \mapsto \tilde{w}(x)$  konvex ist. Siehe hierzu auch Davis [22].

Ist  $\tilde{w}$  konvex und positiv, so ist  $X \mapsto \text{tr}(T_{\tilde{w}}(X))$  konvex und positiv<sup>21</sup>. Verkettung mit der auf  $\mathbb{R}_+$  konvexen Funktion  $x \mapsto \alpha x^2$  ergibt nach Proposition 2.11.5 eine konvexe Funktion.

Ist  $\tilde{w} = ax + b$ , so ist  $D \alpha \text{tr}(T_{\tilde{w}}(X)) = \alpha a \cdot \mathbb{1}$  und damit  $\langle \nabla_X \alpha \text{tr}(T_{\tilde{w}}(X)) - \nabla_X \alpha \text{tr}(T_{\tilde{w}}(Y)), X - Y \rangle = \langle 0, X - Y \rangle = 0$ . ■

**Folgerung 2.11.10.** Sei  $\widehat{w} \in C^1(\mathbb{R}_+)$  eine konvexe Funktion mit  $\widehat{w}'(1) = 0$  und  $\widehat{f} \in C^2(\mathbb{R}_+)$  erfülle die Bedingungen aus Satz 2.11.6 und  $\widehat{f}'(1) = 0$ . Dann ist die durch

$$\widehat{W}(C) = \widehat{w}(\lambda_1^2) + \widehat{w}(\lambda_2^2) + \widehat{w}(\lambda_3^2) + \widehat{f}(\det C) \quad \text{für alle } C \in \text{Sym}^+(3) \tag{2.11.9}$$

definierte Funktion  $\widehat{W} \in C^1(\text{Sym}^+(3))$  konvex. Zudem erfüllt sie  $D \widehat{W}(\mathbb{1}) = 0$ . Hierbei sind  $\lambda_1^2, \lambda_2^2, \lambda_3^2 \in \mathbb{R}_+$  die Eigenwerte von  $C$ .

**Folgerung 2.11.11.** Seien  $\alpha, \zeta \in \mathbb{R}$ , sowie  $\beta, \gamma, \delta \in \mathbb{R}_+ \cup \{0\}$  und sei  $\widehat{W}: \text{Sym}^+ \rightarrow \mathbb{R}$  mit

$$\widehat{W}(C) = \alpha \text{tr}(C) + \beta \text{tr}(C)^2 + \gamma \text{tr}(C^2) - \delta \log \det C + \zeta. \tag{2.11.10}$$

Dann ist  $\widehat{W}$  konvex.

---

<sup>21</sup>Die Spur entspricht bei symmetrischen Matrizen der Summe aller Eigenwerte. Sind diese das Bild einer positiven Funktion, so ist die Spur selbst positiv.

## 2. Matrix-Analysis

**Lemma 2.11.12.** Sei  $\widehat{W}$  gegeben wie in Folgerung 2.11.11. Zudem sei  $\Omega \subset \mathbb{R}^n$  ein Gebiet und  $c^+ \in \mathbb{R}_+$  eine Konstante. Dann ist das Energiefunktional

$$I(\varphi) := \int_{\Omega} W(\nabla\varphi(x)) \, dx \quad (2.11.11)$$

für alle  $\varphi \in W^{1,4}(\Omega, \mathbb{R}^n)$ , die  $\det \nabla\varphi(x) \geq c^+$  für alle  $x \in \Omega$  erfüllen, wohldefiniert.

*Beweis.* Es ist

$$\begin{aligned} \operatorname{tr}(C) &= \operatorname{tr}(\nabla\varphi(x)^T \nabla\varphi(x)) = \|\nabla\varphi(x)\|^2 \leq 1 + \|\nabla\varphi(x)\|^4, \\ \operatorname{tr}(C)^2 &= \operatorname{tr}(\nabla\varphi(x)^T \nabla\varphi(x))^2 = \|\nabla\varphi(x)\|^4, \\ \operatorname{tr}(C^2) &\leq \operatorname{tr}(C)^2 = \|\nabla\varphi(x)\|^4, \\ -\log \det C &= -2 \log \underbrace{\det(\nabla\varphi(x))}_{\geq c^+} \leq -2 \log c^+ \end{aligned}$$

und damit

$$\begin{aligned} W(F) &= \widehat{W}(C) = \alpha \operatorname{tr}(C) + \beta \operatorname{tr}(C)^2 + \gamma \operatorname{tr}(C^2) - \gamma \log \det C + \zeta \\ &\leq \alpha (1 + \|\nabla\varphi(x)\|^4) + \beta \|\nabla\varphi(x)\|^4 + \gamma \|\nabla\varphi(x)\|^4 - 2 \delta \log c^+ + \zeta \\ &= (\alpha + \beta + \gamma) \|\nabla\varphi(x)\|^4 + (\alpha - 2 \delta \log c^+ + \zeta). \end{aligned}$$

Mit Folgerung 2.10.22 folgt nun die Wohldefiniertheit des Energiefunktionals für alle  $\varphi \in W^{1,4}(\Omega, \mathbb{R}^n)$ . ■

## 2.12. Majorization und Schur-Konvexität

Das Konzept der *Majorization* ist von großer Wichtigkeit für die Sum-of-Squared-Logarithms-Inequality und verwandte Ungleichungen (siehe Kapitel 3). Im Folgenden stellen wir die grundlegenden Definitionen und einige fundamentale Eigenschaften der Majorization vor.

Für einen größeren Überblick siehe Marshall, Olkin und Arnold [45]. Der Begriff *majorisiert* folgt Hardy, Littlewood und Pólya [34], der Begriff *schwach majorisiert* Tomić [76] und Weyl [81].

**Definition 2.12.1** (Majorization). Sei  $x, y \in \mathbb{R}^n$ .

- i) Wir bezeichnen  $x$  als *schwach majorisiert von unten* durch  $y$  oder *submajorisiert*, geschrieben  $x \prec_w y$ , wenn

$$\sum_{i=1}^k x_i^\downarrow \leq \sum_{i=1}^k y_i^\downarrow \quad \text{für alle } k \in \{1, \dots, n\}. \quad (2.12.1)$$

- ii) Wir bezeichnen  $x$  als *schwach majorisiert von oben* durch  $y$  oder *supermajorisiert*, geschrieben  $x \prec^w y$ , wenn

$$\sum_{i=k}^n x_i^\downarrow \geq \sum_{i=k}^n y_i^\downarrow \quad \text{für alle } k \in \{1, \dots, n\}. \quad (2.12.2)$$

*Bemerkung:* Es ist einfach zu sehen, dass  $\sum_{i=k}^n x_i^\downarrow \geq \sum_{i=k}^n y_i^\downarrow$  genau dann wenn  $\sum_{i=1}^k x_i^\uparrow \geq \sum_{i=1}^k y_i^\uparrow$  für alle  $k \in \{1, \dots, n\}$ .

- iii) Wir bezeichnen  $x$  als *majorisiert* durch  $y$ , geschrieben  $x \prec y$ , wenn  $x \prec_w y$  und  $x \prec^w y$ . Dies ist genau dann der Fall wenn  $x \prec_w y$  oder  $x \prec^w y$  gilt und zusätzlich  $e_1(x) = e_1(y)$ , also  $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ .



## 2. Matrix-Analysis

Für ein besseres Verständnis: Seien  $x_1 \geq x_2 \geq \dots \geq x_n > 0$  und  $y_1 \geq y_2 \geq \dots \geq y_n > 0$ . Dann wird  $x$  schwach durch  $y$  von unten majorisiert ( $x \prec_w y$ ), wenn

$$\begin{aligned} x_1 &\leq y_1 \\ x_1 + x_2 &\leq y_1 + y_2 \\ &\vdots \\ x_1 + x_2 + \dots + x_{n-1} &\leq y_1 + y_2 + \dots + y_{n-1} \\ x_1 + x_2 + \dots + x_{n-1} + x_n &\leq y_1 + y_2 + \dots + y_{n-1} + y_n \end{aligned} \quad (2.12.3)$$

und wir haben  $x \prec y$ , wenn Gleichheit in der letzten Zeile gilt.

**Lemma 2.12.2.** Seien  $x_1 \geq x_2 \geq \dots \geq x_n > 0$  und  $y_1 \geq y_2 \geq \dots \geq y_n > 0$ . Dann ist die Bedingung

$$\sum_{i=1}^k x_i \leq \sum_{i=1}^k y_i \quad \text{für alle } k \in \{1, \dots, n-1\} \quad \text{und} \quad \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \quad (2.12.4)$$

äquivalent zu

$$\sum_{i=1}^k x_i - \sum_{i=k+1}^n x_i \leq \sum_{i=1}^k y_i - \sum_{i=k+1}^n y_i \quad \text{für alle } k \in \{0, 1, \dots, n\}. \quad (2.12.5)$$

*Beweis.* Die Bedingung

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i \quad (2.12.6)$$

ist äquivalent zu  $\sum_{i=1}^n x_i \leq \sum_{i=1}^n y_i$  und  $\sum_{i=1}^n x_i \geq \sum_{i=1}^n y_i$ , deshalb

$$\sum_{i=1}^n x_i - \sum_{i=n+1}^n x_i \leq \sum_{i=1}^n y_i - \sum_{i=n+1}^n y_i \quad \text{und} \quad \sum_{i=1}^0 x_i - \sum_{i=0+1}^n x_i \leq \sum_{i=1}^0 y_i - \sum_{i=0+1}^n y_i,$$

woraus

$$\sum_{i=1}^k x_i - \sum_{i=k+1}^n x_i \leq \sum_{i=1}^k y_i - \sum_{i=k+1}^n y_i \quad \text{für } k \in \{0, n\} \quad (2.12.7)$$

folgt. Wenn nun (2.12.6) oder (2.12.7) erfüllt ist, dann erhalten wir für  $k \in \{1, \dots, n-1\}$ :

$$\begin{aligned} \sum_{i=1}^k x_i \leq \sum_{i=1}^k y_i &\Leftrightarrow 2 \sum_{i=1}^k x_i - \sum_{i=1}^n x_i \leq 2 \sum_{i=1}^k y_i - \sum_{i=1}^n y_i \\ &\Leftrightarrow \sum_{i=1}^k x_i - \sum_{i=k+1}^n x_i \leq \sum_{i=1}^k y_i - \sum_{i=k+1}^n y_i. \quad \blacksquare \end{aligned}$$

**Proposition 2.12.3.** Seien  $x, y \in \mathbb{R}_+^n$ . Dann ist  $x \prec y$  genau dann, wenn

$$\sum_{i=1}^n |t - x_i| \leq \sum_{i=1}^n |t - y_i| \quad \text{für alle } t \in \mathbb{R}. \quad (2.12.8)$$

*Beweis.* Seien o.B.d.A.  $x_1 \geq x_2 \geq \dots \geq x_n > 0$  und  $y_1 \geq y_2 \geq \dots \geq y_n > 0$ .

Nehmen wir zunächst an, dass  $x \prec y$ , bzw. mit Lemma 2.12.2:

$$\sum_{i=1}^k x_i - \sum_{i=k+1}^n x_i \leq \sum_{i=1}^k y_i - \sum_{i=k+1}^n y_i \quad \text{für alle } k \in \{0, 1, \dots, n\}.$$

## 2. Matrix-Analysis

Sei zudem  $t \in \mathbb{R}$  gegeben. Wir wählen  $k \in \{0, \dots, n\}$  so, dass  $t > x_1$  gilt, falls  $k = 0$ , dass  $x_n \geq t$  gilt, falls  $k = n$ , ansonsten gilt  $x_k \geq t > x_{k+1}$ . Dann ist  $x_i \geq t$  für alle  $i \leq k$  und  $x_i < t$  für alle  $i > k$ , deshalb gilt

$$|x_i - t| = x_i - t \quad \text{für alle } i \leq k \quad \text{und} \quad |x_i - t| = t - x_i \quad \text{für alle } i > k.$$

Wir folgern mit der Dreiecksungleichung:

$$\begin{aligned} \sum_{i=1}^k x_i - \sum_{i=k+1}^n x_i &\leq \sum_{i=1}^k y_i - \sum_{i=k+1}^n y_i \Leftrightarrow \sum_{i=1}^k (x_i - t) + \sum_{i=k+1}^n (t - x_i) \leq \sum_{i=1}^k (y_i - t) + \sum_{i=k+1}^n (t - y_i) \\ \Leftrightarrow \sum_{i=1}^k \underbrace{|x_i - t|}_{\geq 0} + \sum_{i=k+1}^n \underbrace{|t - x_i|}_{\geq 0} &\leq \sum_{i=1}^k (y_i - t) + \sum_{i=k+1}^n (t - y_i) \\ \Leftrightarrow \sum_{i=1}^k |x_i - t| + \sum_{i=k+1}^n |t - x_i| &\leq \left| \sum_{i=1}^k (y_i - t) + \sum_{i=k+1}^n (t - y_i) \right| \\ \Rightarrow \sum_{i=1}^k |x_i - t| + \sum_{i=k+1}^n |t - x_i| &\leq \sum_{i=1}^k |y_i - t| + \sum_{i=k+1}^n |t - y_i| \Leftrightarrow \sum_{i=1}^k |x_i - t| \leq \sum_{i=1}^k |y_i - t|. \end{aligned} \quad (2.12.9)$$

Nun nehmen wir andersherum an, dass

$$\sum_{i=1}^n |t - x_i| \leq \sum_{i=1}^n |t - y_i| \quad \text{für alle } t \in \mathbb{R}.$$

Sei zudem  $k \in \{0, \dots, n\}$  gegeben. Wir wählen  $t \in \mathbb{R}$  so, dass  $t > y_1$ , falls  $k = 0$ , dass  $y_n \geq t$  gilt, falls  $k = n$ , ansonsten gilt  $y_k \geq t > y_{k+1}$ . Dann ist  $y_i > t$  für alle  $i \leq k$  und  $y_i < t$  für alle  $i > k$ , deshalb gilt

$$|y_i - t| = y_i - t \quad \text{für alle } i \leq k \quad \text{und} \quad |y_i - t| = t - y_i \quad \text{für alle } i > k.$$

Mit der Ungleichung  $|\xi| \geq \xi$  für alle  $\xi \in \mathbb{R}$  folgern wir:

$$\begin{aligned} \sum_{i=1}^n |x_i - t| \leq \sum_{i=1}^n |y_i - t| &\Leftrightarrow \sum_{i=1}^k |x_i - t| + \sum_{i=k+1}^n |t - x_i| \leq \sum_{i=1}^k |y_i - t| + \sum_{i=k+1}^n |t - y_i| \\ \Leftrightarrow \sum_{i=1}^k |x_i - t| + \sum_{i=k+1}^n |t - x_i| &\leq \sum_{i=1}^k (y_i - t) + \sum_{i=k+1}^n (t - y_i) \\ \Rightarrow \sum_{i=1}^k (x_i - t) + \sum_{i=k+1}^n (t - x_i) &\leq \sum_{i=1}^k (y_i - t) + \sum_{i=k+1}^n (t - y_i) \quad (2.12.10) \\ \Leftrightarrow \sum_{i=1}^k x_i - \sum_{i=k+1}^n x_i &\leq \sum_{i=1}^k y_i - \sum_{i=k+1}^n y_i. \quad \blacksquare \end{aligned}$$

Das nachstehende Lemma, welches grundlegende Eigenschaften der sogenannten *schwachen logarithmischen Majorization* zeigt, folgt direkt aus den Logarithmusgesetzen und der Monotonie des Logarithmus.

**Lemma 2.12.4** (Logarithmische Majorization). *Seien  $x, y \in \mathbb{R}_+^n$ . Dann*

$$\begin{aligned} \log x \prec_w \log y &\quad \text{genau dann wenn} \quad \prod_{i=1}^k x_i^\downarrow \leq \prod_{i=1}^k y_i^\downarrow \quad \text{für alle } k \in \{1, \dots, n\}, \\ \log x \prec^w \log y &\quad \text{genau dann wenn} \quad \prod_{i=k}^n x_i^\downarrow \geq \prod_{i=k}^n y_i^\downarrow \quad \text{für alle } k \in \{1, \dots, n\}, \\ \log x \prec \log y &\quad \text{genau dann wenn} \quad \log x \prec_w \log y \quad \text{und} \quad e_n(x) = e_n(y), \end{aligned}$$

wobei wir  $\log z := (\log z_1, \log z_2, \dots, \log z_n)$  für  $z \in \mathbb{R}_+^n$  abkürzend schreiben.

## 2. Matrix-Analysis

**Proposition 2.12.5.** Seien  $x, y \in \mathbb{R}^n$  mit  $x \prec_w y$  (bzw.  $x \prec^w y$ ). Dann existiert ein  $z \in \mathbb{R}^n$  mit  $z \prec y$  und  $x \leq z$  (bzw.  $z \prec y$  und  $x \geq z$ ).

*Beweis.* Siehe Marshall, Olkin, Arnold [45, A.9. und A.9.a. S. 177]. ■

**Definition 2.12.6** (Schur-Konvexität). Wir nennen eine Funktion  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$  *Schur-konvex* (bzw. *Schur-konkav*), wenn

$$\varphi(x) \leq \varphi(y) \quad (\text{bzw. } \varphi(x) \geq \varphi(y)) \quad \text{für alle } x, y \in \mathbb{R}^n \text{ mit } x \prec y \quad (2.12.11)$$

gilt.

**Proposition 2.12.7.** Sei  $\varphi: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  Schur-konvex. Wenn  $\varphi$  zudem monoton wachsend (bzw. monoton fallend) ist, dann

$$\varphi(x) \leq \varphi(y) \quad \text{für alle } x, y \in D \text{ mit } x \prec_w y \text{ (bzw. } x \prec^w y). \quad (2.12.12)$$

*Beweis.* Aufgrund von Proposition 2.12.5 existiert ein  $z \in D$  mit  $z \prec y$  und  $x \leq z$  ( $x \geq z$ ). Weil  $g$  monoton wachsend (fallend) ist,  $f(x) \leq f(z)$ . Und mit der Schur-Konvexität von  $g$  folgt  $f(z) \leq f(y)$ . Siehe auch Marshall, Olkin, Arnold [45, A.8. S. 87]. ■

**Proposition 2.12.8.** Seien  $g: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x, y \in \mathbb{R}_+^n$ . Falls

- i)  $g$  konvex ist und  $x \prec y$ , oder
- ii)  $g$  konvex und monoton wachsend ist und  $x \prec_w y$ , oder
- iii)  $g$  konvex und monoton fallend ist und  $x \prec^w y$ ,

dann ist

$$\sum_{i=1}^n g(x_i) \leq \sum_{i=1}^n g(y_i). \quad (2.12.13)$$

**Bemerkung 2.12.9.** Die Funktion  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $\varphi(x) = \sum_{i=1}^n g(x_i)$  ist Schur-konvex, falls  $g: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x, y \in \mathbb{R}_+^n$  eine konvexe Funktion ist. Ist  $g$  zudem monoton wachsend (bzw. monoton fallend), dann ist  $\varphi$  monoton wachsend (bzw. monoton fallend).

*Beweis.* Siehe Mitrinović, Pečarić, Fink [50, S. 686] und Marshall, Olkin, Arnold [45, C.1.a. C.1.b. S. 92]. ■

**Proposition 2.12.10.** Seien  $x, y \in \mathbb{R}_+^n$ . Dann

$$\begin{aligned} \log x \prec_w \log y & \text{ impliziert } x \prec_w y \\ \text{und} & \\ x \prec^w y & \text{ impliziert } \log x \prec^w \log y. \end{aligned}$$

*Beweis.* Siehe Marshall, Olkin, Arnold [45, A.2.b A.2.c S. 168]. ■

## 2.13. Nullstellen und Koeffizienten von Polynomen

**Definition 2.13.1.** Wir bezeichnen für  $k \in \{1, \dots, n\}$  das  $k$ -te elementarsymmetrische Polynom als die Abbildung  $e_k: \mathbb{C}^n \rightarrow \mathbb{C}$  mit

$$e_k(z) = \sum_{1 \leq i_1 < \dots < i_k \leq n} z_{i_1} z_{i_2} \dots z_{i_k}. \quad (2.13.1)$$

Für  $z \in \mathbb{C}^n$  bezeichnen wir mit  $z^{(i)}$  den Vektor, der aus  $z$  hervorgeht, wenn wir die  $i$ -te Komponente streichen. Es können auch mehrere Komponenten gestrichen werden. So ist für  $z = (z_1, \dots, z_n)$  damit  $z^{(i)} = (z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_n)$  und  $z^{(ij)} = (z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_{j-1}, z_{j+1}, \dots, z_n)$ , usw.

## 2. Matrix-Analysis

**Lemma 2.13.2.** Die elementarsymmetrischen Polynome lassen sich rekursiv formulieren:

$$e_k(z) := z_i e_{k-1}(z^{(i)}) + e_k(z^{(i)}), \quad (2.13.2)$$

wobei  $e_0(z) = 1$  und  $e_{n+1}(z) = 0$  gilt. ■

**Bemerkung 2.13.3.** Mit der rekursiven Beschreiben lassen sich ganz einfach die partiellen Ableitungen der elementarsymmetrischen Polynome angeben:

$$\frac{\partial e_k}{\partial x_i}(z) = e_{k-1}(z^{(i)}). \quad (2.13.3)$$

**Definition 2.13.4.** Wir bezeichnen eine Menge  $D \subseteq \mathbb{C}^n$  als *symmetrisch*, wenn für jedes Element  $(x_1, \dots, x_n) \in D$  und jede Permutation  $(i_1, \dots, i_n)$  auch  $(x_{i_1}, \dots, x_{i_n}) \in D$  gilt, d.h. mit jedem Element ist auch jede seiner Umordnungen in der Menge.

**Definition 2.13.5.** Sei  $D \subseteq \mathbb{C}^n$ . Wir bezeichnen eine Funktion  $f: D \rightarrow \mathbb{C}$  als *symmetrisch*, wenn sie invariant gegenüber Komponentenvertauschungen ist, das heißt, wenn für jede Permutation  $(i_1, \dots, i_n)$  gilt

$$f(z_{i_1}, \dots, z_{i_n}) = f(z_1, \dots, z_n). \quad (2.13.4)$$

**Beispiel 2.13.6.** Die elementarsymmetrischen Polynome  $e_1, \dots, e_n: \mathbb{C}^n \rightarrow \mathbb{C}^n$  sind symmetrische Funktionen.

**Beispiel 2.13.7.** Sei  $D \subseteq \mathbb{C}$  und  $\phi: D \rightarrow \mathbb{C}$ . Dann ist  $f: D^n \rightarrow \mathbb{C}$  mit  $f(z_1, \dots, z_n) := \sum_{i=1}^n \phi(z_i)$  eine symmetrische Funktion.

**Definition 2.13.8.** Wir bezeichnen  $z_1, \dots, z_n \in \mathbb{C}$  als *Nullstellen* und  $e_1, \dots, e_n \in \mathbb{C}$  als *Koeffizienten*, wenn es ein normiertes Polynom

$$\begin{aligned} p &= (X - z_1)(X - z_2) \dots (X - z_n) \\ &= X^n - e_1 X^{n-1} + e_2 X^{n-2} + \dots + (-1)^n e_n \end{aligned} \quad (2.13.5)$$

gibt. Alle paarweise verschiedenen  $z_i$  nennen wir *einfache Nullstellen*, die übrigen nennen wir *mehrfache Nullstellen*.

**Bemerkung 2.13.9.** Die Abbildung der Nullstellen auf die zugehörigen Koeffizienten ist genau die Abbildung der elementarsymmetrischen Polynome, d.h.  $e_k = e_k(z_1, \dots, z_n)$  für gegebene  $(z_1, \dots, z_n)$ . So sind aufgrund der Symmetrie die Abbildungen  $e_k: \mathbb{C}^n \rightarrow \mathbb{C}^n$  nicht injektiv. So ist zum Beispiel  $e_1(1, 2) = e_1(2, 1)$ .

**Lemma 2.13.10.** Sei  $\mathcal{U} \subset \mathbb{C}^{n\uparrow}$  eine Menge mit eindeutig sortierten Komponenten<sup>22</sup>. Dann ist die Funktion  $e: \mathcal{U} \rightarrow \mathbb{C}^n$ , welche den geordneten Nullstellenvektor  $(z_1, \dots, z_n)$  auf den Koeffizientenvektor  $(e_1, \dots, e_n)$  abbildet, d.h.

$$e(z_1, \dots, z_n) := (e_1(z_1, \dots, z_n), \dots, e_n(z_1, \dots, z_n)) \quad (2.13.6)$$

homöomorph, d.h. bijektiv und stetig, mit stetiger Umkehrfunktion  $z: \mathbb{C}^n \rightarrow \mathbb{C}^{n\uparrow}$ .

*Beweis.* Die Stetigkeit der elementarsymmetrischen Polynome ist klar. Nach dem Fundamentalsatz der Algebra gibt es zu jedem Koeffizientenvektor eindeutig eine Menge von  $n$  Nullstellen, unter Sortierung eindeutig einen Nullstellenvektor aus  $\mathbb{C}^{n\uparrow}$ . Ein Beweis der Stetigkeit dieser Nullstellenabbildung findet sich in Cucker und Corbalan [20]. ■

**Bemerkung 2.13.11.** Für  $n > 4$  lässt sich  $z$  nicht explizit angeben<sup>23</sup>.

<sup>22</sup>So können wir mit  $\mathbb{C}^{n\uparrow} := \{z \in \mathbb{C}^n \mid \operatorname{Re}(z_1) \geq \dots \geq \operatorname{Re}(z_n), \operatorname{Re} z_i = \operatorname{Re} z_{i+1} \Rightarrow \operatorname{Im} z_i \geq \operatorname{Im} z_{i+1} \forall i \in \{1, \dots, n-1\}\}$  die Menge der geordneten Vektoren komplexer Zahlen definieren.

<sup>23</sup>für  $n = 2$  haben wir die  $p$ - $q$ -Formel, für  $n \in \{3, 4\}$  sind dies die Cardanischen Formeln

## 2. Matrix-Analysis

**Bemerkung 2.13.12.** Aufgrund des bijektiven Zusammenhangs werden wir in dieser Arbeit, wenn es aus dem Kontext ersichtlich ist, verkürzt  $e_k$  für das  $k$ -te elementarsymmetrische Polynom, ausgewertet an  $(z_1, \dots, z_n)$ , schreiben und es als den  $k$ -ten Koeffizienten bezeichnen, d.h.

$$e_k := e_k(z_1, \dots, z_n) \quad \text{für alle } k \in \{1, \dots, n\}, \quad (2.13.7)$$

und verkürzt  $z_i$  für die  $i$ -te Komponente der Umkehrfunktion  $z$ , ausgewertet an  $(e_1, \dots, e_n)$ , schreiben und es als die  $i$ -te Nullstelle bezeichnen, d.h.

$$z_i := z_i(e_1, \dots, e_n) \quad \text{für alle } i \in \{1, \dots, n\}. \quad (2.13.8)$$

In der nachstehenden Folgerung wenden wir diese Schreibweise das erste Mal an:

**Folgerung 2.13.13.** Für  $D \subseteq \mathbb{C}^n$  sei  $f: D \rightarrow \mathbb{C}$  eine symmetrische Funktion. Dann lässt sich  $f$  als Funktion in den elementarsymmetrischen Polynomen ausdrücken. Das heißt, es existiert eine Funktion  $\mathcal{F}: \mathbb{C}^n \rightarrow \mathbb{C}$  mit

$$f(z_1, \dots, z_n) = \mathcal{F}(e_1, \dots, e_n) \quad \text{für alle } (z_1, \dots, z_n) \in D. \quad (2.13.9)$$

*Beweis.* Sei  $(z_{i_1}, \dots, z_{i_n}) \in \mathbb{C}^{n\uparrow}$  die Umsortierung der Nullstellen zu einem geordneten Vektor. Setzen wir  $\mathcal{F} := f \circ z$ , dann ist

$$\mathcal{F}(e_1, \dots, e_n) = f(z(e_1, \dots, e_n)) = f(z_{i_1}, \dots, z_{i_n}) \stackrel{f \text{ sym.}}{=} f(z_1, \dots, z_n). \quad \blacksquare$$

**Lemma 2.13.14.** Die Abbildung der Nullstellenvektoren auf die Koeffizientenvektoren  $e: \mathbb{C}^{n\uparrow} \rightarrow \mathbb{C}^n$  ist differenzierbar. Die partiellen Ableitungen der Komponenten lauten

$$\frac{\partial e_k}{\partial z_i}(z) = \sum_{\substack{1 \leq i_1 < \dots < i_k \leq n \\ i_j \neq i}} z_{i_1} z_{i_2} \dots z_{i_{k-1}}.$$

*Beweis.* Zur Ableitung der Koeffizienten  $e_k$  nach den Nullstellen  $z_i$ : Wie in der Bemerkung zur rekursiven Darstellung der elementarsymmetrischen Polynome schon angegeben, lautet die Ableitung

$$\frac{\partial}{\partial z_i} e_k(z) = e_{k-1}(z^{(i)}) = \sum_{\substack{1 \leq i_1 < \dots < i_k \leq n \\ i_j \neq i}} z_{i_1} z_{i_2} \dots z_{i_{k-1}}. \quad \blacksquare$$

**Lemma 2.13.15.** Die Abbildung der Koeffizientenvektoren auf die Nullstellenvektoren  $z: \mathbb{C}^n \rightarrow \mathbb{C}^{n\uparrow}$  ist differenzierbar für alle Koeffizientenvektoren  $(e_1, \dots, e_n)$  mit paarweise unterschiedlichen Nullstellen  $(z_1, \dots, z_n)$ . Die partiellen Ableitungen der Komponenten lauten

$$\frac{\partial z_i}{\partial e_k}(e) = \frac{(-1)^{k+1} z_i^{n-k}}{\prod_{\substack{j=1 \\ j \neq i}}^n (z_i - z_j)} \quad \text{für alle } (e_1, \dots, e_n) \text{ mit } z_j \neq z_i, \text{ wenn } j \neq i. \quad (2.13.10)$$

*Beweis.* Zur Ableitung der Nullstellen  $z_i$  nach den Koeffizienten  $e_k$ :

Wir betrachten das Polynom  $h$  als Funktion in  $(e_1, \dots, e_n, t)$  und erhalten so die implizite Funktion

$$0 = h(e_1, \dots, e_n, z_i(e_1, \dots, e_n)) \quad \text{für alle } e = (e_1, \dots, e_n) \in \mathbb{R}_+^n.$$

Dem Satz der impliziten Funktionen folgend ist  $z_i$  an der Stelle  $e \in \mathbb{R}^n$  differenzierbar, falls  $\frac{\partial h}{\partial t}(e, z_i(e)) \neq 0$ . Wir erinnern uns: Ein Polynom  $h$  hat an der Stelle  $t$  genau dann eine mehrfache Nullstelle, wenn nicht nur  $h$ , sondern auch seine Ableitung an dieser Stelle null werden. Somit ist  $z_i$  an der Stelle  $e \in \mathbb{R}^n$  differenzierbar, falls die zu  $e$  korrespondierenden Nullstellen alle unterschiedlich sind.

## 2. Matrix-Analysis

Wir können nun die implizite Funktion per Kettenregel ableiten:

$$\begin{aligned} 0 &= \frac{d}{de_k} h(e, z_i(e)) = \sum_{j=1}^n \left( \frac{\partial h}{\partial e_j} \right) (e, z_i(e)) \cdot \delta_{j,k} + \frac{\partial h}{\partial t} (e, z_i(e)) \cdot \frac{\partial z_i}{\partial e_k} (e) \\ &= (-1)^k z_i(e)^{n-k} \cdot 1 + \frac{\partial h}{\partial t} (e, z_i(e)) \cdot \frac{\partial z_i}{\partial e_k} (e). \end{aligned}$$

Umstellen ergibt nun die zu zeigende Ableitung, da  $\frac{\partial h}{\partial t} (e, z_i(e)) = \prod_{\substack{j=1 \\ j \neq i}}^n (z_i(e) - z_j(e))$ . ■

**Lemma 2.13.16.** Sei  $U \subset \mathbb{C}^n$  und  $f: U \rightarrow \mathbb{C}$  mit  $f(z_1, \dots, z_n) = \mathcal{F}(e_1, \dots, e_n)$  eine symmetrische Funktion. Für Koeffizienten  $(e_1, \dots, e_n) \in \mathbb{C}^n$  mit paarweise unterschiedlichen Nullstellen  $(z_1, \dots, z_n) \subset U$  lauten die partiellen Ableitungen von  $f$  nach den Koeffizienten  $e_k$

$$\frac{\partial \mathcal{F}}{\partial e_k} (e_1, \dots, e_n) = \sum_{i=1}^n \frac{(-1)^{k+1} z_i^{n-k}}{\prod_{\substack{j=1 \\ j \neq i}}^n (z_i - z_j)} \frac{\partial f}{\partial z_i} (z_1, \dots, z_n) \quad \text{für alle } (e_1, \dots, e_n) \text{ mit } z_j \neq z_i, \text{ wenn } j \neq i.$$

*Beweis.* Ableitung per Kettenregel. ■

**Folgerung 2.13.17.** Im Falle von  $f(z_1, \dots, z_n) = \sum_{i=1}^n \phi(z_i)$  vereinfachen sich die partiellen Ableitungen zu

$$\frac{\partial \mathcal{F}}{\partial e_k} (e_1, \dots, e_n) = \sum_{i=1}^n \frac{(-1)^{k+1} z_i^{n-k}}{\prod_{\substack{j=1 \\ j \neq i}}^n (z_i - z_j)} \phi'(z_i) \quad \text{für alle } (e_1, \dots, e_n) \text{ mit } z_j \neq z_i, \text{ wenn } j \neq i.$$

**Definition 2.13.18.** Die *Diskriminante* eines normierten Polynoms

$$\begin{aligned} p &= (X - z_1)(X - z_2) \dots (X - z_n) \\ &= X^n - e_1 X^{n-1} + e_2 X^{n-2} + \dots + (-1)^n e_n \end{aligned} \tag{2.13.11}$$

mit Nullstellen  $z_1, \dots, z_n \in \mathbb{C}$  und Koeffizienten  $e_1, \dots, e_n \in \mathbb{C}$  ist definiert

$$D(p) = \prod_{1 \leq i < j \leq n} (z_i - z_j)^2.$$

**Proposition 2.13.19.** Wir können die Diskriminante als symmetrisches Polynom in den Nullstellen  $\mathfrak{d}: \mathbb{C}^n \rightarrow \mathbb{C}$  und als Polynom in den Koeffizienten  $\widehat{\mathfrak{d}}: \mathbb{C}^n \rightarrow \mathbb{C}$  darstellen, d.h.

$$D(p) = \mathfrak{d}(z_1, \dots, z_n) = \widehat{\mathfrak{d}}(e_1, \dots, e_n). \tag{2.13.12}$$

So lautet die Diskriminante im Fall  $n = 3$ :  $D(p) = e_1^2 e_2^2 - 4 e_2^3 - 4 e_2^3 e_3 + 18 e_1 e_2 e_3 - 27 e_3^2$ .

*Beweis.* Die Abbildungsvorschrift von  $\mathfrak{d}$  folgt direkt der Definition der Diskriminante<sup>24</sup>. Offenkundig ist  $\mathfrak{d}$  symmetrisch. Mit Folgerung 2.13.13 ergibt sich die Existenz von einer Funktion in den Koeffizienten  $\widehat{\mathfrak{d}}$ . Über die Art des Ausmultiplizierens sämtlicher  $z_i$  in der Diskriminante wird klar, dass  $\widehat{\mathfrak{d}}$  selbst ein Polynom ist. ■

**Proposition 2.13.20** (Identitätssatz für Polynome in mehreren Veränderlichen). Seien  $f, g: \mathbb{C}^n \rightarrow \mathbb{C}$  Polynome. Stimmen  $f$  und  $g$  auf einer offenen Teilmenge  $U \subseteq \mathbb{C}^n$  überein, so gilt  $f = g$ .

<sup>24</sup>deren Notation formal korrekt eigentlich über Funktionen  $p \mapsto z_i(p)$  hätte erfolgen müssen, was der leichteren Verständlichkeit geopfert wurde.

## 2. Matrix-Analysis

*Beweis.* Wir wählen ein  $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n) \in \mathcal{U}$  und  $\varepsilon > 0$ , sodass  $U_\varepsilon(\tilde{x}) := (\tilde{x} + z \mid z \in \mathbb{C}^n \text{ mit } |z| < \varepsilon) \subseteq \mathcal{U}$ , also  $(f-g)(x) = 0$  für alle  $x \in U_\varepsilon(\tilde{x})$ . Angenommen, es gäbe ein  $\tilde{y} \in \mathbb{C}^n$  mit  $f(\tilde{y}) \neq g(\tilde{y})$ , also  $(f-g)(\tilde{y}) \neq 0$ . Wir betrachten nun das Polynom  $\psi: \mathbb{C} \rightarrow \mathbb{C}$  mit

$$\psi(t) = (f-g)((1-t)\tilde{x} + t\tilde{y}). \quad (2.13.13)$$

Es ist  $\psi(0) = (f-g)(\tilde{x}) = 0$  und  $\psi(1) = (f-g)(\tilde{y}) \neq 0$ . Für alle  $t \in \mathbb{C}$  mit  $|t| < \frac{\varepsilon}{|\tilde{y}-\tilde{x}|}$  ist  $|t(\tilde{y}-\tilde{x})| < \varepsilon$ , also  $(1-t)\tilde{x} + t\tilde{y} \in U_\varepsilon(\tilde{x})$  und damit  $\psi(t) = 0$ . Ein Polynom in einer Veränderlichen, das nicht das Nullpolynom ist, kann aber höchstens  $n$  Nullstellen besitzen. Somit muss  $\psi$  das Nullpolynom sein, ein Widerspruch. ■

**Proposition 2.13.21.** *Sei  $\mathcal{M} \subseteq \mathbb{C}^n$  eine Menge von Koeffizientenvektoren und*

$$\mathcal{U} := \{(e_1, \dots, e_n) \in \mathcal{M} \mid z_i \neq z_j \text{ für } i \neq j\} \quad (2.13.14)$$

*die Teilmenge aller Elemente mit paarweise unterschiedlichen Nullstellen. Dann liegt  $\mathcal{U}$  dicht in  $\mathcal{M}$ , d.h.  $\overline{\mathcal{U}} = \mathcal{M}$ .*

*Beweis.* Wir zeigen, dass die Menge  $\mathcal{M} \setminus \mathcal{U}$  kein inneres hat. Angenommen, es gäbe einen inneren Punkt  $x_0$  in dieser Menge, so gäbe es eine offene Menge  $U_\varepsilon(x_0) \subseteq \mathcal{M}$  so dass all deren Elemente  $x \in U_\varepsilon(x_0)$  auf mehrfache Nullstellen führten, d.h.

$$\mathfrak{d}(x) = 0 \quad \text{für alle } x \in U_\varepsilon(x_0),$$

wobei  $\hat{\mathfrak{d}}$  die Diskriminante als Polynom in den Koeffizienten bezeichnet. Mit dem Identitätssatz für Polynome, Prop 2.13.20, wäre die Diskriminante dann das Nullpolynom, ein Widerspruch. ■

**Lemma 2.13.22.** *Sei  $M \subseteq \mathbb{R}^n$  eine konvexe, offene und symmetrische Menge und  $T := ((e_1, \dots, e_n) \in \mathbb{R}^n \mid (z_1, \dots, z_n) \in M)$ . Zudem sei eine symmetrische Funktion  $f: M \rightarrow \mathbb{R}$  gegeben. Ist nun  $f$  auf  $\overline{M}$  (dem Abschluss der Menge der Nullstellenvektoren)  $n \cdot r$ -mal stetig differenzierbar, dann ist  $\mathcal{F}$  auf  $\overline{T}$  (dem Abschluss der Menge der Koeffizientenvektoren)  $r$ -mal stetig differenzierbar.*

*Hierbei ist  $\mathcal{F}: T \rightarrow \mathbb{R}$  die nach Folgerung 2.13.13 existierende Funktion mit  $f(z_1, \dots, z_n) = \mathcal{F}(e_1, \dots, e_n)$*

*Beweis.* Siehe Ball [3, Theorem 3.2, S. 705]. ■

**Folgerung 2.13.23.** *Sei eine symmetrische und unendlich oft differenzierbare Funktion  $f: \mathbb{R}_+^n \rightarrow \mathbb{R}$  gegeben. Zudem sei  $\mathcal{F}: T \rightarrow \mathbb{R}$  mit  $T := ((e_1, \dots, e_n) \in \mathbb{R}_+^n \mid (z_1, \dots, z_n) \in \mathbb{R}_+^n)$  die nach Folgerung 2.13.13 existierende Funktion mit  $f(z_1, \dots, z_n) = \mathcal{F}(e_1, \dots, e_n)$ . Dann ist auch  $\mathcal{F}$  unendlich oft differenzierbar.*

*Beweis.* Für alle  $\varepsilon > 0$  liefert die Differenzierbarkeit von  $f$  auf  $(\varepsilon, \infty)^n$  mit Lemma 2.13.22 die Differenzierbarkeit von  $\mathcal{F}$  auf  $T_\varepsilon := ((e_1, \dots, e_n) \in \mathbb{R}^n \mid (z_1, \dots, z_n) \in (\varepsilon, \infty)^n)$ . Somit ist  $\mathcal{F}$  unendlich oft auf  $T = \bigcup_\varepsilon ((e_1, \dots, e_n) \in \mathbb{R}^n \mid (z_1, \dots, z_n) \in (\varepsilon, \infty)^n) = ((e_1, \dots, e_n) \in \mathbb{R}^n \mid (z_1, \dots, z_n) \in \mathbb{R}_+^n)$  differenzierbar. ■

**Folgerung 2.13.24.** *Im Falle von  $f(z_1, \dots, z_n) = \sum_{i=1}^n \phi(z_i)$  ist  $\mathcal{F}$  unendlich oft auf allen  $(e_1, \dots, e_n)$ , für die  $(z_1, \dots, z_n) \in \mathbb{R}_+^n$  gilt, differenzierbar, wenn  $\phi$  auf  $\mathbb{R}_+$  differenzierbar ist.*

Wenden wir zudem Lemma 2.5.26 an, erhalten wir

**Lemma 2.13.25** (Ableitung in den Invarianten). *Sei  $\psi: \text{Sym}^+(n) \rightarrow \mathbb{R}$  eine isotrope, unendlich oft differenzierbare Funktion und sei  $T := \{(I_1, \dots, I_n) \mid X \in \text{Sym}^+(n)\}$  die Menge der Invariantenvektoren, die auf Matrizen aus  $\text{Sym}^+(n)$  führen. Dann existiert eine unendlich oft differenzierbare Darstellung in den Invarianten  $\Psi: T \rightarrow \mathbb{R}$  mit  $\Psi(X) = \Psi(I_1, \dots, I_n)$ .*

## 2. Matrix-Analysis

*Beweis.* Lemma 2.5.26 garantiert uns, dass es wegen der Isotropie von  $\psi$  eine eindeutige symmetrische Funktion  $g: \mathbb{R}_+^n \rightarrow \mathbb{R}$  und eine Funktion  $\Psi: \mathbb{R}_+^n \rightarrow \mathbb{R}$  mit  $(\lambda_1, \dots, \lambda_n) \mapsto g(\lambda_1, \dots, \lambda_n) = \psi(X)$  und  $(I_1, \dots, I_n) \mapsto \Psi(I_1, \dots, I_n) = \psi(X)$  für alle  $X \in \text{Sym}^+(n)$  mit Eigenwerten  $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$  und Invarianten  $I_1, \dots, I_n \in \mathbb{R}_+$  gibt.

Die Differenzierbarkeit von  $g$  wird impliziert durch die Differenzierbarkeit von  $\psi$  auf den Diagonalmatrizen, die partiellen Ableitungen  $\frac{\partial \psi}{\partial x_{ii}}$  liefern uns die Ableitung

$$\text{Dg}(\lambda_1, \dots, \lambda_n) = \left( \frac{\partial \psi}{\partial x_{11}} \quad \frac{\partial \psi}{\partial x_{22}} \quad \dots \quad \frac{\partial \psi}{\partial x_{nn}} \right) (\text{diag}(\lambda_1, \dots, \lambda_n)).$$

Mit Folgerung 2.13.23 erhalten wir die Differenzierbarkeit von  $\Psi$ . ■

**Beispiel 2.13.26.** Für die Funktion  $\psi: \text{Sym}^+(n) \rightarrow \mathbb{R}$  mit  $\psi(X) = \|\log X\|^2$  existiert eine unendlich oft differenzierbare Funktion in den Invarianten, d.h.  $\|\log X\|^2 = \Psi(I_1, \dots, I_n)$ .

**Lemma 2.13.27.** Für alle  $(e_1, \dots, e_n) \in \mathbb{R}_+^n$  besteht  $(z_1, \dots, z_n)$  ausschließlich aus positiven reellen Zahlen und konjugiert komplexen Zahlenpaaren. Zudem gilt

$$(t + z_1) \cdot \dots \cdot (t + z_n) \geq \min\{t^n, e_n\} \quad \text{für alle } t \in \mathbb{R}_+. \quad (2.13.15)$$

*Beweis.* Nach dem Fundamentalsatz der Algebra hat ein Polynom mit reellen Koeffizienten

$$h = X^n - e_1 X^{n-1} + e_2 X^{n-2} + \dots + (-1)^n e_n$$

genau  $n$  Nullstellen; sie kommen als reelle Zahlen oder als konjugiert komplexe Zahlenpaare vor. Dass die reellen Nullstellen positiv sind sehen wir am Einfachsten, wenn wir  $-X$  in  $h$  einsetzen und damit

$$h(-X) = X^n + e_1 X^{n-1} + \dots + e_{n-1} X + e_n,$$

erhalten, ein Polynom mit ausschließlich reellen Koeffizienten. Mit der somit verwendbaren Vorzeichenregel von Descartes wird die maximale Anzahl der positiven reellen Nullstellen durch die Anzahl der Vorzeichenwechsel in den Koeffizienten beschränkt. Da  $h(-X)$  über keine Vorzeichenwechsel verfügt (alle  $e_k$  sind schließlich positiv), hat es keine positiven, und damit  $h$  keine negativen Nullstellen.

Setzen wir  $-t$  in  $h$  ein, erhalten wir für alle  $t > 0$

$$(t + z_1) \cdot \dots \cdot (t + z_n) = t^n + \underbrace{e_1 t^{n-1} + \dots + e_{n-1} t + e_n}_{\geq 0} \geq \min\{t^n, e_n\}. \quad \blacksquare$$

Wir erinnern uns:

**Definition 2.13.28.** Sei  $X \in \text{Sym}(n)$ . Wir bezeichnen mit  $I_k(X) := e_k(\lambda_1, \dots, \lambda_n)$  die  $k$ -te Invariante von  $X$ . Hierbei sind  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Eigenwerte von  $X$  und  $e_k$  das  $k$ -te elementarsymmetrische Polynom.

**Bemerkung 2.13.29.** Die Invarianten sind gemäß ihres Namens invariant gegenüber Basistransformationen, ergo sind die  $e_k$  isotrope Funktionen mit  $e_k(Q^T X Q) = e_k(X)$  für alle  $X \in \text{Sym}(n)$  und alle  $Q \in O(n)$ .

**Bemerkung 2.13.30.** Die erste Invariante einer Matrix ist stets deren Spur, d.h.  $I_1(X) = \text{tr}(X)$ , die  $n - 1$ -te Invariante ist stets die Spur ihres Kofaktors, d.h.  $I_{n-1}(X) = \text{tr}(\text{Cof } X)$ , die  $n$ -te Invariante ist stets deren Determinante, d.h.  $I_n(X) = \det(X)$ .

**Definition 2.13.31.** Das *charakteristische Polynom* einer Matrix  $X \in \text{Sym}(n)$  ist definiert als

$$P_X(t) := \det(t \cdot \mathbb{1} - X). \quad (2.13.16)$$

$P_X$  ist ein normiertes Polynom  $n$ -ten Grades. Wir bezeichnen die Koeffizienten  $I_1, \dots, I_n \in \mathbb{R}$  des charakteristischen Polynoms  $P_X$  in der Darstellung

$$P_X(t) = t^n - I_1 t^{n-1} + I_2 t^{n-2} + \dots + (-1)^n I_n \quad (2.13.17)$$

als die *Invarianten von  $X$* .



## 2. Matrix-Analysis

Wir erinnern uns: Die Invarianten von  $X$  sind per Definition über die elementarsymmetrischen Polynome differenzierbare Funktionen in den Eigenwerten  $\lambda_1, \dots, \lambda_n$  von  $X$ , es gilt

$$I_k(X) = e_k(\lambda_1, \dots, \lambda_n) \quad \text{und} \quad \frac{\partial I_k}{\partial \lambda_i}(z) = e_{k-1}(\lambda_1, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_n).$$

Aber wir können auch eine umgekehrte Aussage treffen:

**Lemma 2.13.32.** *Sei  $X \in \text{Sym}(n)$  mit Eigenwerten  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  und Invarianten  $I_1, \dots, I_n \in \mathbb{R}$ .*

*Die Eigenwerte von  $X$  sind stetige Funktionen in den Invarianten von  $X$ . Für Invarianten, die auf paarweise unterschiedliche Eigenwerte führen, sind die Eigenwerte sogar differenzierbar mit partiellen Ableitungen*

$$\frac{\partial \lambda_i}{\partial I_k}(I_1, \dots, I_n) = \frac{(-1)^{k+1} \lambda_i^{n-k}}{\prod_{\substack{j=1 \\ j \neq i}}^n (\lambda_i - \lambda_j)} \quad \text{für alle } I_1, \dots, I_n \text{ mit } \lambda_i \neq \lambda_j \text{ falls } i \neq j. \quad (2.13.18)$$

*Beweis.* Sei  $X = Q^T \text{diag}(\lambda_1, \dots, \lambda_n)Q$  mit  $Q \in O(n)$ . Mit der Isotropie der Determinanten und der Definition der Invarianten erhalten wir

$$\begin{aligned} P_X(t) &= \det(t \cdot \mathbb{1} - X) \\ &= \det(t \cdot \mathbb{1} - Q^T \text{diag}(\lambda_1, \dots, \lambda_n)Q) \\ &= \det(Q^T(t \cdot \mathbb{1} - \text{diag}(\lambda_1, \dots, \lambda_n)Q)) \\ &= \det \text{diag}(t - \lambda_1, t - \lambda_2, \dots, t - \lambda_n) \\ &= (t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_n) \\ &= t^n - e_1(\lambda_1, \dots, \lambda_n)t^{n-1} + e_2(\lambda_1, \dots, \lambda_n)t^{n-2} + \dots + e_n(\lambda_1, \dots, \lambda_n) \\ &= t^n - I_1 t^{n-1} + I_2 t^{n-2} + \dots + I_n. \end{aligned}$$

Die Eigenwerte  $\lambda_1, \dots, \lambda_n$  sind somit die Nullstellen eines normierten Polynoms  $n$ -ten Grades mit Koeffizienten  $I_1, \dots, I_n$ . Mit Lemma 2.13.15 folgt nun die Aussage. ■

### 2.13.1. Ungleichungen zu elementarsymmetrischen Polynomen

Es gibt viele wichtige Ungleichungen zu elementarsymmetrischen Polynomen. Als Beispiele stellen wir die Newton-Ungleichung und die MacLaurin-Ungleichung vor:

**Lemma 2.13.33** (Newton-Ungleichung). *Sei  $a \in \mathbb{R}_+^n$ . Dann gilt*

$$\frac{e_{k-1}(a)}{\binom{n}{k-1}} \cdot \frac{e_{k+1}(a)}{\binom{n}{k+1}} \leq \left( \frac{e_k(a)}{\binom{n}{k}} \right)^2, \quad (2.13.19)$$

*mit Gleichheit genau dann, wenn  $a_1 = a_2 = a_3 = \dots = a_n$ .*

*Beweis.* Siehe z.B. Bernstein [7, Fact 1.17.11 S. 55]. ■

**Lemma 2.13.34** (MacLaurin-Ungleichung). *Sei  $a \in \mathbb{R}_+^n$ . Dann gilt*

$$\frac{e_1(a)}{\binom{n}{1}} \geq \sqrt{\frac{e_2(a)}{\binom{n}{2}}} \geq \sqrt[3]{\frac{e_3(a)}{\binom{n}{3}}} \geq \dots \geq \sqrt[n]{\frac{e_n(a)}{\binom{n}{n}}}, \quad (2.13.20)$$

*mit Gleichheit genau dann, wenn  $a_1 = a_2 = a_3 = \dots = a_n$ .*

*Beweis.* Siehe z.B. Biler und Witkowski [8, Problem 1.60 S. 5]. ■

Die MacLaurin-Ungleichung können wir auf die Invarianten von Matrizen mit Determinante eins übertragen:

**Folgerung 2.13.35.** Sei  $X \in \text{Sym}^+(n)$ . Dann gilt

$$I_k\left(\frac{X}{\det X^{1/n}}\right) \geq \binom{n}{k} \quad \text{für alle } k \in \{1, \dots, n\}, \quad (2.13.21)$$

wobei Gleichheit genau für alle  $X \in \mathbb{R} \cdot \mathbb{1}$  eintritt.

**Bemerkung 2.13.36.** Für den dreidimensionalen Fall erhalten wir dann

$$\text{tr}\left(\frac{X}{\det X^{1/3}}\right) \geq 3 \quad \text{und} \quad \text{tr}\left(\text{Cof} \frac{X}{\det X^{1/3}}\right) \geq 3 \quad \text{für alle } X \in \text{Sym}(3) \quad (2.13.22)$$

bzw.

$$\text{tr}(X) \geq 3 \quad \text{und} \quad \text{tr}(\text{Cof } X) \geq 3 \quad \text{für alle } X \in \text{Sym}(3) \cap \text{SL}(3). \quad (2.13.23)$$

Aus der Newton-Ungleichung und der MacLaurin-Ungleichung sind im Fall  $n = 3$  zudem zwei ähnliche Ungleichungen zu folgern, welche sich ebenfalls im Anhang von Hartmann und Neff [35] finden lassen:

**Folgerung 2.13.37.** Sei  $X \in \text{Sym}^+(3)$ . Dann gilt

$$3 I_1 I_3 \leq I_2^2 \quad \text{und} \quad 3 I_2 \leq I_1^2, \quad (2.13.24)$$

wobei  $I_1, I_2, I_3$  die Invarianten von  $X$  sind.

## 2.13.2. Eigenschaften von elementarsymmetrischen Polynomen

**Lemma 2.13.38.** Seien  $x \in \mathbb{R}_+^n$  und  $i, j \in \{1, \dots, n\}$  mit  $x_i \geq x_j$ . Dann ist  $e_k(x^{(i)}) \leq e_k(x^{(j)})$  für alle  $k \in \{1, \dots, n\}$ .

*Beweis.* Für den Beweis verwenden wir die rekursive Darstellung der elementarsymmetrischen Polynome:

$$e_k(x^{(i)}) = x_j e_{k-1}(a^{(ij)}) + e_k(x^{(ij)}) \leq a_i e_{k-1}(a^{(ij)}) + e_k(a^{(ij)}) = e_k(a^{(j)}). \quad \blacksquare$$

**Folgerung 2.13.39.** Mit  $a_1 \geq a_2 \geq \dots \geq a_n > 0$  erhalten wir  $e_k(a^{(1)}) \leq e_k(a^{(2)}) \leq \dots \leq e_k(a^{(n)})$  für alle  $k \in \{1, \dots, n\}$ .

**Lemma 2.13.40.** Eine Funktion  $\phi: \mathbb{R}^n \rightarrow \mathbb{R}$  ist genau dann Schur-konvex, wenn  $\phi$  symmetrisch ist und wenn

$$(x_i - x_j) \left( \frac{\partial}{\partial x_i} \phi(x) - \frac{\partial}{\partial x_j} \phi(x) \right) \geq 0 \quad (2.13.25)$$

für alle  $x \in \mathbb{R}^n$  und  $i, j \in \{1, \dots, n\}$ . Analog dazu ist  $\phi$  genau dann Schur-konkav, wenn  $\phi$  symmetrisch ist und die Ungleichung (2.13.25) mit  $\leq 0$  wird erfüllt.

*Beweis.* Siehe Marshall, Olkin, Arnold [45, A.1. S. 84]. ■

**Satz 2.13.41.** Die elementarsymmetrischen Polynome  $e_k: \mathbb{R}^n \rightarrow \mathbb{R}$  sind Schur-konkav und monoton fallend für alle  $k \in \{1, \dots, n\}$  und sogar streng Schur-konkav für  $k \geq 2$ .

*Beweis.* Die Monotonie ist klar.

Für  $k = 1$  und  $i, j \in \{1, \dots, n\}$  haben wir

$$(x_i - x_j) \left( \frac{\partial}{\partial x_i} e_k(x) - \frac{\partial}{\partial x_j} e_k(x) \right) = (x_i - x_j) (1 - 1) = 0.$$

## 2. Matrix-Analysis

Für  $k \geq 2$  und  $i = j$  haben wir  $x_i = x_j$  und damit ist (2.13.25) mit Gleichheit erfüllt. Für  $i \neq j$  haben wir

$$\begin{aligned} (x_i - x_j) \left( \frac{\partial}{\partial x_i} e_k(x) - \frac{\partial}{\partial x_j} e_k(x) \right) &= (x_i - x_j) \left( e_{k-1}(x^{(i)}) - e_{k-1}(x^{(j)}) \right) \\ &= (x_i - x_j) \left( x_j e_{k-2}(x^{(ij)}) - x_i e_{k-2}(x^{(ij)}) \right) = \underbrace{(x_i - x_j)}_{\geq 0} \underbrace{(x_j - x_i)}_{\leq 0} \underbrace{e_{k-2}(x^{(ij)})}_{> 0} \leq 0. \end{aligned}$$

Somit ist  $e_k$  Schur-konkav für alle  $k \in \{1, \dots, n\}$ .

Wir haben noch die strenge Schur-Konkavität für  $k \geq 2$  zu zeigen, also  $e_k(a) > e_k(b)$  für alle  $a \prec b$  mit  $a^\downarrow \neq b^\downarrow$ :

Nehmen wir o.B.d.A.  $a_1 \geq a_2 \geq \dots \geq a_n > 0$ ,  $b_1 \geq b_2 \geq \dots \geq b_n > 0$  und  $a \prec b$  mit  $a \neq b$  an, dann gibt es einen kleinsten Index  $s$  mit  $a_s < b_s$  und einen größten index  $r < s$  mit  $a_r > b_r$ . Sei  $\alpha := \min\{b_r - a_r, a_s - b_s\}$ , dann ist

$$(a_1, \dots, a_n) \prec (a_1, \dots, a_r + \alpha, \dots, a_s - \alpha, \dots, a_n) \prec (b_1, \dots, b_n),$$

oder  $a \prec \tilde{a} \prec b$  mit  $\tilde{a} := (a_1, \dots, a_r + \alpha, \dots, a_s - \alpha, \dots, a_n)$ . Wir erhalten

$$\begin{aligned} e_k(\tilde{a}) &= (a_r + \alpha) e_{k-1}(\tilde{a}^{(r)}) + e_k(\tilde{a}^{(r)}) \\ &= (a_r + \alpha) \left( (a_s - \alpha) e_{k-2}(a^{(rs)}) + e_{k-1}(a^{(rs)}) \right) + (a_s - \alpha) e_{k-1}(a^{(rs)}) + e_k(a^{(rs)}) \\ &= (a_r + \alpha)(a_s - \alpha) e_{k-2}(a^{(rs)}) + (a_r + a_s) e_{k-1}(a^{(rs)}) + e_k(a^{(rs)}) \\ &= (-\alpha(a_r - a_s) - \alpha^2) e_{k-2}(a^{(rs)}) + a_r(a_s e_{k-2}(a^{(rs)}) + e_{k-1}(a^{(rs)})) + (a_s e_{k-1}(a^{(rs)}) + e_k(a^{(rs)})) \\ &= (-\alpha(a_r - a_s) - \alpha^2) e_{k-2}(a^{(rs)}) + a_r e_{k-1}(a^{(r)}) + e_k(a^{(r)}) \\ &= \underbrace{(-\alpha(a_r - a_s) - \alpha^2)}_{< 0} \underbrace{e_{k-2}(a^{(rs)})}_{> 0} + e_k(a), \end{aligned} \tag{2.13.26}$$

somit  $e_k(a) > e_k(\tilde{a})$ . Wegen  $\tilde{a} \prec b$  und der Schur-Konkavität von  $e_k$  ist die Ungleichung  $e_k(\tilde{a}) \geq e_k(b)$  wahr, also  $e_k(a) > e_k(b)$ . ■

**Beispiel 2.13.42.** Seien  $a := (9, 5, 4, 3, 1)$  und  $b := (9, 8, 3, 1, 1)$ . Damit ist  $r = 2$ ,  $s = 4$  und  $\alpha = \min\{b_2 - a_2, a_4 - b_4\} = 2$ , und dann

$$(9, 5, 4, 3, 1) \prec \tilde{a} = (9, 7, 4, 1, 1) \prec (9, 8, 3, 1, 1).$$

So sehen wir

$$e_5(\tilde{a}) = 9(5+2)4(3-2)1 = 5 \cdot 5 \cdot 4 \cdot 3 \cdot 1 - 9 \cdot 4 \cdot 1 \left( (5-3)2 + 2^2 \right) < 9 \cdot 5 \cdot 4 \cdot 3 \cdot 1 = e_5(a).$$

**Folgerung 2.13.43.** Seien  $x, y \in \mathbb{R}_+^n$ . Dann gilt: Aus  $x \prec y$  folgt  $e_k(x) \geq e_k(y)$  für alle  $k \in \{1, \dots, n\}$ . Jedoch impliziert  $e_k(x) \geq e_k(y)$  im Allgemeinen nicht  $x \prec y$ : Hierzu betrachten wir  $a = (2, 2, 2)$  und  $b = (1, 1, 1)$ .

**Folgerung 2.13.44.** Seien  $x, y \in \mathbb{R}_+^n$ . Dann gilt: Aus  $x \prec y$  folgt  $e_k(x) > e_k(y)$  für alle  $k \in \{2, \dots, n\}$  wenn  $x^\downarrow \neq y^\downarrow$ .

**Folgerung 2.13.45.** Wenn  $x \prec y$  und  $e_k(x) = e_k(y)$  für irgendein  $k \in \{2, \dots, n\}$ , dann ist  $x^\downarrow = y^\downarrow$  und somit  $e_k(x) = e_k(y)$  für alle  $k \in \{2, \dots, n\}$ .

# 3. Sum-of-Squared-Logarithms-Ungleichung

Wesentliche Teile dieses Kapitels wurden bereits in [12] und [21] veröffentlicht.

Wir werden in diesem Kapitel die SSLI in ihren verschiedenen Erscheinungsformen vorstellen. Zudem werden wir erläutern, wie sich die chronologische Entwicklung seit ihrem ersten Auftreten als wissenschaftliche Fragestellung bis zum Zeitpunkt der Fertigstellung dieser Arbeit vollzogen hat. Dann werden wir einen vollständigen Beweis für den  $n$ -dimensionalen Fall vorstellen, der mit Grundkenntnissen der Analysis nachvollzogen werden kann. Anschließend wollen wir noch andere Bedingungen untersuchen, welche ebenfalls die SSLI implizieren, sowie auch Bedingungen anderer Ungleichungen beweisen, die der SSLI ähnlich sind und dabei einen Bezug zum Konzept der Majorization herstellen.

## 3.1. Einleitung

### 3.1.1. Die Ungleichung

**Satz 3.1.1** (Sum-of-Squared-Logarithms-Ungleichung (SSLI)). *Seien  $n \in \mathbb{N}$  und  $x, y \in \mathbb{R}_+^n$ , sodass*

$$e_k(x) \leq e_k(y) \quad \text{für alle } k \in \{1, \dots, n-1\}, \quad (3.1.1)$$

und

$$e_n(x) = e_n(y). \quad (3.1.2)$$

Dann gilt

$$\sum_{i=1}^n (\log x_i)^2 \leq \sum_{i=1}^n (\log y_i)^2.$$

**Bemerkung 3.1.2.** Für  $n = 3$  bedeutet dies:

Seien  $x_1, x_2, x_3, y_1, y_2, y_3 > 0$ , sodass

$$\begin{aligned} x_1 + x_2 + x_3 &\leq y_1 + y_2 + y_3, \\ x_1 x_2 + x_1 x_3 + x_2 x_3 &\leq y_1 y_2 + y_1 y_3 + y_2 y_3, \\ x_1 x_2 x_3 &= y_1 y_2 y_3. \end{aligned} \quad (3.1.3)$$

Dann gilt  $(\log x_1)^2 + (\log x_2)^2 + (\log x_3)^2 \leq (\log y_1)^2 + (\log y_2)^2 + (\log y_3)^2$ .

**Satz 3.1.3** (Sum-of-Squared-Logarithms-Ungleichung (SSLI) - Matrixform). *Seien  $U, \tilde{U} \in \text{Sym}^+(n)$ , sodass*

$$I_k(U) \leq I_k(\tilde{U}) \quad \text{für alle } k \in \{1, \dots, n-1\}, \quad (3.1.4)$$

und

$$\det U = \det \tilde{U}. \quad (3.1.5)$$

Dann gilt

$$\|\log U\|^2 \leq \|\log \tilde{U}\|^2.$$

In Folgerung 2.13.35 auf Seite 69 haben wir gezeigt, dass für  $U \in \text{Sym}^+(3)$  mit  $I_3 = \det U = 1$  stets  $I_1 = \text{tr}(U) \geq 3$  und  $I_2 = \text{tr}(\text{Cof } U) \geq 3$  gilt, deshalb können wir die SSLI eingeschränkt auf die Menge  $\text{Sym}^+(3) \cap \text{SL}(3)$  umformulieren:

### 3. Sum-of-Squared-Logarithms-Ungleichung

**Folgerung 3.1.4.** Seien  $U, \tilde{U} \in \text{Sym}^+(3)$  mit  $\det(U) = \det(\tilde{U}) = 1$ . Sind zudem

$$|I_1(U) - 3| \leq |I_1(\tilde{U}) - 3| \quad \text{und} \quad |I_2(U) - 3| \leq |I_2(\tilde{U}) - 3|, \quad (3.1.6)$$

erfüllt, dann gilt

$$\|\log U\|^2 \leq \|\log \tilde{U}\|^2.$$

**Bemerkung 3.1.5.** Diese Umformulierung für  $U, \tilde{U} \in \text{Sym}^+(3) \cap \text{SL}(3)$  lässt eine geometrische Interpretation der Bedingungen der SSLI zu: So bemisst  $4 \cdot |I_1(U) - 3| = 4 \cdot |\lambda_1 + \lambda_2 + \lambda_3 - 3|$  die Gesamtkantenverzerrung und  $2 \cdot |I_2(U) - 3| = 2 \cdot |\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3 - 3|$  die Gesamtflächenverzerrung eines mittels  $U$  verzerrten Einheitswürfels. Verwenden wir nun  $W: \text{Sym}^+(3) \cap \text{SL}(3) \rightarrow \mathbb{R}$  mit  $W(U) = \|\log U\|^2$  als Energiefunktion, dann besagt die SSLI: Ist für eine volumenerhaltende Deformation  $U$  sowohl die Kanten- als auch Flächenverzerrung kleiner gleich einer volumenerhaltenden Deformation  $\tilde{U}$ , dann ist auch die Deformationsenergie von  $U$  kleiner gleich der Deformationsenergie von  $\tilde{U}$ .

#### 3.1.2. Die SSLI als Schlüssel zur Lösung eines Minimierungsproblems

Die Sum-of-Squared-Logarithms-Ungleichung (SSLI) hat wichtige Anwendungen in der Matrix-Analysis und der nichtlinearen Elastizitätstheorie [9, 51, 52, 57, 58, 60]. Einige davon werden wir im weiteren Verlauf dieser Arbeit untersuchen.

Als erstes tauchte die SSLI 2013 in einer Arbeit von Neff, Nakatsukasa und Fischle [59] als isolierte Fragestellung auf, mit der das folgende Resultat direkt bewiesen werden kann.

**Lemma 3.1.6.** Sei  $F \in \text{GL}^+(3)$  mit Polarzerlegung  $F = RU$  mit  $R \in \text{SO}(3)$  und  $U \in \text{Sym}^+(3)$ , dann gilt

$$\min_{Q \in \text{SO}(3)} \|\text{sym Log } Q^T F\|^2 = \min \|\text{sym Log } R^T F\|^2 = \|\log U\|^2, \quad (3.1.7)$$

d.h. genau  $Q = R$  minimiert den Ausdruck.

*Beweis.* Aufgrund der Mengengleichheit  $\{Q^T U \mid Q \in \text{SO}(3)\} = \{Q^T F \mid Q \in \text{SO}(3)\}$  und  $\text{sym Log } R^T F = \text{sym Log } U = \log U^1$  können wir die zu zeigende Aussage umformen zu

$$\min_{Q \in \text{SO}(3)} \|\text{sym Log } Q^T U\|^2 = \|\log U\|^2. \quad (3.1.8)$$

offenkundig gilt

$$\min_{Q \in \text{SO}(3)} \|\text{sym Log } Q^T U\|^2 \leq \|\text{sym Log } \mathbb{1}^T U\|^2 = \|\log U\|^2,$$

sodass uns

$$\min_{Q \in \text{SO}(3)} \|\text{sym Log } Q^T U\|^2 \geq \|\text{sym Log } \mathbb{1}^T U\|^2 = \|\log U\|^2, \quad (3.1.9)$$

zu zeigen bleibt.

Setzen wir  $X_Q := \exp(\text{sym Log } Q^T U)$ . Dann ist  $\|\text{sym Log } Q^T U\|^2 = \|\text{Log } X_Q\|^2$  und die noch zu zeigende Aussage ist  $\|\text{Log } X_Q\|^2 \geq \|\log U\|^2$ , aufgrund von Folgerung 2.6.11 gleichbedeutend mit  $\frac{1}{2} \|\text{Log } X_Q^2\|^2 \geq \frac{1}{2} \|\log U^2\|^2$ , wir haben somit noch zu zeigen

$$\|\text{Log } X_Q^2\|^2 \geq \|\log U^2\|^2.$$

Genau hier kommt die SSLI zum Tragen: Denn sie sichert uns in der Matrixvariante Satz 3.1.3 die Aussage unter den folgenden Bedingungen an die Invarianten

$$I_1(X_Q^2) \geq I_1(U^2), \quad I_2(X_Q^2) \geq I_2(U^2) \quad \text{und} \quad I_3(X_Q^2) = I_3(U^2), \quad (3.1.10)$$

---

<sup>1</sup>Für  $U \in \text{Sym}(3)$  ist der Logarithmus als Menge der Lösungen  $\exp(\text{Log } U) = U$  eindeutig und gleich der primären Matrixlogarithmusfunktion, d.h.  $\text{Log } U = \log U$ .

### 3. Sum-of-Squared-Logarithms-Ungleichung

die wir nun prüfen werden. Als Hilfsausdruck setzen wir  $A_Q := \text{Log } Q^T U$ , sodass gilt  $\exp(A_Q) = Q^T U$  und

$$\begin{aligned} X_Q^2 &= \exp(\text{sym } 2A_Q) = \exp(A_Q + A_Q^T), \\ X_Q^{-2} &= \exp(-\text{sym } 2A_Q) = \exp(-A_Q - A_Q^T). \end{aligned}$$

Für die dritte Invariante gilt:

$$\det X_Q^2 = \det \exp(\text{sym } 2A_Q) \stackrel{\text{Folg. 2.7.6}}{=} \det \exp(2A_Q) = (\det \exp(A_Q))^2 = \det(Q^T U)^2 = \det U^2$$

und damit  $I_3(X_Q^2) = I_3(U^2)$ . Für die erste Invariante gilt:

$$\begin{aligned} \text{tr } X_Q^2 &= \text{tr}(\exp(A_Q + A_Q^T)) \\ &\stackrel{\text{Prop. 2.7.2}}{\geq} \text{tr}(\exp(A_Q) \cdot \exp(A_Q^T)) = \text{tr}(\exp(A_Q) \cdot (\exp(A_Q))^T) \\ &= \text{tr}((Q^T U)(Q^T U)^T) = \text{tr}(Q^T U^2 Q) = \text{tr}(U^2) \end{aligned}$$

und damit  $I_1(X_Q^2) \geq I_1(U^2)$ . Für die zweite Invariante benötigen wir erst noch eine Hilfsabschätzung

$$\begin{aligned} \text{tr } X_Q^{-2} &= \text{tr}(\exp(-A_Q - A_Q^T)) \\ &\stackrel{\text{Prop. 2.7.2}}{\geq} \text{tr}(\exp(-A_Q) \cdot \exp(-A_Q^T)) = \text{tr}(\exp(A_Q)^{-1} \cdot \exp(A_Q)^{-T}) \\ &= \text{tr}((Q^T U)^{-1}(Q^T U)^{-T}) = \text{tr}(U^{-1} Q Q^T U^{-1}) = \text{tr}(U^{-2}), \end{aligned}$$

mit der wir dann folgern können

$$\begin{aligned} I_2(X_Q^2) &= \text{tr}(\text{Cof } X_Q^2) = \text{tr}(\det X_Q^2 \cdot X_Q^{-2T}) = \det X_Q^2 \cdot \text{tr}(X_Q^{-2}) \\ &\geq \det U^2 \cdot \text{tr}(U^{-2}) = \text{tr}(\text{Cof } U^2) = I_2(U^2), \end{aligned}$$

das heißt, auch für die zweite Invariante gilt  $I_2(X_Q^2) \geq I_2(U^2)$ .

Somit gelten alle drei Bedingungen in Satz 3.1.10 und somit  $\|\text{Log } X_Q^2\|^2 \geq \|\log U^2\|^2$ , was noch zu zeigen war. ■

#### 3.1.3. Chronologie der Lösung

Die SSLI war Gegenstand mehrerer wissenschaftlicher Arbeiten.

Wie bereits in Abschnitt 3.1.2 beschrieben, tauchte sie 2013 erstmalig in einer Arbeit von Neff, Nakatsukasa und Fischle [59] als isolierte Fragestellung auf.

Ebenfalls im Jahr 2013 untersuchten Birsan, Neff und Lankeit [9] sie näher und präsentierten einen Beweis für  $n = 2$  und  $n = 3$ .

Pompe und Neff [64] veröffentlichten 2014 einen Beweis für den Fall  $n = 4$ , der mit der Existenz einer Fortführung von Pfaden in der nichtkonvexen Nullstellenmenge arbeitet und sich so nicht auf höhere Dimensionen übertragen lässt.

Im März 2015, nach einem Forschungsaufenthalt von Ionel Roventa an der Universität Duisburg-Essen, arbeiteten Thiel und Roventa an einem Ansatz, die SSLI mit Mitteln der Majorization zu lösen. Für den Fall  $n = 2$  funktioniert der resultierende Beweis auf eine sehr elementare Weise. Leider erwies sich die Übertragung via Induktion auf höhere Dimensionen als nicht zielführend. Den Beweis für  $n = 2$  werden wir in Abschnitt 3.1.4, die verfolgten Ideen in Abschnitt 3.1.5 präsentieren.

Im Mai 2015 veröffentlichte Patrizio Neff das Problem auf der Internet-Plattform *MathOverflow* [55]. Lev Borisov erweiterte es auf die komplexe Zahlenebene und präsentierte eine Beweisidee, die sich ausgearbeitet

### 3. Sum-of-Squared-Logarithms-Ungleichung

im Anhang von Borisov, Neff, Sra und Thiel [12] wiederfindet. Im Hauptteil der gleichen Arbeit findet sich schließlich der Beweis der SSLI für beliebige Dimensionen, den wir in überarbeiteter Form in Abschnitt 3.1.6 vorstellen und der aus der so hergestellten Zusammenarbeit mit Borisov, Neff und Sra erwuchs.

Miroslav Šilhavý, welcher auch über die Veröffentlichung und private Korrespondenz mit Neff vom Problem erfahren hatte, veröffentlichte eine auf der Pick-Nevalinna-Theorie basierende Charakterisierung der Funktionen, welche die bei der SSLI relevante Monotonie in den elementarsymmetrischen Polynomen aufweisen [74]. Bereits Jozsa und Mitchison [41] ahnten den auf der Pick-Funktion basierten Ansatz zur Beweisführung solcher Ungleichungen voraus, arbeiteten ihn aber nicht aus.

#### 3.1.4. Ein elementarer Beweis auf Basis der Konvexität für $n = 2$

Wir haben in Abschnitt 2.12 die Schur-Konvexität kennengelernt. Mit ihr können wir direkt ein Kriterium für die Gültigkeit der SSLI angeben.

**Proposition 3.1.7.** *Seien  $x, y \in \mathbb{R}_+^n$  Angenommen,  $x$  wird logarithmisch von  $y$  majorisiert, d.h. es gilt*

$$(\log x_1, \dots, \log x_n) \prec (\log y_1, \dots, \log y_n) \quad (3.1.11)$$

*dann folgt daraus die Gültigkeit der SSLI*

$$\sum_{i=1}^n (\log x_i)^2 \leq \sum_{i=1}^n (\log y_i)^2. \quad (3.1.12)$$

*Beweis.* Die Aussage ist gleichbedeutend mit der Schur-Konvexität der Abbildung  $z \mapsto \sum_{i=1}^n z_i^2$ . In Proposition 2.12.8 auf Seite 62 haben wir gezeigt, dass für konvexe Funktionen  $\phi: \mathbb{R} \rightarrow \mathbb{R}$  die Abbildung  $z \mapsto \sum_{i=1}^n \phi(z_i)$  schurkonvex ist. Die Abbildung  $x \mapsto x^2$  ist konvex. ■

Für  $n = 2$  können wir tatsächlich elementar beweisen, dass für alle  $x, y \in \mathbb{R}_+^n$ , welche die Bedingungen an die elementarsymmetrischen Polynome

$$e_k(x) \leq e_k(y) \quad \text{für alle } k \in \{1, \dots, n-1\},$$

und

$$e_n(x) = e_n(y)$$

erfüllen, die logarithmische Majorisierung (3.1.11) gilt. Für  $n \geq 3$  kann der Beweis nicht gelingen, wie das Gegenbeispiel  $x = (15, 3, 2)$ ,  $y = (10, 9, 1)$  belegt. Selbst eine Abschwächung auf die Aussage  $(|\log x_1|, \dots, |\log x_n|) \prec (|\log y_1|, \dots, |\log y_n|)$ , aus der immer noch die SSLI folgen würde, wird durch das Gegenbeispiel widerlegt.

**Proposition 3.1.8.** *Seien  $x_1, x_2, y_1, y_2 \in \mathbb{R}_+$  gegeben, sodass*

$$\begin{aligned} x_1 + x_2 &\leq y_1 + y_2, \\ x_1 x_2 &= y_1 y_2. \end{aligned} \quad (3.1.13)$$

*Dann gilt  $(\log x_1, \log x_2) \prec (\log y_1, \log y_2)$ .*

*Beweis.* Sei o.B.d.A.  $x_1 \geq x_2 > 0$  und  $y_1 \geq y_2 > 0$ . Unter der Voraussetzung  $x_1 x_2 = y_1 y_2$  gilt dann

$$\begin{aligned} x_1 + x_2 \leq y_1 + y_2 &\Leftrightarrow (x_1 + x_2)^2 \leq (y_1 + y_2)^2 \\ &\Leftrightarrow x_1^2 + 2x_1 x_2 + x_2^2 \leq y_1^2 + 2y_1 y_2 + y_2^2 \\ &\Leftrightarrow x_1^2 - 2x_1 x_2 + x_2^2 \leq y_1^2 - 2y_1 y_2 + y_2^2 \\ &\Leftrightarrow (x_1 - x_2)^2 \leq (y_1 - y_2)^2 \Leftrightarrow x_1 - x_2 \leq y_1 - y_2. \end{aligned}$$

### 3. Sum-of-Squared-Logarithms-Ungleichung

Die Addition von  $x_1 + x_2 \leq y_1 + y_2$  mit  $x_1 - x_2 \leq y_1 - y_2$  führt so auf  $x_1 \leq y_1$ . Unter Ausnutzung der Monotonie der Logarithmusfunktion erhalten wir

$$\begin{aligned}\log x_1 &\leq \log y_1, \\ \log x_1 + \log x_2 &\leq \log y_1 + \log y_2,\end{aligned}$$

was gleichbedeutend mit  $(\log x_1, \log x_2) \prec (\log y_1, \log y_2)$  (kurz geschrieben  $\log x \prec \log y$ ) ist. ■

#### 3.1.5. Gescheiterter Beweisversuch auf Basis von Majorization

Die Idee von Roventa war es, den Zusammenhang zwischen Nullstellen und kritischen Stellen eines Polynoms und den darauf ausgebauten Majorization-Begriff für einen Beweis der SSLI zu verwenden.

Seien  $x, y \in \mathbb{R}_+^n$  vorgegeben. Dazu sei  $P$  das normierte Polynom  $n$ -ten Grades mit den Nullstellen  $x_1, \dots, x_n$  und  $Q$  das normierte Polynom  $n$ -ten Grades mit den Nullstellen  $y_1, \dots, y_n$ . Wir bezeichnen den Vektor der Nullstellen der  $i$ -ten Ableitung von  $P$  bzw. von  $Q$  mit  $x^{(i)}$  bzw.  $y^{(i)}$  (im Falle  $i = 1$  mit  $x'$  bzw.  $y'$ ).

Dann gilt (im Sinne des verallgemeinerten Majorization-Begriffs)  $x \prec x'$  und  $y \prec y'$ . Zudem:  $x \prec y$  impliziert  $x' \prec y'$ , sowie  $x \prec_w y$  impliziert  $x' \prec_w y'$ .

Die Idee von Roventa war es nun, die Aussage

$$\text{Aus } e_k(x) \leq e_k(y) \text{ für alle } k \in \{1, \dots, n-1\} \text{ und } e_n(x) = e_n(y) \text{ folgt } |\log x| \prec_w |\log y|$$

zu zeigen, die wir in diesem Abschnitt bereits für  $n = 3$  (mit  $x = (15, 3, 2)$ ,  $y = (10, 9, 1)$ ) widerlegt haben. Wie dort auch gesehen, würde aus der Gültigkeit dieser Aussage sofort die Gültigkeit der SSLI folgen.

Der Beweis sollte per Induktion erfolgen und auf der Aussage

$$\text{Aus } e_k(x) \leq e_k(y) \text{ für alle } k \in \{1, \dots, n-1\} \text{ und } e_n(x) = e_n(y) \text{ folgt } |\log x^{(n-2)}| \prec_w |\log y^{(n-2)}|$$

beruhen. Für  $n = 3$  lautet diese Aussage

$$\text{Aus } e_1(x) \leq e_1(y), e_2(x) \leq e_2(y) \text{ und } e_3(x) = e_3(y) \text{ folgt } |\log x'| \prec_w |\log y'|.$$

Leider zeigte sich mit dem Gegenbeispiel  $x = (\frac{25}{9}, \frac{4}{3}, \frac{2}{45})$  und  $y = (\frac{10}{9}, \frac{2}{3}, \frac{2}{9})$  (welche auf die kritischen Punkte  $x' = \frac{187 \pm \sqrt{11359}}{135}$  und  $y' = \frac{18 \pm \sqrt{48}}{27}$  führen) auch hier, dass die Aussage falsch ist.

Der Umstand, dass  $x \prec x'$ ,  $y \prec y'$  und zudem  $x \prec y \Rightarrow x' \prec y'$ ,  $x \prec_w y \Rightarrow x' \prec_w y'$  gilt, nährt die Annahme, dass die Aussage

$$\begin{aligned}\text{Seien } x, y \in \mathbb{R}_+^n \text{ und } a, b \in \mathbb{R}_+^{n-1} \text{ mit } x \prec a \text{ und } y \prec b. \\ \text{Aus } x \prec y \text{ folgt nun } a \prec b \text{ und aus } x \prec_w y \text{ folgt } a \prec_w b\end{aligned}$$

wahr ist. Es war nun angedacht, die Induktion mittels dieser Aussage durchzuführen. Das Gegenbeispiel  $x = (\frac{5}{2}, 2, \frac{3}{2})$ ,  $y = (3, 2, 1)$ ,  $a = (\frac{7}{3}, \frac{5}{3})$  und  $b = (2, 2)$  zeigt nur leider, dass die Annahme inkorrekt ist.

Insgesamt sehen wir, dass sich zwar einige interessante Zusammenhänge zwischen Majorization und der SSLI auftun, der Beweis der SSLI mittels Majorization – sofern überhaupt möglich – nicht so einfach ist.

#### 3.1.6. Der Beweis für beliebige Dimension

Wir haben in Folgerung 2.13.13 auf Seite 64 gesehen, dass wir jede symmetrische<sup>2</sup> Funktion  $f: M \rightarrow \mathbb{C}$  mit  $M \subseteq \mathbb{C}^n$  als Funktion in den elementarsymmetrischen Polynomen ausdrücken können, es also eine Funktion  $\mathcal{F}: \mathbb{C}^n \rightarrow \mathbb{C}$  gibt mit

$$f(z_1, \dots, z_n) = \mathcal{F}(e_1, \dots, e_n) \quad \text{für alle } z \in M. \quad (3.1.14)$$

<sup>2</sup>Die Funktion ist invariant gegenüber Vertauschungen von Komponenten ihres Arguments.



### 3. Sum-of-Squared-Logarithms-Ungleichung

Wie in Definition 2.13.8 eingeführt, bezeichnen wir  $(z_1, \dots, z_n)$  als Nullstellen und  $(e_1, \dots, e_n)$  als Koeffizienten. Der Zusammenhang wird über das Polynom  $n$ -ten Grades

$$\begin{aligned} h &= (X - z_1)(X - z_2) \dots (X - z_n) \\ &= X^n - e_1 X^{n-1} + e_2 X^{n-2} + \dots + (-1)^n e_n \end{aligned} \quad (3.1.15)$$

hergestellt.

Definieren wir nun  $f(z) := \sum_{i=1}^n (\log z_i)^2$  für eine Menge  $M \subseteq \mathbb{C}^n$ , die garantiert, dass  $f(z) \in \mathbb{R}$  wohldefiniert ist, dann ist  $f$  vom Typ  $f(z) = \sum_{i=1}^n \phi(z_i)$  und somit symmetrisch und wir können mittels Folgerung 2.13.13 eine Funktion  $\mathcal{F}: T \rightarrow \mathbb{R}$  mit

$$\mathcal{F}(e_1, \dots, e_n) = \sum_{i=1}^n (\log z_i)^2 \quad \text{für alle } (e_1, \dots, e_n) \in T \quad (3.1.16)$$

mit  $T := \{(e_1, \dots, e_n) \mid (z_1, \dots, z_n) \in M\}$  definieren. Die Frage der Gültigkeit der SSLI wird somit zur Frage nach dem monotonen Wachstum der Funktion (3.1.16) in ihren ersten  $n - 1$ -Komponenten.

**Bemerkung 3.1.9.** Bevor wir fortfahren, machen wir uns zunächst klar, dass für jedes  $T \subseteq \mathbb{R}_+$  die Funktion  $\mathcal{F}$  aus (3.1.16) wohldefiniert ist: Nach Lemma 2.13.27 besteht  $(z_1, \dots, z_n)$  für alle  $(e_1, \dots, e_n) \in \mathbb{R}_+^n$  ausschließlich aus positiven reellen Zahlen und konjugiert komplexen Zahlenpaaren, insbesondere existiert zu jedem Koeffizientenvektor ein Nullstellenvektor. Die positiven reellen Zahlen sind unkritisch. Zudem ist  $(\log \bar{x})^2 = (\overline{\log x})^2 = (\log x)^2$ , so gilt für diese Paare  $x, \bar{x} \in \mathbb{C}$  somit  $(\log \bar{x})^2 + (\log x)^2 = 2 \operatorname{Re}(\log x)^2 \in \mathbb{R}$ .

Auf Basis der letzten Bemerkung können wir  $\mathcal{F}$  auch direkt über die Vorgabe von  $T \subseteq \mathbb{R}_+^n$  definieren:

**Definition 3.1.10.** Für die Menge der Koeffizientenvektoren  $T \subseteq \mathbb{R}_+^n$  definieren wir  $\mathcal{F}: T \rightarrow \mathbb{R}$  als die Funktion, für die

$$\mathcal{F}(e_1, \dots, e_n) = \sum_{i=1}^n (\log z_i)^2 \quad (3.1.17)$$

gilt. Die Menge der mit  $T$  korrespondierenden Nullstellenvektoren bezeichnen wir als

$$M = \{(z_1, \dots, z_n) \in \mathbb{C}^n \mid e(z) \in T\}, \quad (3.1.18)$$

also  $M = z(T)$  und  $T = e(M)$ . Definieren wir  $f: M \rightarrow \mathbb{R}$  mit  $f(z_1, \dots, z_n) := \sum_{i=1}^n (\log z_i)^2$ ,<sup>3</sup> so ist  $\mathcal{F} = f \circ z$  die Verkettung von  $f$  mit der Nullstellenabbildung  $z$ .

**Bemerkung 3.1.11.** Bei aller impliziten Definition von  $\mathcal{F}$  mittels komplexer Argumente: Ganz gleich, was „innerhalb“ von  $\mathcal{F}$  geschieht, die Funktion  $\mathcal{F}: T \subseteq \mathbb{R}_+^n \rightarrow \mathbb{R}$  ist eine rein reelle Funktion!

#### Beweisziel

Können wir für  $\mathcal{F}$  aus Definition 3.1.10 das monotone Wachstum in den ersten  $n - 1$ -Komponenten zeigen, dann folgt für vorgegebene  $x_1, \dots, x_n, \tilde{x}_1, \dots, \tilde{x}_n \in \mathbb{R}_+$  mit  $(e_1, \dots, e_n) = e(x_1, \dots, x_n)$  und  $(\tilde{e}_1, \dots, \tilde{e}_n) = e(\tilde{x}_1, \dots, \tilde{x}_n)$ , welche

$$e_k \leq \tilde{e}_k \quad \text{für alle } k \in \{1, \dots, n - 1\} \quad \text{und} \quad e_n = \tilde{e}_n \quad (3.1.19)$$

erfüllen, dass

$$\mathcal{F}(e_1, \dots, e_n) \leq \mathcal{F}(\tilde{e}_1, \dots, \tilde{e}_n), \quad (3.1.20)$$

gilt.

Diese Monotonie von  $\mathcal{F}$  in den Komponenten  $e_k$  wollen wir nun anhand der  $k$ -ten partiellen Ableitung von  $\mathcal{F}$  untersuchen. Können wir überhaupt sicherstellen, dass  $\mathcal{F}$  partiell differenzierbar ist? Erinnern wir uns an Folgerung 2.13.24 auf Seite 66, geschlossen aus einer Aussage von Ball [3, Theorem 3.2, S. 705], wonach eine Funktion  $\mathcal{F}(e_1, \dots, e_n) = \sum_{i=1}^n \phi(z_i)$  unendlich oft auf allen  $(e_1, \dots, e_n)$ , für die  $(z_1, \dots, z_n) \in \mathbb{R}_+^n$  gilt, differenzierbar ist, wenn  $\phi$  auf  $\mathbb{R}_+$  differenzierbar ist.

Die Menge  $T = \{(e_1, \dots, e_n) \mid z(e) \in \mathbb{R}_+^n\}$  auf der wir so die Differenzierbarkeit von  $\mathcal{F}$  zeigen können, reicht – wie wir im folgenden Unterabschnitt sehen werden – für den Beweis der SSLI nicht aus!

<sup>3</sup>Hierbei ist  $\log$  der Hauptzweig des komplexen Logarithmus, definiert als  $\log(x + iy) := \log|x + iy| + i \cdot \arctan \frac{y}{x}$

### 3. Sum-of-Squared-Logarithms-Ungleichung

#### Beweisschritte

Zu gegebenem  $T \subseteq \mathbb{R}_+^n$  unterteilt sich der Beweis der SSLI in zwei Beweisschritte:

- 1) Sind die partiellen Ableitungen von  $\mathcal{F}$  positiv? Also

$$\frac{\partial \mathcal{F}}{\partial e_k}(e_1, \dots, e_n) \geq 0 \quad \text{für alle } (e_1, \dots, e_n) \in T?$$

- 2) Gibt es in der Menge Nullstellenvektoren  $M$  einen Weg  $\gamma$  von  $x$  nach  $\tilde{x}$ , sodass zugehörige Koeffizientenvektoren in  $T$  bleiben und dabei alle  $e_k$  monoton wachsen? Also

$$\gamma: [0, 1] \rightarrow M \quad \text{mit} \quad \begin{cases} \gamma(0) = x, \gamma(1) = \tilde{x} \\ \frac{d}{ds} e_k(\gamma(s)) \geq 0 \quad \text{für alle } k \in \{1, \dots, n\} \end{cases} ?$$

Grundsätzlich ist Bedingung 1) nicht ausreichend. Um dies zu verstehen, schauen wir uns in Abbildung 3.1 den Graphen einer Funktion  $g: D \rightarrow \mathbb{R}$  mit nichtkonvexem Definitionsbereich  $D \subseteq \mathbb{R}^n$  an:

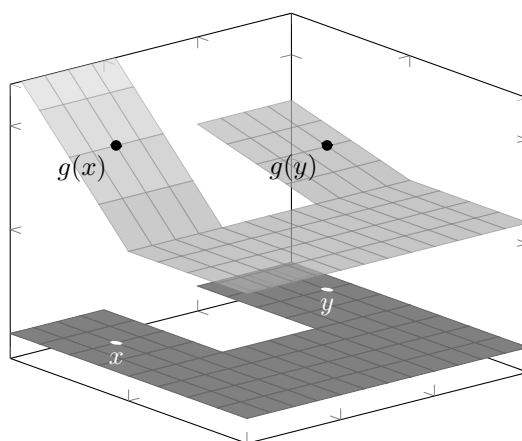


Abbildung 3.1.: Nicht monoton wachsende Funktion mit nichtkonvexem Definitionsbereich und überall nichtnegativen partiellen Ableitungen

Wie wir sehen, erfüllt die Funktion trotz ihrer nichtnegativen partiellen Ableitungen *nicht* für jedes Paar von Stellen  $x, y \in D$  mit (komponentenweise)  $x \leq y$  auch  $g(x) \leq g(y)$ , insbesondere finden wir dann keinen  $x$  mit  $y$  verbindenden, in den Komponenten wachsenden Weg.

So reicht die nichtkonvexe Menge  $T = \{(e_1, \dots, e_n) \mid z(e) \in \mathbb{R}_+^n\}$  nicht aus. Aber auch ohne die Folgerung aus der Aussage von Ball können wir die partielle Differenzierbarkeit einer für den Beweis hinreichend großen Menge  $T$  belegen. Für  $k = n$  scheitert der Beweis für das folgende Lemma, wäre dies nicht so, dann wäre die Bedingung  $e_n(x) = e_n(y)$  der SSLI nicht nötig, deren Notwendigkeit aber einfache Gegenbeispiele belegen.

#### 1) Positivität der partiellen Ableitungen

**Lemma 3.1.12.** Für alle Koeffizientenvektoren  $(e_1, \dots, e_n) \in \mathbb{R}_+^n$ , die auf paarweise unterschiedliche Nullstellen (d.h.  $j \neq i \Rightarrow z_j \neq z_i$  für alle  $i, j \in \{1, \dots, n\}$ ) führen, sind die ersten  $n - 1$  partiellen Ableitungen von  $\mathcal{F}$  stets positiv, d.h. sei  $\mathbf{k} \in \{1, \dots, n - 1\}$ , dann gilt

$$\frac{\partial \mathcal{F}}{\partial e_k}(e_1, \dots, e_n) > 0 \quad \text{für alle } (e_1, \dots, e_n) \in \mathbb{R}_+^n \text{ mit } z_j \neq z_i \text{ wenn } j \neq i. \quad (3.1.21)$$

### 3. Sum-of-Squared-Logarithms-Ungleichung

**Bemerkung 3.1.13.** Es sei trotz allem nochmal darauf hingewiesen, in der so gestalteten Menge  $T \subseteq \mathbb{R}_+^n$  fehlen alle  $(e_1, \dots, e_n)$  mit mehrfachen Nullstellen! Dadurch ist Konvexität der Menge  $T$  aber verletzt. Diese Lücke werden wir mit Dichtheitsargumenten schließen und die Aussage schlussendlich unter Einbüßen der Striktheit auf alle Koeffizientenvektoren aus  $\mathbb{R}_+^n$  ohne Einschränkung paarweise unterschiedlicher Nullstellen vervollständigen.

Wir haben in Folgerung 2.13.17 auf Seite 65 gezeigt, dass sich  $\mathcal{F}(e_1, \dots, e_n) = \sum_{i=1}^n \phi(z_i)$  im Falle paarweise unterschiedlicher Nullstellen per Kettenregel partiell differenzieren lässt mit

$$\frac{\partial \mathcal{F}}{\partial e_k}(e_1, \dots, e_n) = \sum_{i=1}^n \frac{(-1)^{k+1} z_i^{n-k}}{\prod_{\substack{j=1 \\ j \neq i}}^n (z_i - z_j)} \phi'(z_i) \quad \text{für alle } (e_1, \dots, e_n) \text{ mit } z_j \neq z_i, \text{ wenn } j \neq i. \quad (3.1.22)$$

Für  $\phi(z) = \log z$  ergibt sich die folgende Proposition sofort.

**Proposition 3.1.14.** *Für alle Koeffizientenvektoren  $(e_1, \dots, e_n) \in \mathbb{R}_+^n$ , die auf paarweise unterschiedliche Nullstellen (d.h.  $j \neq i \Rightarrow z_j \neq z_i$  für alle  $i, j \in \{1, \dots, n\}$ ) führen, ist  $\mathcal{F}$  partiell differenzierbar mit der Ableitung*

$$\frac{\partial \mathcal{F}}{\partial e_k}(e_1, \dots, e_n) = -2 \sum_{i=1}^n \frac{(-z_i)^{n-k-1}}{\prod_{j \neq i} (z_j - z_i)} \log z_i \quad \text{für } k \in \{1, \dots, n\}. \quad (3.1.23)$$

Wir wollen nun zeigen, dass die partiellen Ableitungen von  $\mathcal{F}$  positiv sind. Nun taucht der Ausdruck aus Proposition 3.1.14 schon in anderen Arbeiten auf. So stellen beispielsweise Mitchison und Josza im Jahr 2004 im Anhang von [49] fest, dass

$$(-1)^q \sum_{i=1}^n \frac{z_i^{n-q} \log z_i}{\prod_{\substack{j=1 \\ j \neq i}}^n (z_j - z_i)} \geq 0 \quad (3.1.24)$$

für alle  $q \in 2, \dots, n$  gilt. Damit scheinen wir unser Ziel bereits erreicht zu haben. Jedoch beweisen Mitchison und Josza die Ungleichung (3.1.24) nur für den Fall, dass  $z_1, \dots, z_n$  Eigenwerte einer Gram-Matrix sind, welche notwendigerweise symmetrisch und positiv definit ist. So kann die Ungleichung nur auf positive reelle Zahlen  $z_1, \dots, z_n$  angewandt werden. Dies reicht uns nicht aus, wir benötigen die Aussage auch für konjugiert komplexe Paare. Abhilfe schafft nun die folgende Proposition.

**Proposition 3.1.15.** *Sei  $(z_1, \dots, z_n) \in M$ , dann gilt für alle  $r \in \{0, \dots, n-2\}$*

$$-\sum_{i=1}^n \frac{(-z_i)^r}{\prod_{\substack{j=1 \\ j \neq i}}^n (z_j - z_i)} \log z_i = \int_0^\infty \frac{t^r}{\prod_{j=1}^n (t + z_j)} dt > 0. \quad (3.1.25)$$

*Beweis.* Nach Lemma 2.13.27 ist  $\prod_{j=1}^n (t + z_j) \geq \min\{t^n, \prod_{j=1}^n z_j\} > 0$  für alle  $t \in \mathbb{R}_+$ . Somit ist, da  $r \leq n-2$ ,

$$\int_0^\infty \left| \frac{t^r}{\prod_{j=1}^n (t + z_j)} \right| dt \leq \int_0^1 \frac{t^r}{e_n} dt + \int_1^\infty \underbrace{t^{r-n}}_{\leq t^{-2}} dt \leq \frac{1}{e_n} + 1. \quad (3.1.26)$$

Also konvergiert das Integral für  $r \in \{0, \dots, n-2\}$  und wegen  $\frac{t^r}{\prod_{j=1}^n (t+z_j)} > 0$  sogar gegen einen positiven

### 3. Sum-of-Squared-Logarithms-Ungleichung

Wert. Wir nutzen nun Lemma 2.9.6<sup>4</sup> und ziehen anschließend noch den Logarithmus auseinander:

$$\begin{aligned}
 & \int_0^\infty \frac{t^r}{\prod_j (t+z_j)} dt \stackrel{\text{Lemma 2.9.6}}{=} \int_0^\infty \sum_{i=1}^n \frac{1}{t+z_i} \cdot \frac{(-z_i)^r}{\prod_{j \neq i} ((-z_i) - (-z_j))} dt \\
 &= \sum_{i=1}^n \frac{(-z_i)^r}{\prod_{j \neq i} (z_j - z_i)} \cdot \log(t+z_i) \Big|_{t=0}^{t \rightarrow \infty} \\
 &= \sum_{i=1}^n \frac{(-z_i)^r}{\prod_{j \neq i} (z_j - z_i)} \cdot \left( \log t + \log\left(1 + \frac{z_i}{t}\right) \right) \Big|_{t=0}^{t \rightarrow \infty} \\
 &= \lim_{t \rightarrow \infty} \left( \sum_{i=1}^n \frac{(-z_i)^r}{\prod_{j \neq i} (z_j - z_i)} \cdot \log t + \sum_{i=1}^n \frac{(-z_i)^r}{\prod_{j \neq i} (z_j - z_i)} \cdot \log\left(1 + \frac{z_i}{t}\right) - \sum_{i=1}^n \frac{(-z_i)^r}{\prod_{j \neq i} (z_j - z_i)} \log z_i \right)
 \end{aligned}$$

Damit ist das Lemma bewiesen, denn:

Offenkundig ist  $\lim_{t \rightarrow \infty} \log\left(1 + \frac{z_i}{t}\right) = 0$ .

Allerdings ist auch  $\sum_{i=1}^n \frac{(-z_i)^r}{\prod_{j \neq i} (z_j - z_i)} = 0$ , ansonsten würde  $\lim_{t \rightarrow \infty} \sum_{i=1}^n \frac{(-z_i)^r}{\prod_{j \neq i} (z_j - z_i)} \log t$  divergieren, obwohl das Integral selbst konvergiert. Ein Widerspruch. Also ist

$$\int_0^\infty \frac{t^r}{\prod_j (t+z_j)} dt = - \sum_{i=1}^n \frac{(-z_i)^r}{\prod_{j \neq i} (z_j - z_i)} \log z_i$$

und wir erhalten die Aussage. ■

*Beweis von Lemma 3.1.12.* Wir können nun direkt zusammensetzen: Es gilt für alle  $n-k-1 \in \{0, \dots, n-2\}$ , also für  $k \in \{1, \dots, n-1\}$ :

$$\begin{aligned}
 \frac{\partial \mathcal{F}}{\partial e_k}(e_1, \dots, e_n) & \stackrel{\text{Prop. 3.1.14}}{=} -2 \sum_{i=1}^n \frac{(-z_i)^r}{\prod_{j \neq i} (z_j - z_i)} \log z_i \\
 & \stackrel{\text{Prop. 3.1.15}}{=} 2 \int_0^\infty \frac{t^r}{\prod_j (t+z_j)} dt > 0. \quad \blacksquare
 \end{aligned}$$

## 2) Weg durch $T$

Wir betrachten den Weg  $e^s \subseteq \mathbb{R}_+^n$  für  $s \in [0, 1]$ , der für  $s = 0$  bei  $e(x)$  startet und dann linear in den Koeffizienten mit wachsendem  $s$  ansteigt, um für  $s = 1$  bei  $e(y)$  zu enden:

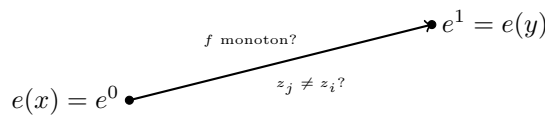


Abbildung 3.2.: in den Koeffizienten linearer Weg  $e^s$  von  $e(x)$  nach  $e(y)$

Wir definieren also

$$e_k^s := (1-s)e_k(x) + se_k(y) \quad \text{für } s \in [0, 1]. \quad (3.1.27)$$

<sup>4</sup>also eine Partialbruchzerlegung der Form  $\frac{t^r}{(t-a)(t-b)(t-c)} = \frac{a^r}{(a-b)(a-c)(t-a)} + \frac{b^r}{(b-a)(b-c)(t-b)} + \frac{c^r}{(c-a)(c-b)(t-c)}$

### 3. Sum-of-Squared-Logarithms-Ungleichung

Dann ist  $e^0 = e(x)$  und  $e^1 = e(y)$ , sowie

$$e_k(x) \leq e_k^s < e_k^{s'} \leq e_k(y) \quad \text{für alle } 0 \leq s < s' \leq 1 \text{ und alle } k \in \{1, \dots, n-1\}$$

und

$$e_n(x) = e_n^s = e_n(y) \quad \text{für alle } s \in [0, 1].$$

#### Bedingungen 1) und 2) nicht (ganz) erfüllt

Wir haben nun Bedingung 1) auf beinahe ganz  $T = \mathbb{R}_+^n$  gezeigt. Das Einzige, was uns Probleme bereitet, sind  $(e_1, \dots, e_n) \in T$ , die auf mehrfache Nullstellen führen.  $T$  ist zwar konvex, sodass unser Bedingung 2) erfüllende Weg  $e^s$  sicher ganz durch  $T$  verläuft. Dennoch können auf ihm mehrfache Nullstellen liegen. Wir werden nun im Folgenden zeigen, dass diese nicht ganz erfüllten Bedingungen schon ausreichend sind, um die Gültigkeit der SSLI zu beweisen.

#### Schließen der Lücken

Zunächst machen wir uns daran, die Anzahl der  $s$  für die  $e^s$  auf mehrfache Nullstellen führt, auf eine endliche Anzahl zu beschränken. Den Schlüssel hierzu bietet folgende Proposition.

**Proposition 3.1.16.** *Seien  $p_0$  und  $p_1$  Polynome  $n$ -ten Grades, die nicht beide mehrfache Nullstellen haben. Dann hat das Polynom  $p_s := (s-1)p_0 + sp_1$  nur für endliche viele  $s \in [0, 1]$  mehrfache Nullstellen.*

*Beweis.* Sei  $p = \alpha \prod_{i=1}^n (X - x_i) = \sum_{i=0}^n a_i X^{n-i}$  ein Polynom vom Grad  $n$  mit den Nullstellen  $z_1, \dots, z_n$  und den Koeffizienten  $a_0, \dots, a_n$ . Die Diskriminante von  $p$  ist definiert als

$$D(p) := \prod_{1 \leq i < j \leq n} (z_i - z_j)^2 \quad (3.1.28)$$

und ist genau dann Null, wenn  $p$  mehrfache Nullstellen besitzt. Die Diskriminante ist als ein in den  $z_1, \dots, z_n$  symmetrisches homogenes Polynom auch darstellbar als ein homogenes Polynom in den  $e_k(z_1, \dots, z_n)$  und somit in den Koeffizienten  $a_k$ .<sup>5</sup>

Die Koeffizienten von  $p_s$  sind Polynome in  $s$ , so ist auch  $D(s) := D(p_s)$  ein Polynom in  $s$ . Ist  $D(s)$  das Nullpolynom, so gilt  $D(0) = D(1) = 0$ , also haben sowohl  $p_0$  als auch  $p_1$  mehrfache Nullstellen. Umgekehrt folgt: Haben  $p_0$  und  $p_1$  nicht beide mehrfache Nullstellen, dann ist  $D(s)$  nicht das Nullpolynom und nur für endlich viele  $s \in [0, 1]$  kann  $f_s$  mehrfache Nullstellen besitzen. ■

Zum Preis einer weiteren Einschränkung haben wir somit gezeigt

**Folgerung 3.1.17.** *Für gegebene  $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}_+$  mit  $x_1 > x_2 > \dots > x_n$ ,  $y_1 \geq y_2 \geq \dots \geq y_n$  und den Weg*

$$e_k^s := (1-s)e_k(x) + se_k(y) \quad \text{für } s \in [0, 1] \quad (3.1.29)$$

*gibt es nur endliche viele  $s \in [0, 1]$ , für die  $e^s$  auf mehrfache Nullstellen führt.*

**Bemerkung 3.1.18.** Die Sortierung der  $x_i$  und der  $y_i$  ist aufgrund der Symmetrie von  $f$  keine weitere Einschränkung, sondern dient nur der einfacheren Formulierung.

Der Weg  $e^s$  ist somit in den elementarsymmetrischen Polynomen monoton wachsend, für alle bis auf endlich viele  $s \in [0, 1]$  sind die partiellen Ableitungen von  $\mathcal{F}$  positiv, es gelten somit die Bedingungen 1) und 2) für alle bis auf endlich viele  $s \in [0, 1]$ . Nun verwenden wir noch die Stetigkeit der Verkettung von  $\mathcal{F}$  mit  $e^s$ .

<sup>5</sup>So hat beispielsweise das Polynom  $q(t) = t^3 - at^2 + bt - c$  die Diskriminante  $D(q) = a^2b^2 - 4b^3 - 4a^3c + 18abc - 27c^2$ .

### 3. Sum-of-Squared-Logarithms-Ungleichung

**Lemma 3.1.19.** Seien  $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}_+$  mit  $x_1 > x_2 > \dots > x_n$ ,  $y_1 \geq y_2 \geq \dots \geq y_n$  und  $e_k(x) < e_k(y)$  für alle  $k \in \{1, \dots, n-1\}$  und  $e_n(x) = e_n(y)$ . Dann gilt  $f(x) < f(y)$ , also

$$\sum_{i=1}^n (\log x_i)^2 < \sum_{i=1}^n (\log y_i)^2. \quad (3.1.30)$$

*Beweis.* Wir betrachten die Verkettung  $s \mapsto f(s) := f(z(e^s))$ . Nach Lemma 2.13.10 ist die Abbildung der  $(e_1^s, \dots, e_n^s)$  auf die zugehörigen (eindeutig sortierten) Nullstellen stetig, zudem ist die Abbildung  $f(z) = \sum_{i=1}^n (\log z_i)^2$  offenkundig für alle  $z \in \mathbb{R}_+^n$  stetig. Damit ist  $s \mapsto f(s)$  stetig. Da nun  $s \mapsto e^s$  in den ersten  $n-1$  Komponenten streng monoton wächst (und in der letzten konstant) und  $\frac{\partial \mathcal{F}}{\partial e_k}$  für alle  $k \in \{1, \dots, n-1\}$  echt positiv ist, wächst  $s \mapsto f(s)$  für alle bis auf endlich viele  $s \in [0, 1]$  streng monoton. Die Stetigkeit von  $s \mapsto f(s)$  liefert monotoneres Wachstum auf ganz  $s \in [0, 1]$ . Somit gilt

$$f(x) = f(z(e^0)) = f(0) < f(1) = f(z(e^1)) = f(y). \quad \blacksquare$$

Im nächsten Schritt nutzen wir aus, dass die Menge der  $(e_1, \dots, e_n) \in T = \mathbb{R}_+^n$  mit paarweise unterschiedlichen  $(z_1, \dots, z_n)$  dicht in  $T$  liegt und schaffen es so, die Einschränkung  $x_1 > x_2 > \dots > x_n$  aus Lemma 3.1.19 zu umgehen.

**Lemma 3.1.20.** Seien  $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}_+$  mit  $e_k(x) < e_k(y)$  für alle  $k \in \{1, \dots, n-1\}$  und  $e_n(x) = e_n(y)$ . Dann gilt  $f(x) < f(y)$ , also

$$\sum_{i=1}^n (\log x_i)^2 < \sum_{i=1}^n (\log y_i)^2,$$

*Beweis.* Sind die Komponenten von  $x = (x_1, \dots, x_n)$  paarweise unterschiedlich, so können wir mit einer Umsortierung  $x_1 > x_2 > \dots > x_n$  und  $y_1 \geq y_2 \geq \dots \geq y_n$  erreichen, dann ist die Aussage bewiesen.

Im anderen Fall ist die Idee, die Komponenten von  $x$  jeweils paarweise um ein hinreichend kleines Stück zu bewegen und für das so entstandene  $x'$  immer noch  $e_k(x') < e_k(y)$  für alle  $k \in \{1, \dots, n-1\}$  und  $e_n(x') = e_n(y)$  zu erhalten. Erfolgt diese paarweise Bewegung von den wertgleichen  $x_i$  und  $x_j$  durch ein  $x'_i := x_i(1 + \varepsilon)$  und  $x'_j := x_j \frac{1}{1 + \varepsilon}$  mit  $\varepsilon > 0$ , so erhalten wir

$$\begin{aligned} (\log x'_i)^2 + (\log x'_j)^2 &= (\log x_i + \log(1 + \varepsilon))^2 + (\log x_j - \log(1 + \varepsilon))^2 \\ &= 2(\log x_i)^2 + 2(\log(1 + \varepsilon))^2 > (\log x_i)^2 + (\log x_j)^2. \end{aligned}$$

Lösen wir alle paarweise Gleichheiten auf diese Art auf, so folgt unter Verwendung von Lemma 3.1.19

$$\sum_{i=1}^n (\log x_i)^2 < \sum_{i=1}^n (\log x'_i)^2 \stackrel{3.1.19}{<} \sum_{i=1}^n (\log y_i)^2. \quad \blacksquare$$

Es ist jetzt noch möglich, dass

$$e_k(x) = e_k(y) \quad \text{für ein } k \in \{1, \dots, n-1\}$$

gilt und damit Lemma 3.1.20 trotz Erfülltheit aller SSLI Voraussetzungen nicht anwendbar ist. Der Schlüssel zur Lösung dieses Problems liegt in der Folge  $(e^m)_{m \in \mathbb{N}}$  mit

$$e_k^m := e_k(x) - \frac{1}{m} \quad \text{für alle } k \in \{1, \dots, n-1\} \quad \text{und} \quad e_n^m := e_n(x),$$

wie in Abbildung 3.3 angedeutet.

Unter Aufgabe der Stricktheit  $\sum (\log x_i) < \sum (\log y_i)$  und Ausnutzung der Stetigkeit des Zusammenhangs zwischen Koeffizienten und Nullstellen, sind wir nun abschließend in der Lage, Proposition 3.1.20 zu ergänzen und damit die SSLI in aller Gänze und ohne jede Einschränkung zu beweisen:

### 3. Sum-of-Squared-Logarithms-Ungleichung

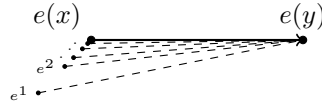


Abbildung 3.3.: Verwendung des Wegs von  $e^1$  nach  $e(y)$  anstelle von  $e(x)$  nach  $e(y)$  um Strikheit in allen elementarsymmetrischen Polynomen zu gewährleisten

Seien  $n \in \mathbb{N}$  und  $x, y \in \mathbb{R}_+^n$  sodass

$$e_k(x) \leq e_k(y) \quad \text{für alle } k \in \{1, \dots, n-1\},$$

und

$$e_n(x) = e_n(y).$$

Dann gilt

$$\sum_{i=1}^n (\log x_i)^2 \leq \sum_{i=1}^n (\log y_i)^2,$$

*Beweis der Sum-of-Squared-Logarithms-Ungleichung.* Wie oben schon beschrieben definieren wir

$$e_k^m := e_k(x) - \frac{1}{m} \quad \text{für alle } k \in \{1, \dots, n-1\} \quad \text{und} \quad e_n^m := e_n(x).$$

Für hinreichend große  $m \in \mathbb{N}$  erhalten wir damit  $0 < e_k^m < e_k(y)$ , sodass (zusammen mit  $e_n^m := e_n(x)$ ) Proposition 3.1.20 die Ungleichung

$$\sum_{i=1}^n \left( \log(z_i(e_1^m, \dots, e_n^m)) \right)^2 < \sum_{i=1}^n (\log y_i)^2$$

liefert. Die Nullstellenabbildung ist stetig in den Koeffizienten, so gilt

$$\lim_{m \rightarrow \infty} z_i(e_1^m, \dots, e_n^m) = x_i.$$

und wir erhalten schlussendlich

$$\sum_{i=1}^n (\log x_i)^2 = \lim_{m \rightarrow \infty} \sum_{i=1}^n \left( \log(z_i(e_1^m, \dots, e_n^m)) \right)^2 \leq \sum_{i=1}^n (\log y_i)^2. \quad \blacksquare$$

## 3.2. Verbindungen der SSLI zur Entropie

Jozsa und Mitchison [41] studieren Entropie und „Subentropie“ aus der Perspektive der Quanteninformationstheorie. Sie führen partielle Ableitungen dieser Größen nach  $e_k$  ein und untersuchen diese. Sie betrachten Entropie als eine Funktion in den elementarsymmetrischen Polynomen  $e_k$  und die analytische Fortsetzung um die Definition auf die ganze Menge der nichtnegativen reellen  $e_k$  zu erweitern. Folglich untersuchen sie Integralrepräsentationen der Entropie (und Subentropie), aus der die gewünschten Eigenschaften der Nichtnegativität folgen. Noch interessanter, sie studieren höhere partielle Ableitungen nach den elementarsymmetrischen Polynomen und stellen die komplette Monotonie der Entropie und Subentropie her. Es können ähnliche Monotonieeigenschaften höherer Ordnung für  $f(x) = \sum_{i=1}^n (\log x_i)^2$  hergestellt werden.

Dannan, Neff und Thiel [21] diskutieren Anwendungen der SSLI hinsichtlich der Entropie von Wahrscheinlichkeitsverteilungen. Wir beweisen nun eine Aussage, bezogen auf einen Ausdruck, der dem Entropieausdruck zweier Vektoren sehr ähnlich ist.

### 3. Sum-of-Squared-Logarithms-Ungleichung

Wir wiederholen unser Vorgehen des letzten Abschnitts.

Gegeben seien  $T = \mathbb{R}_+^n$  und  $M = \{(z_1, \dots, z_n) \in \mathbb{C}^n \mid e(z) \in T\}$  wie im Beweis der SSLI. Nun definieren wir  $\mathcal{G}: T \rightarrow \mathbb{R}$  und  $g: M \rightarrow \mathbb{R}$  mit

$$\mathcal{G}(e_1, \dots, e_n) := g(z_1, \dots, z_n) := - \sum_{i=1}^n z_i \log z_i. \quad (3.2.1)$$

Analog zum Beweis der SSLI verwenden wir die Nullstellenabbildung  $z: T \rightarrow M$ , um die Funktion  $\mathcal{G} = g \circ z$  als Verkettung zu definieren.

Die Wohldefiniertheit von  $\mathcal{G}$  ist sichergestellt: Wir wissen, nach Lemma 2.13.27 besteht  $(z_1, \dots, z_n)$  für alle  $(e_1, \dots, e_n) \in \mathbb{R}_+^n$  ausschließlich aus positiven reellen Zahlen und konjugiert komplexen Zahlenpaaren, insbesondere existiert zu jedem Koeffizientenvektor ein Nullstellenvektor. Für positive reelle Zahlen  $x \in \mathbb{R}_+$  ist  $x \log x \in \mathbb{R}$ , für konjugiert komplexe Paare  $x, \bar{x}$  gilt

$$x \log x + \bar{x} \log \bar{x} = x \log x + \bar{x} \cdot \overline{\log x} = x \log x + \overline{x \log x} \in \mathbb{R}.$$

$\mathcal{G}: T \rightarrow \mathbb{R}$  mit

$$\mathcal{G}(e_1, \dots, e_n) = \sum_{i=1}^n z_i \log z_i \quad \text{für alle } (e_1, \dots, e_n) \in T \quad (3.2.2)$$

Für alle  $(e_1, \dots, e_n) \in \mathbb{R}_+^n$ , die auf  $(z_1, \dots, z_n)$  mit  $z_j \neq z_i$  für  $j \neq i$  führen, gilt

$$\begin{aligned} \frac{\partial \mathcal{G}}{\partial e_k}(e_1, \dots, e_n) &= \sum_{i=1}^n \frac{(-1)^{k+1} z_i^{n-k}}{\prod_{\substack{j=1 \\ j \neq i}}^n (z_i - z_j)} \phi'(z_i) \\ &= - \sum_{i=1}^n \frac{(-1)^{k+1} z_i^{n-k}}{\prod_{\substack{j=1 \\ j \neq i}}^n (z_i - z_j)} (\log z_i + 1) \\ &= - \sum_{i=1}^n \frac{(-1)^{k+1} z_i^{n-k}}{\prod_{\substack{j=1 \\ j \neq i}}^n (z_i - z_j)} \log z_i - \underbrace{\sum_{i=1}^n \frac{(-1)^{k+1} z_i^{n-k}}{\prod_{\substack{j=1 \\ j \neq i}}^n (z_i - z_j)}}_{=0 \text{ für } k \in \{2, \dots, n\}}. \end{aligned} \quad (3.2.3)$$

Hierbei machen wir uns klar, dass der zweite Summand in (3.2.3) für alle  $k \in \{2, \dots, n\}$ <sup>6</sup> aufgrund von Folgerung 2.9.7<sup>7</sup> verschwindet und wir einen Ausdruck erhalten, dessen Positivität im letzten Kapitel schon eine entscheidende Rolle eingenommen hat. Er korrespondiert (paarweise unterschiedliche Nullstellen vorausgesetzt) mit Proposition 3.1.15

$$2 \frac{\partial \mathcal{G}}{\partial e_{k+1}}(e_1, \dots, e_n) = \frac{\partial \mathcal{F}}{\partial e_k}(e_1, \dots, e_n) > 0 \quad \text{für alle } k \in \{1, \dots, n-1\}. \quad (3.2.4)$$

Wie beim Beweis der SSLI können wir nun analog zu den dort geführten Beweisschritten die Monotonie von  $\mathcal{G}$ , diesmal in den letzten  $n-1$  Komponenten, zeigen. So erhalten wir das folgende Resultat:

**Folgerung 3.2.1.** *Seien  $n \in \mathbb{N}$  und  $x, y \in \mathbb{R}_+^n$ , sodass*

$$e_1(x) = e_1(y). \quad (3.2.5)$$

und

$$e_k(x) \leq e_k(y) \quad \text{für alle } k \in \{2, \dots, n\}, \quad (3.2.6)$$

Dann gilt

$$- \sum_{i=1}^n x_i \log x_i \leq - \sum_{i=1}^n y_i \log y_i. \quad (3.2.7)$$

<sup>6</sup>was gleichbedeutend ist mit  $n-k \in \{0, \dots, n-2\}$

<sup>7</sup>also eine Identität der Form  $\frac{a^2}{(a-b)(a-c)(a-d)} + \frac{b^2}{(b-a)(b-c)(b-d)} + \frac{c^2}{(c-a)(c-b)(c-d)} + \frac{d^2}{(d-a)(d-b)(d-c)} = 0$



### 3. Sum-of-Squared-Logarithms-Ungleichung

Wenn wir die erste Bedingung abschwächen und nur noch  $e_k(x) \leq e_k(y)$  für alle  $k \in \{1, \dots, n\}$  fordern, dann ergeben sich offenkundig ähnliche Gegenbeispiele, wie bei der SSLI: Für  $x = (e^{-1}, \dots, e^{-1}) \in \mathbb{R}_+^n$  und  $y = (1, \dots, 1) \in \mathbb{R}_+^n$  ist die abgeschwächte Bedingung erfüllt, aber es ist  $-\sum_{i=1}^n x_i \log x_i = \frac{n}{e} > 0 = -\sum_{i=1}^n y_i \log y_i$ .

### 3.3. Andere Bedingungen für die SSLI und die Sum-of-Powered-Logarithms-Ungleichung

Zunächst wollen wir in den folgenden Theoremen weitere Bedingungen angeben, welche die Gültigkeit der SSLI sicherstellen.

**Satz 3.3.1.** *Seien  $a, b \in \mathbb{R}_+^n$ . Weiter existiere eine Umordnung von  $a$  und  $b$ , welche*

$$\frac{b_1}{a_1} \geq \frac{b_2}{a_2} \geq \dots \geq \frac{b_n}{a_n} \quad \text{und} \quad a_1 b_1 \geq a_2 b_2 \geq \dots \geq a_n b_n \quad (3.3.1)$$

erfüllt. Wenn wir nun zusätzlich eine der folgenden beiden Bedingungen annehmen

$$e_n(a) \leq e_n(b) \quad \text{und} \quad e_n(a) \cdot e_n(b) \geq 1, \quad (3.3.2a)$$

oder

$$e_n(a) \geq e_n(b) \quad \text{und} \quad e_n(a) \cdot e_n(b) \leq 1, \quad (3.3.2b)$$

dann erhalten wir die Sum-of-Squared-Logarithms-Ungleichung (3.1.1):

$$\sum_{i=1}^n (\log a_i)^2 \leq \sum_{i=1}^n (\log b_i)^2.$$

*Beweis.* Zunächst nehmen wir an, dass Bedingung (3.3.2a) gilt:

Wegen der Monotonie des Logarithmus folgt aus den Annahmen, dass

$$\log \frac{b_1}{a_1} \geq \log \frac{b_2}{a_2} \geq \dots \geq \log \frac{b_n}{a_n} \quad \text{und} \quad \log(a_1 b_1) \geq \log(a_2 b_2) \geq \dots \geq \log(a_n b_n).$$

Nun können wir mittels der Tschebyscheff-Summenungleichung (2.9.3) mit  $\tilde{a}_{n+k-1} := \log \frac{b_i}{a_i}$  und  $\tilde{b}_{n+k-1} := \log a_i b_i$  abschätzen:

$$\begin{aligned} \sum_{i=1}^n (\log b_i)^2 - \sum_{i=1}^n (\log a_i)^2 &= \sum_{i=1}^n ((\log b_i)^2 - (\log a_i)^2) = \sum_{i=1}^n (\log b_i - \log a_i)(\log b_i + \log a_i) \\ &= \sum_{i=1}^n \log \frac{b_i}{a_i} \cdot \log b_i a_i \stackrel{(2.9.3)}{\geq} \frac{1}{n} \left( \sum_{i=1}^n \log \frac{b_i}{a_i} \right) \left( \sum_{i=1}^n \log b_i a_i \right) \\ &= \frac{1}{n} \left( \log \underbrace{\prod_{i=1}^n \frac{b_i}{a_i}}_{e_n(b)/e_n(a)} \right) \left( \log \underbrace{\prod_{i=1}^n b_i a_i}_{e_n(a) e_n(b)} \right) \geq \frac{1}{n} \log 1 \log 1 = 0. \end{aligned}$$

Schließlich ist  $\sum_{i=1}^n (\log y_i)^2 - \sum_{i=1}^n (\log a_i)^2 \geq 0$  äquivalent zu (3.1.1).

Als nächstes nehmen wir an, dass Bedingung (3.3.2b) gilt:

Wir setzen  $\tilde{a}, \tilde{b} \in \mathbb{R}_+^n$  mit  $\tilde{a}_k := \frac{1}{a_{n+1-k}}$  und  $\tilde{b}_k := \frac{1}{b_{n+1-k}}$  für  $k \in \{1, \dots, n\}$ , sodass wir  $(\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n) = (a_n^{-1}, a_{n-1}^{-1}, \dots, a_1^{-1})$  und  $(\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_n) = (b_n^{-1}, b_{n-1}^{-1}, \dots, b_1^{-1})$  erhalten. Dann

$$\frac{\tilde{b}_k}{\tilde{a}_k} = \frac{\frac{1}{b_{n+1-k}}}{\frac{1}{a_{n+1-k}}} = \frac{a_{n+1-k}}{b_{n+1-k}} \quad \text{und} \quad \tilde{a}_k \tilde{b}_k = \frac{1}{a_{n+1-k} b_{n+1-k}}.$$

### 3. Sum-of-Squared-Logarithms-Ungleichung

Daher

$$\frac{\tilde{b}_1}{\tilde{a}_1} = \frac{a_n}{b_n} \geq \frac{\tilde{b}_2}{\tilde{a}_2} = \frac{a_{n-1}}{b_{n-1}} \geq \dots \geq \frac{\tilde{b}_n}{\tilde{a}_n} \quad \text{und} \quad \tilde{a}_1 \tilde{b}_1 = \frac{1}{a_n b_n} \geq \tilde{a}_2 \tilde{b}_2 = \frac{1}{a_{n-1} b_{n-1}} \geq \dots \geq \tilde{a}_n \tilde{b}_n.$$

Außerdem

$$e_n(\tilde{b}) = \frac{1}{e_n(\tilde{a})} \geq \frac{1}{e_n(a)} = e_n(\tilde{a}) \quad \text{und} \quad e_n(\tilde{a})e_n(\tilde{b}) = \frac{1}{e_n(a)} \cdot \frac{1}{e_n(b)} \geq 1,$$

somit erfüllen  $\tilde{a}$  und  $\tilde{b}$  Bedingung (3.3.2a). Deshalb ist die (3.1.1) für  $\tilde{a}$  und  $\tilde{b}$  erfüllt, und wir können umformen:

$$\begin{aligned} \sum_{i=1}^n (\log \tilde{a}_i)^2 &\leq \sum_{i=1}^n (\log \tilde{b}_i)^2 &\Leftrightarrow &\sum_{i=1}^n \left( \log \frac{1}{a_{n+1-i}} \right)^2 \leq \sum_{i=1}^n \left( \log \frac{1}{b_{n+1-i}} \right)^2 \\ \Leftrightarrow \sum_{i=1}^n \left( \log \frac{1}{a_i} \right)^2 &\leq \sum_{i=1}^n \left( \log \frac{1}{b_i} \right)^2 &\Leftrightarrow &\sum_{i=1}^n (-\log a_i)^2 \leq \sum_{i=1}^n (-\log b_i)^2 &(3.3.3) \\ \Leftrightarrow \sum_{i=1}^n (\log a_i)^2 &\leq \sum_{i=1}^n (\log b_i)^2. &&& \blacksquare \end{aligned}$$

**Beispiel 3.3.2.** Weder sind die Bedingungen der SSLI strenger als in Satz 3.3.1 noch umgekehrt:

- i) Mit  $a = (14, 2, 10)$  und  $b = (20, 2, 7)$  erhalten wir  $e_k(a) \leq e_k(b)$  für alle  $k \in \{1, \dots, n-1\}$  und  $e_n(a) = e_n(b)$ , aber es existiert keine Umordnung von  $a$  und  $b$ , die  $\frac{b_1}{a_1} \geq \frac{b_2}{a_2} \geq \frac{b_3}{a_3}$  und  $a_1 b_1 \geq a_2 b_2 \geq a_3 b_3$  erfüllt.
- ii) Mit  $a = (6, 5, 7)$  und  $b = (10, 8, 3)$  erhalten wir  $\frac{10}{6} \geq \frac{8}{5} \geq \frac{3}{7}$  und  $6 \cdot 10 \geq 5 \cdot 8 \geq 7 \cdot 3$ . Wegen  $e_n(a) = 210$  und  $e_n(b) = 240$  erhalten wir  $e_n(a) \leq e_n(b)$  und  $e_n(a)e_n(b) \geq 1$ , aber nicht  $e_n(a) = e_n(b)$ .
- iii) Mit  $a = (2, 2, 2)$  und  $b = (4, 2, 1)$  erhalten wir  $e_k(a) \leq e_k(b)$  für alle  $k \in \{1, \dots, n-1\}$  und  $e_n(a) = e_n(b)$ . Außerdem  $\frac{4}{2} \geq \frac{2}{2} \geq \frac{1}{2}$ ,  $2 \cdot 4 \geq 2 \cdot 2 \geq 2 \cdot 1$  and  $e_n(a)e_n(b) \geq 1$ .

**Satz 3.3.3** (Sum-of-Powered-Logarithms-Ungleichung). *Seien  $a, b \in \mathbb{R}^n$  und  $p \in \mathbb{R}$  mit  $a_i > 1$ ,  $b_i > 1$  und  $p < 0$ . Angenommen  $a \prec^w b$ , dann gilt*

$$\sum_{i=1}^n (\log a_i)^p \leq \sum_{i=1}^n (\log b_i)^p. \quad (3.3.4)$$

**Bemerkung 3.3.4.** Wir setzen  $a_i > 1$  und  $b_i > 1$  für die Wohldefiniertheit von  $(\log a_i)^p$  und  $(\log b_i)^p$  für alle  $p \in \mathbb{R}$  voraus.

*Beweis.* Aus  $a \prec^w b$  mit Proposition 2.12.10 erhalten wir

$$\sum_{i=1}^k \log a_i^\uparrow \geq \sum_{i=1}^k \log b_i^\uparrow \quad \text{für alle } k \in \{1, \dots, n\}.$$

Seien  $x, y \in \mathbb{R}_+^n$  mit  $x := \log a$ ,  $y := \log b$  (damit  $x_i := \log a_i$  und  $y_i := \log b_i$  für alle  $k \in \{1, \dots, n\}$ ), dann  $x \prec^w y$ . Betrachten wir nun die Funktion  $g: \mathbb{R}_+ \rightarrow \mathbb{R}$  mit  $g(z) = z^p$ , dann

$$g'(z) = \underbrace{p}_{<0} \cdot \underbrace{z^{p-1}}_{>0} < 0 \quad \text{und} \quad g''(z) = \underbrace{p(p-1)}_{>0} \cdot \underbrace{z^{p-2}}_{>0} > 0.$$

Aus diesem Grund ist  $g$  monoton fallend und konvex. Mittels Proposition 2.12.8 ii) erhalten wir, dass  $\varphi: \mathbb{R}_+^n \rightarrow \mathbb{R}$  mit  $\varphi(\xi) = \sum_{i=1}^n \xi_i^p$  monoton fallend und Schur-konvex ist. Mit Proposition 2.12.7 erhalten wir  $\sum_{i=1}^n x_i^p \leq \sum_{i=1}^n y_i^p$ . Rücksubstitution liefert uns direkt die Aussage.  $\blacksquare$

### 3. Sum-of-Squared-Logarithms-Ungleichung

**Bemerkung 3.3.5.**  $a \prec b$  ist eine zu schwache und  $a \prec b$  plus  $e_n(a) = e_n(b)$  eine zu starke Bedingung für (3.1.1):

Für  $a = (3, 2, 2)$  und  $b = (4, 2, 1)$  erhalten wir  $a \prec b$ , aber  $2.17 \approx \sum (\log a_i)^2 < \sum (\log b_i)^2 \approx 2.40$ .

Für  $a = (4, 4, 4)$  und  $b = (10, 1, 1)$  erhalten wir  $a \prec b$ , aber  $5.77 \approx \sum (\log a_i)^2 > \sum (\log b_i)^2 \approx 5.30$ .

Durch Lemma 2.13.43 haben wir gezeigt, dass  $a \prec b$  impliziert  $e_k(a) \geq e_k(b)$  (Es sei auf die umgekehrte Ungleichung hingewiesen!).

Wie sieht es aus,  $a \prec b$  und  $e_n(a) = e_n(b)$  als hinreichende Bedingungen für (3.1.1) zu fordern?

Wir können recht einfach zeigen:  $a \prec b$  und  $e_n(a) = e_n(b)$  implizieren die logarithmische Majorization  $\log a \prec \log b$ . Weil die Abbildung  $t \mapsto t^2$  konvex ist, folgt mit Proposition 2.12.8, dass die Ungleichung  $\sum_{i=1}^n (\log a_i)^2 \leq \sum_{i=1}^n (\log b_i)^2$  erfüllt ist. Jedoch, für  $a \prec b$ , können wir Folgerung 2.13.43 anwenden, um  $e_k(a) \geq e_k(b)$  zu finden, daher impliziert die SSLI  $\sum_{i=1}^n (\log a_i)^2 \geq \sum_{i=1}^n (\log b_i)^2$  und damit  $\sum_{i=1}^n (\log a_i)^2 = \sum_{i=1}^n (\log b_i)^2$ . Dieser Umstand ist nicht überraschend: Wir wissen bereits (siehe Bemerkung 2.13.45), dass

$$a \prec b \quad \text{und} \quad e_n(a) = e_n(b) \quad \Rightarrow \quad a^\downarrow = b^\downarrow$$

gilt. Somit sind die Vektoren  $a$  und  $b \in \mathbb{R}_+^n$  bis auf Permutation identisch.

## 3.4. Verwandte Ungleichungen

**Proposition 3.4.1.** Seien  $x, y \in \mathbb{R}_+^n$  und  $m \in \mathbb{R}_+$ . Nehmen wir zudem  $\log x \prec_w \log y$  an, dann gilt

$$\sum_{i=1}^n e^{m x_i} \leq \sum_{i=1}^n e^{m y_i}. \quad (3.4.1)$$

*Beweis.* Wir setzen  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $\varphi(x) = \sum_{i=1}^n e^{m x_i}$ . Da  $x \mapsto e^{m x}$  konvex und monoton wachsend ist, ist die Funktion  $\varphi$  gemäß Proposition 2.12.8 ii) Schur-konvex und monoton wachsend. Mit Proposition 2.12.10 erhalten wir  $x \prec_w y$ , so zeigt (2.12.7) i), dass  $\varphi(x) \leq \varphi(y)$ . ■

**Satz 3.4.2.** Wenn die reellen Zahlen  $a, b, c, x, y, z$

$$a + b + c = x + y + z = 0 \quad (3.4.2)$$

und

$$a^2 + b^2 + c^2 = x^2 + y^2 + z^2 \neq 0, \quad (3.4.3)$$

erfüllen, dann

$$e^{xy} + e^{yz} + e^{zx} < (e^{x^2} + e^{y^2} + e^{z^2}) \exp\left(-3\sqrt[3]{\frac{1}{4}a^2b^2c^2}\right) \quad (3.4.4)$$

und

$$e^{ab} + e^{bc} + e^{ca} < (e^{a^2} + e^{b^2} + e^{c^2}) \exp\left(-3\sqrt[3]{\frac{1}{4}x^2y^2z^2}\right). \quad (3.4.5)$$

*Beweis.* Für  $\alpha, \beta, \gamma \in \mathbb{R}$  erhalten wir  $(\alpha + \beta + \gamma)^2 = \alpha^2 + \beta^2 + \gamma^2 + 2\alpha\beta + 2\beta\gamma + 2\gamma\alpha$ , damit gilt unter den Voraussetzungen

$$\begin{aligned} ab + bc + ca &= \frac{1}{2}((a + b + c)^2 - (a^2 + b^2 + c^2)) \\ &= \frac{1}{2}((x + y + z)^2 - (x^2 + y^2 + z^2)) = xy + yz + zx \end{aligned}$$

und wir können

$$p := ab + bc + ca = xy + yz + zx.$$

### 3. Sum-of-Squared-Logarithms-Ungleichung

setzen. Weiterhin

$$x^3 + px - xyz = x^3 + x^2y + xyz + x^2z - xyz = x^2(x + y + z) = x^2 \cdot 0 = 0$$

und analog

$$y^3 + py - xyz = 0 \quad \text{und} \quad z^3 + pz - xyz = 0.$$

Somit hat die kubische Gleichung  $X^3 + pX - xyz = 0$  genau die drei Lösungen  $X \in \{x, y, z\}$ .

Mit den Cardanischen Formeln (siehe Cardano [14]), hat die kubische Gleichung  $X^3 + pX + q = 0$  genau dann drei reelle Lösungen, wenn

$$D := \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 < 0.$$

Somit erhalten wir

$$\left(\frac{-xyz}{2}\right)^2 + \left(\frac{p}{3}\right)^3 = \frac{x^2y^2z^2}{4} + \frac{p^3}{27} < 0,$$

also  $p^3 < -27\frac{x^2y^2z^2}{4}$  und schließlich

$$p = xy + yz + zx < -3\sqrt[3]{\frac{x^2y^2z^2}{4}}. \quad (3.4.6)$$

Auf analoge Weise erhalten wir aus  $X^3 + pX - abc = 0$

$$p = ab + bc + ca < -3\sqrt[3]{\frac{a^2b^2c^2}{4}}. \quad (3.4.7)$$

Also

$$p = p + x^2 - x^2 = xy + yz + zx + x^2 - x^2 = yz - x^2,$$

und damit

$$yz = x^2 + p$$

und analog  $xz = y^2 + p$  und  $xy = z^2 + p$ . Gemäß (3.4.7) und der Monotonie der Exponentialfunktion haben wir

$$e^{yz} < e^{x^2} \exp\left(-3\sqrt[3]{\frac{a^2b^2c^2}{4}}\right), \quad e^{zx} < e^{y^2} \exp\left(-3\sqrt[3]{\frac{a^2b^2c^2}{4}}\right), \quad e^{xy} < e^{z^2} \exp\left(-3\sqrt[3]{\frac{a^2b^2c^2}{4}}\right)$$

und durch Summenbildung erhalten wir (3.4.4). Der Beweis von (3.4.5) erfolgt analog. ■

**Satz 3.4.3.** Für  $I \subseteq \mathbb{R}$  seien Funktionen  $f_1, \dots, f_n: I \rightarrow \mathbb{R}$  mit

$$\sum_{i=1}^n f_i(t) = 0 \quad \text{und} \quad f_1(t) \leq f_2(t) \leq \dots \leq f_n(t) \quad \text{für alle } t \in I \quad (3.4.8)$$

und  $g: I \rightarrow \mathbb{R}$  mit

$$g(t) = \sum_{i=1}^n e^{f_i(t)} \quad \text{für alle } t \in I \quad (3.4.9)$$

gegeben.

- i) Falls  $f'_1(t) \leq f'_2(t) \leq \dots \leq f'_n(t)$  für alle  $t \in I$ , dann ist  $g'(t) \geq 0$  für alle  $t \in I$  und  $g$  ist monoton wachsend.
- ii) Falls  $f'_1(t) \geq f'_2(t) \geq \dots \geq f'_n(t)$  für alle  $t \in I$ , dann ist  $g'(t) \leq 0$  für alle  $t \in I$  und  $g$  ist monoton fallend.

### 3. Sum-of-Squared-Logarithms-Ungleichung

*Beweis.* Die Bedingung  $\sum_{i=1}^n f_i(t) = 0$  für alle  $t \in I$  impliziert

$$\sum_{i=1}^n f_i'(t) = \frac{d}{dt} \left( \sum_{i=1}^n f_i(t) \right) = 0 \quad \text{für alle } t \in I.$$

Nehmen wir als erstes an, dass gilt:  $f_1'(t) \leq f_2'(t) \leq \dots \leq f_n'(t)$  für alle  $t \in I$ . Mit der Monotonie der Exponentialfunktion und der Tschebyscheff-Summenungleichung mit  $a_i := e^{f_i(t)}$ ,  $b_i := f_i'(t)$  erhalten wir

$$g'(t) = \sum_{i=1}^n f_i'(t) \cdot e^{f_i(t)} \stackrel{(2.9.3)}{\geq} \frac{1}{n} \underbrace{\left( \sum_{i=1}^n f_i'(t) \right)}_{=0} \left( \sum_{i=1}^n e^{f_i(t)} \right) = 0.$$

Nun nehmen wir stattdessen an, dass  $f_1'(t) \geq f_2'(t) \geq \dots \geq f_n'(t)$  für alle  $t \in I$  gilt. Mit der Monotonie der Exponentialfunktion und der Tschebyscheff-Summenungleichung mit  $a_i := e^{f_i(t)}$ ,  $b_{n+k-i} := f_i'(t)$  erhalten wir

$$g'(t) = \sum_{i=1}^n f_i'(t) \cdot e^{f_i(t)} \stackrel{(2.9.3)}{\leq} \frac{1}{n} \underbrace{\left( \sum_{i=1}^n f_i'(t) \right)}_{=0} \left( \sum_{i=1}^n e^{f_i(t)} \right) = 0. \quad \blacksquare$$

**Beispiel 3.4.4.** Gegeben seien die Funktionen  $f_1, f_2, f_3: \mathbb{R} \rightarrow \mathbb{R}$  mit

$$f_1(x) = -x^2 + 1, \quad f_2(x) = x - 1, \quad f_3(x) = x^2 - x$$

und

$$f_1'(x) = -2x, \quad f_2'(x) = 1, \quad f_3'(x) = 2x - 1.$$

Dann gilt

$$f_1(x) + f_2(x) + f_3(x) = 0 \quad \text{für alle } x \in \mathbb{R}.$$

Zudem ist

$$\begin{aligned} f_1(x) \leq f_2(x) \leq f_3(x) \quad \text{und} \quad f_1'(x) \leq f_2'(x) \leq f_3'(x) & \quad \text{für alle } x \in [1, \infty), \\ f_2(x) \leq f_3(x) \leq f_1(x) \quad \text{und} \quad f_1'(x) \geq f_2'(x) \geq f_3'(x) & \quad \text{für alle } x \in \left[ \frac{1}{4}, 1 \right]. \end{aligned}$$

Definieren wir nun  $g: \mathbb{R} \rightarrow \mathbb{R}$  mit

$$g(x) = e^{-x^2+1} + e^{-1+x} + e^{x^2-x} = e^{f_1(x)} + e^{f_2(x)} + e^{f_3(x)},$$

dann können wir mit Satz 3.4.3 schließen:  $g$  ist monoton wachsend auf  $[1, \infty)$  und monoton fallend auf  $[\frac{1}{4}, 1]$ .

Jetzt verallgemeinern wir Satz 3.4.3.

**Satz 3.4.5.** Seien  $I \subseteq \mathbb{R}$  und  $f_1, \dots, f_n: I \rightarrow \mathbb{R}$  mit

$$\sum_{i=1}^n f_i(t) = 0, \quad f_1(t) \leq f_2(t) \leq \dots \leq f_n(t) \quad \text{und} \quad f_1'(t) \leq f_2'(t) \leq \dots \leq f_n'(t) \quad \text{für alle } t \in I \tag{3.4.10}$$

und nehmen wir  $g, h: I \rightarrow \mathbb{R}$ , wobei  $h$  positiv und monoton wachsend ist und

$$g_h(t) = \sum_{i=1}^n e^{h(t)f_i(t)} \quad \text{für alle } t \in I. \tag{3.4.11}$$

Dann ist  $g_h'(t) \geq 0$  für alle  $t \in I$ , woraus folgt, dass  $g_h$  monoton wachsend ist.

### 3. Sum-of-Squared-Logarithms-Ungleichung

*Beweis.* Die Bedingung  $\sum_{i=1}^n f_i(t) = 0$  für alle  $t \in I$  impliziert

$$\sum_{i=1}^n f'_i(t) = \frac{d}{dt} \left( \sum_{i=1}^n f_i(t) \right) = 0 \quad \text{für alle } t \in I.$$

Mit der Tschebyscheff-Summenungleichung, verwendend  $a_i := f_i(t)$  bzw.  $a_i := f'_i(t)$  und  $b_i := e^{h(t)f_i(t)}$  können wir schließen:

$$\begin{aligned} g'_h(t) &= \sum_{i=1}^n (h'(t)f_i(t) + h(t)f'_i(t)) e^{h(t)f_i(t)} \\ &= h'(t) \left( \sum_{i=1}^n f_i(t) \cdot e^{h(t)f_i(t)} \right) + h(t) \left( \sum_{i=1}^n f'_i(t) \cdot e^{h(t)f_i(t)} \right) \\ &\stackrel{(2.9.3)}{\geq} h'(t) \cdot \underbrace{\frac{1}{n} \left( \sum_{i=1}^n f_i(t) \right)}_{=0} \left( \sum_{i=1}^n e^{h(t)f_i(t)} \right) + h(t) \left( \sum_{i=1}^n \underbrace{f'_i(t)}_{=0} \right) \left( \sum_{i=1}^n e^{h(t)f_i(t)} \right) = 0. \quad \blacksquare \end{aligned} \tag{3.4.12}$$

**Satz 3.4.6.** Für  $I \subseteq \mathbb{R}$  sei  $f_1, \dots, f_n: I \rightarrow \mathbb{R}$  mit den Eigenschaften

$$\sum_{i=1}^n f_i(t) = 0, \quad f_1(t) \leq f_2(t) \leq \dots \leq f_n(t) \quad \text{und} \quad f'_1(t) \leq f'_2(t) \leq \dots \leq f'_n(t) \tag{3.4.13}$$

für alle  $t \in I$ . Zusätzlich ist  $h: D \rightarrow \mathbb{R}$  monoton wachsend und konvex vorgegeben. Die Menge  $D \subseteq \mathbb{R}$  stellt hierbei sicher, dass  $h(f_i(t))$  für alle  $i \in \{1, \dots, n\}$  und alle  $t \in I$  wohldefiniert ist. Weiterhin definieren wir  $H: I \rightarrow \mathbb{R}$  mit

$$H(t) = \sum_{i=1}^n e^{h(f_i(t))}.$$

Dann ist  $H'(t) \geq 0$  für alle  $t \in I$  und  $H$  ist monoton wachsend.

*Beweis.* Die Bedingung  $\sum_{i=1}^n f_i(t) = 0$  für alle  $t \in I$  impliziert

$$\sum_{i=1}^n f'_i(t) = \frac{d}{dt} \left( \sum_{i=1}^n f_i(t) \right) = 0 \quad \text{für alle } t \in I.$$

Weil  $h$  monoton wachsend ist, gilt  $h'(x) \geq 0$  und  $h(x) \leq h(y)$  für  $x \leq y$ , sodass

$$h'(f_i(t)) \geq 0 \quad \text{und} \quad h(f_1(t)) \leq h(f_2(t)) \leq \dots \leq h(f_n(t))$$

für alle  $t \in I$  gilt. Unter der Annahme, dass  $h$  konvex, also  $h'$  monoton wachsend ist, erhalten wir  $h'(x) \leq h'(y)$  für  $x \leq y$ , woraus

$$h'(f_1(t)) \leq h'(f_2(t)) \leq \dots \leq h'(f_n(t))$$

für alle  $t \in I$  folgt. Wenn nun für reelle Zahlen  $a_1, a_2, b_1, b_2$  die Ungleichungen  $0 < a_1 \leq a_2$  und  $b_1 \leq b_2$  erfüllt sind, dann ist  $a_1 b_1 \leq a_2 b_2$ . Wenden wir dies iteriert an, so erhalten wir

$$h'(f_1(t)) \cdot f'_1(t) \leq h'(f_2(t)) \cdot f'_2(t) \leq \dots \leq h'(f_n(t)) \cdot f'_n(t)$$

für alle  $t \in I$ . Schließlich können wir sehr einfach die Originalaussage durch Anwendung der Tschebyscheff-Summenungleichung einmal mit  $a_i := h'(f_i(t))f'_i(t)$  und  $b_i := e^{h(f_i(t))}$ , ein zweites Mal mit  $a_i := h'(f_i(t))$

### 3. Sum-of-Squared-Logarithms-Ungleichung

und  $b_i := f'_i(t)$  erhalten. Es ist

$$\begin{aligned}
 H'(t) &= \sum_{i=1}^n (h \circ f_i)'(t) e^{h(f_i(t))} = \sum_{i=1}^n h'(f_i(t)) f'_i(t) \cdot e^{h(f_i(t))} \\
 &\stackrel{(2.9.3)}{\geq} \frac{1}{n} \left( \sum_{i=1}^n h'(f_i(t)) \cdot f'_i(t) \right) \left( \sum_{i=1}^n e^{h(f_i(t))} \right) \\
 &\stackrel{(2.9.3)}{\geq} \frac{1}{n^2} \left( \sum_{i=1}^n h'(f_i(t)) \right) \underbrace{\left( \sum_{i=1}^n f'_i(t) \right)}_{=0} \left( \sum_{i=1}^n e^{h(f_i(t))} \right) = 0. \quad \blacksquare
 \end{aligned} \tag{3.4.14}$$

Das folgende Theorem wurde von Birsan, Neff und Lankeit in [9] bewiesen. Mit Satz 3.4.3 (beziehungsweise den Verallgemeinerungen Satz 3.4.5 oder Satz 3.4.6) können wir nun einen sowohl alternativen als auch sehr elementaren Beweis führen:

**Satz 3.4.7.** Seien  $a, b, c, x, y, z \in \mathbb{R}$  mit

$$a \geq b \geq c \quad \text{und} \quad x \geq y \geq z. \tag{3.4.15}$$

Weiterhin

$$a + b + c = x + y + z = 0 \quad \text{und} \quad a^2 + b^2 + c^2 = x^2 + y^2 + z^2. \tag{3.4.16}$$

Dann ist

$$e^a + e^b + e^c \leq e^x + e^y + e^z \tag{3.4.17}$$

genau dann wenn  $a \leq x$ .

**Bemerkung 3.4.8.** Benutzen wir Satz 3.4.5 oder Satz 3.4.6 anstelle von Satz 3.4.3, erlaubt uns der folgende Beweis sogar die stärkere Aussage unter den Bedingungen von Satz 3.4.7 zu zeigen:

$$e^{ma} + e^{mb} + e^{mc} \leq e^{mx} + e^{my} + e^{mz} \quad \text{genau dann wenn} \quad a \leq x \quad \text{für alle } m \in \mathbb{R}_+.$$

*Beweis.* Halten wir zunächst  $a, b, c$  fest, dann definieren wir zur Vereinfachung  $r \in \mathbb{R}_+$  mit  $r^2 := a^2 + b^2 + c^2$ . Mittels der Bedingungen (3.4.15) und (3.4.16) können wir eindeutig  $y$  und  $z$  in Abhängigkeit von  $x \in \mathbb{R}_+$  bestimmen (Mit (3.4.15)  $x < 0$  impliziert  $a + b + c < 0 + y + z \leq 0 + 0 + 0$ ; ein Widerspruch zu (3.4.16)). Wegen  $z = -x - y$  können wir  $y^2 + (-x - y)^2 + x^2 - r^2 = 0$  sicherstellen. Seien  $x$  und  $r$  vorgegeben, dann erhalten wir eine quadratische Gleichung, die wir mit der  $p$ - $q$ -Formel lösen können:

$$\begin{aligned}
 y^2 + (-x - y)^2 + x^2 - r^2 &= 2y^2 + 2xy + 2x^2 - r^2 = 0 \quad \Leftrightarrow \quad y^2 + xy + x^2 - \frac{1}{2}r^2 = 0 \\
 \Leftrightarrow y &= -\frac{1}{2}x \pm \sqrt{\frac{1}{4}x^2 - x^2 + \frac{1}{2}r^2} = -\frac{1}{2}x \pm \sqrt{-\frac{3}{4}x^2 + \frac{1}{2}r^2}.
 \end{aligned}$$

Setzen wir diese beiden Lösungen in  $z = -x - y$  ein, so bekommen wir

$$z = -\frac{1}{2}x \mp \sqrt{-\frac{3}{4}x^2 + \frac{1}{2}r^2}.$$

Von den beiden Möglichkeiten  $\pm$  für  $y$  bleibt nur der positive Fall, bei  $\mp$  für  $z$  nur der negative Fall übrig, um (3.4.15) zu erfüllen. Beide Gleichungen haben drei reelle Lösungen genau dann, wenn  $-\frac{3}{4}x^2 + \frac{1}{2}r^2 \geq 0$  – dafür muss  $x \leq \sqrt{\frac{2}{3}r^2}$  gelten. Außerdem

$$\begin{aligned}
 x \geq y &\Leftrightarrow x \geq -\frac{1}{2}x + \sqrt{-\frac{3}{4}x^2 + \frac{1}{2}r^2} \Leftrightarrow \frac{3}{2}x \geq \sqrt{-\frac{3}{4}x^2 + \frac{1}{2}r^2} \\
 &\Leftrightarrow \frac{9}{4}x^2 \geq -\frac{3}{4}x^2 + \frac{1}{2}r^2 \Leftrightarrow 3x^2 \geq \frac{1}{2}r^2 \Leftrightarrow x \geq \sqrt{\frac{1}{6}r^2}.
 \end{aligned}$$

### 3. Sum-of-Squared-Logarithms-Ungleichung

Mit  $D_r := \left[ \sqrt{\frac{1}{6}r^2}, \sqrt{\frac{2}{3}r^2} \right]$  erhalten wir die differenzierbaren Funktionen  $y, z: D_r \rightarrow \mathbb{R}$  mit

$$y(x) = -\frac{1}{2}x + \sqrt{-\frac{3}{4}x^2 + \frac{1}{2}r^2} \quad \text{und} \quad z(x) = -\frac{1}{2}x - \sqrt{-\frac{3}{4}x^2 + \frac{1}{2}r^2} \quad (3.4.18)$$

als die eindeutigen Lösungen  $(x, y(x), z(x))$  von (3.4.15) und (3.4.16). Da die vorgegebenen  $a, b, c$  diese Bedingungen erfüllen, sind offensichtlich  $a \in D_r$ ,  $b = y(a)$  und  $c = y(b)$ .

Nun definieren wir für  $m \in \mathbb{R}_+$  die Funktion  $\tilde{g}: D_r \rightarrow \mathbb{R}$  mit  $\tilde{g}(x) = e^{m y(x)} + e^{m z(x)}$ . Wir wissen  $y(x) \geq z(x)$  und

$$y'(x) = \frac{1}{2} - \frac{\frac{3}{4}x}{\sqrt{-\frac{3}{4}x^2 + \frac{1}{2}r^2}} \leq \frac{1}{2} + \frac{\frac{3}{4}x}{\sqrt{-\frac{3}{4}x^2 + \frac{1}{2}r^2}} = z'(x).$$

Also können wir mit Satz 3.4.5 oder Satz 3.4.6 folgern (in beiden Fällen können wir  $h(t) := mt$  setzen. Für  $m = 1$  können wir direkt Satz 3.4.3 benutzen), dass  $\tilde{g}$  monoton wachsend ist. Zudem ist  $x \mapsto e^{m x}$  monoton wachsend, also ist die Summenbildung  $g: \mathbb{R} \rightarrow \mathbb{R}$  mit  $g(x) = e^{m x} + e^{m y(x)} + e^{m z(x)}$  ebenso monoton wachsend. Deshalb

$$g(x) \geq g(a) \quad \text{genau dann, wenn} \quad x \geq a,$$

was äquivalent zur Aussage ist, wenn wir  $m = 1$  setzen. ■

## 3.5. Neue logarithmische Ungleichungen in der Informationstheorie

Zunächst stellen wir die Jensen-Ungleichung (vgl. Mitrinović, Pečarić, [50, eq. (2.1) S. 191]) und die Log-Summenungleichung (vgl. Cover, Thomas [19, 2.7 S. 29]) vor. Danach beweisen wir die Informationsungleichung, welche in abweichenden Schreibweisen unter dem Namen Gibbs-Ungleichung bekannt ist. Die Informationsungleichung (vgl. Cover, Thomas [19, 2.6 S. 28]) ist die grundlegendste Ungleichung in der Informationstheorie. Es wird behauptet, dass die relative Entropie zwischen zwei Wahrscheinlichkeitsverteilungen  $p, q: \Omega \rightarrow [0, 1]$ , definiert als

$$D(p \parallel q) := \sum_{x \in \Omega} p(x) \log \frac{p(x)}{q(x)} \quad (3.5.1)$$

nichtnegativ ist. Für den Fall, dass  $p$  und  $q$  Wahrscheinlichkeitsmaße auf einer endlichen Menge  $\Omega = \{1, \dots, n\}$  sind, ergibt sich die Formel zu:

$$D(p \parallel q) := \sum_{i=1}^n p_i \log \frac{p_i}{q_i}. \quad (3.5.2)$$

**Lemma 3.5.1** (Jensen-Ungleichung). *Sei  $I \subseteq \mathbb{R}$  ein Intervall,  $f: I \rightarrow \mathbb{R}$  eine konvexe Funktion,  $\lambda_1, \dots, \lambda_n$  positive Zahlen mit  $\lambda_1 + \dots + \lambda_n = 1$  und  $x_1, \dots, x_n \in I$ . Dann gilt*

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i). \quad (3.5.3)$$

*Wir bemerken: Mit  $n = 2$  erhalten wir direkt die Definition der Konvexität von  $f$ .*

*Beweis.* Die Ungleichung folgt direkt aus der Definition der Konvexität durch vollständige Induktion: Der Fall  $n = 1$  ist trivial,  $n = 2$  ist wahr, weil  $f$  konvex ist.



### 3. Sum-of-Squared-Logarithms-Ungleichung

Sei nun als Induktionsvoraussetzung (IV) die Gleichung (3.5.3) erfüllt für ein  $n \in \mathbb{N}$  und für  $n = 2$ , dann folgt

$$\begin{aligned} f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) &= f\left(\sum_{i=1}^{n-1} \lambda_i x_i + (\lambda_n + \lambda_{n+1}) \frac{\lambda_n x_n + \lambda_{n+1} x_{n+1}}{\lambda_n + \lambda_{n+1}}\right) \\ &\stackrel{\text{IV } n}{\leq} \sum_{i=1}^{n-1} \lambda_i f(x_i) + (\lambda_n + \lambda_{n+1}) f\left(\frac{\lambda_n}{\lambda_n + \lambda_{n+1}} x_n + \frac{\lambda_{n+1}}{\lambda_n + \lambda_{n+1}} x_{n+1}\right) \\ &\stackrel{\text{IV } 2}{\leq} \sum_{i=1}^{n-1} \lambda_i f(x_i) + \lambda_n f(x_n) + \lambda_{n+1} f(x_{n+1}) = \sum_{i=1}^{n+1} \lambda_i f(x_i), \end{aligned}$$

somit ist (3.5.3) erfüllt für  $n + 1$ . ■

Mit der Jensen-Ungleichung können wir die sogenannte Log-Summenungleichung beweisen:

**Lemma 3.5.2** (strengere Log-Summenungleichung). *Seien  $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}_+$  und  $k \geq 0$ , dann ist*

$$\sum_{i=1}^n a_i \log\left(\frac{a_i}{b_i} + k\right) \geq \left(\sum_{i=1}^n a_i\right) \log\left(\frac{1}{\sum_{i=1}^n b_i} \sum_{i=1}^n a_i + k\right). \quad (3.5.4)$$

*Wir erhalten genau dann Gleichheit, wenn  $\frac{a_1}{b_1} = \frac{a_2}{b_2} = \dots = \frac{a_n}{b_n}$ .*

**Bemerkung 3.5.3.** Für  $k = 0$  haben wir die Log-Summenungleichung:

$$\sum_{i=1}^n a_i \log \frac{a_i}{b_i} \geq \left(\sum_{i=1}^n a_i\right) \log\left(\frac{1}{\sum_{i=1}^n b_i} \sum_{i=1}^n a_i\right). \quad (3.5.5)$$

*Beweis.* Mit  $f(x) = x \log(x + k)$ ,  $\lambda_i := b_i / \sum_{i=1}^n b_i$ ,  $x_i := a_i / b_i$  und (3.5.3) erhalten wir

$$\begin{aligned} \sum_{i=1}^n a_i \log\left(\frac{a_i}{b_i} + k\right) &= \left(\sum_{i=1}^n b_i\right) \frac{\sum_{i=1}^n b_i \frac{a_i}{b_i} \log\left(\frac{a_i}{b_i} + k\right)}{\sum_{i=1}^n b_i} = \sum_{i=1}^n \lambda_i x_i \log(x_i + k) \\ &= \sum_{i=1}^n \lambda_i f(x_i) \stackrel{(3.5.3)}{\geq} f\left(\sum_{i=1}^n \lambda_i x_i\right) = \left(\sum_{i=1}^n b_i\right) \left(\sum_{i=1}^n \lambda_i x_i\right) \log\left(\sum_{i=1}^n \lambda_i x_i + k\right) \\ &= \left(\sum_{i=1}^n b_i\right) \left(\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}\right) \log\left(\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} + k\right) = \left(\sum_{i=1}^n a_i\right) \log\left(\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} + k\right). \end{aligned} \quad (3.5.6)$$

Wenn nun  $c := \frac{a_1}{b_1} = \frac{a_2}{b_2} = \dots = \frac{a_n}{b_n}$ , dann ist  $x_i = c$  und so erhalten wir

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) = f\left(\sum_{i=1}^n \lambda_i c\right) = f(c) = \sum_{i=1}^n \lambda_i f(c) = \sum_{i=1}^n \lambda_i f(x_i).$$

Gibt es ein  $i \in \{1, \dots, n-1\}$  mit  $\frac{a_i}{b_i} \neq \frac{a_{i+1}}{b_{i+1}}$ , dann erhalten wir

$$\sum_{i=1}^n \lambda_i f(x_i) > f\left(\sum_{i=1}^n \lambda_i x_i\right). \quad \blacksquare$$

**Proposition 3.5.4** (Gibbs-Ungleichung / Informationsungleichung). *Bezeichne  $P_n$  die Menge der Wahrscheinlichkeitsmaße auf einer  $n$ -elementigen Menge, also  $P_n = \{p \in \mathbb{R}_+^n \mid \sum_{i=1}^n p_i = 1\}$ . Die folgenden vier Ausdrücke (die ersten drei werden Gibbs-Ungleichung, die vierte Informationsungleichung genannt) sind äquivalent und erfüllt für alle  $a, b \in P_n$ :*

$$\begin{aligned} i) \quad \sup_{\xi \in P_n} \left\{ \prod_{i=1}^n \xi_i^{a_i} \right\} &= \prod_{i=1}^n a_i^{a_i}, & ii) \quad \inf_{\xi \in P_n} \left\{ \sum_{i=1}^n a_i (-\log \xi_i) \right\} &= \sum_{i=1}^n a_i (-\log a_i), \\ iii) \quad \sum_{i=1}^n a_i (-\log b_i) &\geq \sum_{i=1}^n a_i (-\log a_i), & iv) \quad D(a \parallel b) &= \sum_{i=1}^n a_i \log \frac{a_i}{b_i} \geq 0. \end{aligned} \quad (3.5.7)$$

### 3. Sum-of-Squared-Logarithms-Ungleichung

*Beweis.* Zunächst beweisen wir die Äquivalenz der vier Ausdrücke:

$$\begin{aligned}
 & \sup_{\xi \in P_n} \left\{ \prod_{i=1}^n b_i^{a_i} \right\} = \prod_{i=1}^n a_i^{a_i} \Leftrightarrow \inf_{\xi \in P_n} \left\{ -\prod_{i=1}^n b_i^{a_i} \right\} = -\prod_{i=1}^n a_i^{a_i} \Leftrightarrow -\prod_{i=1}^n b_i^{a_i} \geq -\prod_{i=1}^n a_i^{a_i} \\
 \Leftrightarrow & -\log \prod_{i=1}^n b_i^{a_i} \geq -\log \prod_{i=1}^n a_i^{a_i} \Leftrightarrow -\sum_{i=1}^n \log b_i^{a_i} \geq -\sum_{i=1}^n \log a_i^{a_i} \\
 \Leftrightarrow & \sum_{i=1}^n a_i (-\log b_i) \geq \sum_{i=1}^n a_i (-\log a_i) \Leftrightarrow \sum_{i=1}^n a_i (\log a_i - \log b_i) \geq 0 \\
 \Leftrightarrow & \sum_{i=1}^n a_i \log \frac{a_i}{b_i} \geq 0 \Leftrightarrow D(a \parallel b) \geq 0.
 \end{aligned}$$

Nun sei  $a, b \in P_n$ . Mit der strengeren Log-Summenungleichung und wegen  $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i = 1$  erhalten wir

$$\sum_{i=1}^n a_i \log \left( \frac{a_i}{b_i} + k \right) \geq \left( \sum_{i=1}^n a_i \right) \log \left( \frac{1}{\sum_{i=1}^n b_i} \sum_{i=1}^n a_i + k \right) = \log(1+k). \quad (3.5.8)$$

Mit  $k = 0$  erhalten wir die Informationsungleichung aus Fall iv). ■

**Bemerkung 3.5.5.** Analog können wir die Ungleichung (3.5.8) auch als die *strengere Informationsungleichung* nennen.

**Folgerung 3.5.6** (verallgemeinerte Log-Summenungleichung). Für  $a, b \in \mathbb{R}_+^n$  und  $k_1, \dots, k_n \in [0, \infty)$  gilt

$$\sum_{i=1}^n a_i \log \prod_{s=1}^m \left( \frac{a_i}{b_i} + k_s \right) \geq \left( \sum_{i=1}^n a_i \right) \log \prod_{s=1}^m \left( \frac{1}{\sum_{i=1}^n b_i} \sum_{i=1}^n a_i + k_s \right). \quad (3.5.9)$$

*Beweis.* Mit  $m$  Zahlen  $k_1, \dots, k_m \in [0, 1)$  erhalten wir durch  $m$ -fache Summenbildung über die strengere Log-Summenungleichung (3.5.4)

$$\sum_{s=1}^m \left( \sum_{i=1}^n a_i \log \left( \frac{a_i}{b_i} + k_s \right) \right) \geq \sum_{s=1}^m \left( \sum_{i=1}^n a_i \log \left( \frac{1}{\sum_{i=1}^n b_i} \sum_{i=1}^n a_i + k_s \right) \right).$$

Unter Verwendung der Distributivität und den Logarithmengesetzen erhalten wir direkt unser gewünschtes Resultat. ■

**Folgerung 3.5.7** (verallgemeinerte Informationsungleichung). Für  $a, b \in P_n$  und  $k_1, \dots, k_n \in [0, \infty)$  gilt

$$\sum_{i=1}^n a_i \log \prod_{s=1}^m \left( \frac{a_i}{b_i} + k_s \right) \geq \log \prod_{s=1}^m (1 + k_s). \quad (3.5.10)$$

*Beweis.* Wir vereinfachen (3.5.9) unter der Bedingung  $a, b \in P_n$ , somit gilt  $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i = 1$ . ■

**Teil II.**

**Anwendung auf die nichtlineare  
Elastizitätstheorie**

# 4. Einführung in die nichtlineare Elastizitätstheorie

## 4.1. Modellierung über das Kraftgleichgewicht nach Cauchy

Gegeben sei ein physischer Körper elastischen Materials, der Kräften ausgesetzt wird, die ihn verformen. Dies sind zum einen auf seine Oberfläche einwirkende Kräfte wie Ziehen oder Drücken, zum anderen auf sein Volumen wirkende Kräfte wie Gravitation oder Elektromagnetismus. Wir gehen davon aus, dass die Verformung in einem stationären Zustand endet. Die Elastizität des Materials bewirkt eine elastische Verformung. Diese ist nur abhängig vom Ausgangszustand und den einwirkenden Kräften, nicht aber von vorigen Verformungen und hat die Eigenschaft, dass der Körper sich unter Wegfall der verformenden Kräfte wieder zurückverformt. Zudem gehen wir davon aus, dass kein Teil des Materials unendlich dicht zusammengedrückt oder unendlich stark auseinander gezogen werden kann. Damit der Körper in einem stationären Zustand enden kann, müssen zu den von außen einwirkenden Kräften auch innere Kräfte entstehen, welche die äußeren ausgleichen.

Zur mathematischen Modellierung des stationären Zustandes der Verformung definieren wir:

**Definition 4.1.1.** Gegeben sei ein Gebiet  $\Omega \subseteq \mathbb{R}^3$  mit hinreichend glattem Rand und eine hinreichend glatte, sich nicht selbstdurchdringende<sup>1</sup> Deformationsabbildung  $\varphi: \Omega \rightarrow \mathbb{R}^3$  mit  $x \mapsto x^\varphi := \varphi(x)$  und  $\nabla\varphi(x) \in GL^+(3)$ , wobei wir das Bild von  $\varphi$  unter  $\Omega$  als  $\Omega^\varphi$  bezeichnen.

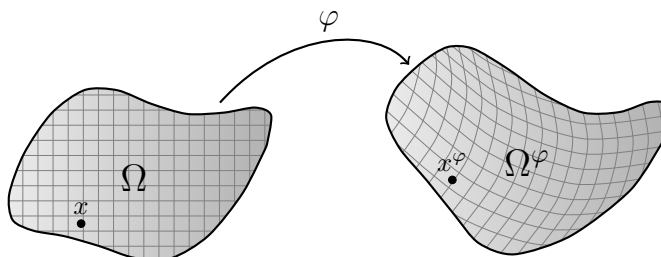


Abbildung 4.1.: Deformation eines Gebietes  $\Omega$

Zudem sei eine *Volumenkraftabbildung*  $f^\varphi: \Omega^\varphi \rightarrow \mathbb{R}^3$  gegeben. Zu einer *Flächenkraftabbildung*  $p^\varphi: \Omega^\varphi \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$  sind ausschließlich die Funktionswerte für einen Teil des Randes  $\Gamma \subseteq \partial\Omega$  vorgegeben, der Rest ist nicht bekannt. Ziel ist es, zumindest Eigenschaften der Flächenkraftabbildung angeben zu können.

**Bemerkung 4.1.2.** Ein hochgestelltes  $\varphi$  drückt immer die Abhängigkeit der Menge oder Abbildung von  $\varphi$  aus.

Die Volumenkraften sind zu verstehen als Kraftdichten, somit Kraft pro Volumen, und sind abhängig vom Ort im verformten Gebiet  $x^\varphi \in \Omega^\varphi$ . Volumenkraften können durch beispielsweise Gravitation oder Elektromagnetismus induziert werden. Hierbei wirkt dann ein Beschleunigungsfeld auf den Körper und die Kraftdichte an einem Ort ergibt sich aus der Beschleunigung und der Dichte des Körpers an diesem Ort.

Die Flächenkräfte sind ebenfalls als Kraftdichten zu verstehen. Sie sind abhängig vom Ort im verformten Gebiet  $x^\varphi \in \Omega^\varphi$  und dem Normalenvektor  $n^\varphi \in \mathbb{R}^3$  der durch den Punkt  $x^\varphi$  gehenden zu betrachtenden

<sup>1</sup>Eine Selbstdurchdringung setzt einen Punkt der unendlichen Kompression voraus, welchen wir ausschließen.

#### 4. Einführung in die nichtlineare Elastizitätstheorie

Fläche. Betrachten wir zwei durch den verformten Körper gehende Flächen, die im Punkt  $x^\varphi$  dieselbe Normale haben, ist die Flächenkraftdichte auch dieselbe. Flächenkräfte sind die von außen wirkenden Kräfte, die zur Verformung beitragen, wie auch die inneren Rückstellkräfte des gezogenen oder gestauchten Materials.

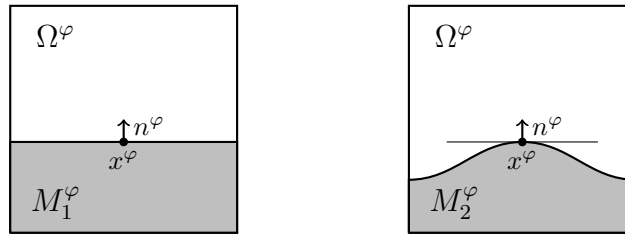


Abbildung 4.2.: zwei Teilgebiete  $M_1^\varphi$  und  $M_2^\varphi$  von  $\Omega^\varphi$  mit gleichem Normalenvektor  $n^\varphi$  im Punkt  $x^\varphi$

Es ist klar, dass sich die Volumenkräfte und Flächenkräfte in der Summe für jedes Teilgebiet  $M^\varphi \subseteq \Omega^\varphi$  ausgleichen müssen, damit der Körper im ruhenden Gleichgewicht ist. Ansonsten würde eine Kraft auf  $M^\varphi$  im Ganzen wirken und dieses bewegen. Damit keine Verschiebung des Körpers oder seiner Teile stattfindet, jeder Teilkörper eine ausgeglichene Kraftbilanz aufweist, muss gelten:

$$\underbrace{\int_{\partial M} p^\varphi(x^\varphi, n^\varphi) dx^\varphi}_{\text{Flächenkräfte}} + \underbrace{\int_M f^\varphi(x^\varphi) dx^\varphi}_{\text{Volumenkräfte}} = 0 \quad \text{für alle } M^\varphi \subseteq \Omega^\varphi. \quad (4.1.1)$$

Damit keine Rotation des Körpers oder seiner Teile stattfindet, jeder Teilkörper eine ausgeglichene Drehmomentbilanz aufweist, muss gelten:

$$\underbrace{\int_{\partial M^\varphi} x^\varphi \times p^\varphi(x^\varphi, n^\varphi) dx^\varphi}_{\text{Flächendrehmomente}} + \underbrace{\int_{M^\varphi} x^\varphi \times f^\varphi(x^\varphi) dx^\varphi}_{\text{Volumendrehmomente}} = 0 \quad \text{für alle } M^\varphi \subseteq \Omega^\varphi. \quad (4.1.2)$$

Hierbei ist  $n^\varphi := n(x^\varphi, \partial M) \in S^1$  stets der äußere Normalenvektor auf  $\partial M^\varphi$  im Punkt  $x^\varphi$ .

**Beispiel 4.1.3.** Betrachten wir als Beispiel einen Zylinder  $Z$ , der durch gleichmäßigen Druck auf seine beiden Stirnflächen auf einen Zylinder von gleicher Grundfläche  $Z^\varphi$  zusammengedrückt wurde. Volumenkräfte sollen keine wirken. Wirkt auf die obere Stirnfläche  $A_1$  die konstante Kraft  $p_1$  und auf die untere Stirnfläche  $A_2$  die konstante Kraft  $p_2$ , dann ergibt sich

$$\int_{A_1} p_1 dx^\varphi + \int_{A_2} p_2 dx^\varphi = 0 \quad (4.1.3)$$

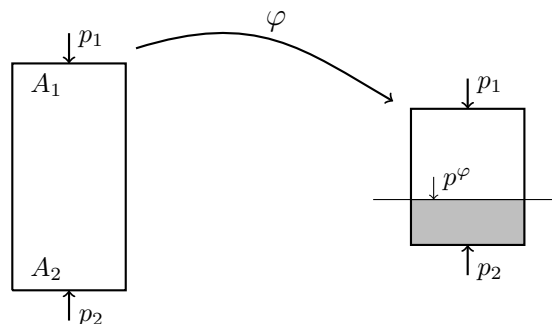


Abbildung 4.3.: Zylinder vor und nach der Verformung in Seitenansicht

#### 4. Einführung in die nichtlineare Elastizitätstheorie

**Axiom 4.1.4** (Spannungsprinzip von Euler und Cauchy). Sei  $\Omega^\varphi \subseteq \mathbb{R}^n$  ein Gebiet mit hinreichend glattem Rand und  $\Gamma^\varphi \subseteq \partial\Omega^\varphi$  ein Teil des Randes. Zu einer vorgegebenen auf dem Abschluss  $\overline{\Omega}^\varphi$  definierten Volumenkraftabbildung  $f^\varphi: \overline{\Omega}^\varphi \rightarrow \mathbb{R}^3$  und einer auf dem Randteil  $\Gamma^\varphi$  definierten Vorgabe der Flächenkräfte  $g^\varphi: \Gamma^\varphi \rightarrow \mathbb{R}^3$  existiert eine Flächenkraftabbildung  $p^\varphi: \overline{\Omega}^\varphi \times S^1 \rightarrow \mathbb{R}^3$ , sodass  $p^\varphi$  auf dem Randteil  $\Gamma^\varphi$  durch

$$p^\varphi(x^\varphi, n^\varphi) = g^\varphi(x^\varphi) \quad \text{für alle } x^\varphi \in \Gamma^\varphi$$

festgelegt ist und zudem für alle Teilgebiete  $M^\varphi$  von  $\Omega^\varphi$  das Gleichgewicht der Kräfte (4.1.1) und das Gleichgewicht der Drehmomente (4.1.2) erfüllt ist. Hierbei ist  $n^\varphi := n(x^\varphi, \partial M) \in S^1 := \{v \in \mathbb{R}^3 \mid \|v\| = 1\}$  der äußere Normalenvektor auf  $\partial M$  im Punkt  $x^\varphi$ .

Dies lässt sich auch als Anfangswertproblem schreiben: Zu wie oben vorgegebenen  $f^\varphi$  und  $g^\varphi$  existiert  $p^\varphi$ , sodass

$$\left. \begin{aligned} \int_{M^\varphi} f^\varphi(x^\varphi) dx^\varphi + \int_{\partial M^\varphi} p^\varphi(x^\varphi, n^\varphi) dS(x^\varphi) &= 0 \\ \int_{M^\varphi} x^\varphi \times f^\varphi(x^\varphi) dx^\varphi + \int_{\partial M^\varphi} x^\varphi \times p^\varphi(x^\varphi, n^\varphi) dS(x^\varphi) &= 0 \\ p^\varphi(x^\varphi, n^\varphi) &= g^\varphi(x^\varphi) \quad \forall x^\varphi \in (M^\varphi \cap \Gamma^\varphi) \end{aligned} \right\} \quad \text{für alle Teilgebiete } M^\varphi \subseteq \Omega^\varphi \quad (4.1.4)$$

gilt. Hierbei ist  $n^\varphi := n^\varphi(x^\varphi) \in S^1$  der äußere Normalenvektor des Randes  $\partial M$  in  $x^\varphi$ .

**Lemma 4.1.5.** Für jedes  $x^\varphi \in \overline{\Omega}^\varphi$  ist die Flächenkraftabbildung  $p^\varphi$  linear in  $n^\varphi \in S^1 := \{v \in \mathbb{R}^3 \mid \|v\| = 1\}$ , d.h. es gibt eine Abbildung  $\sigma: \overline{\Omega}^\varphi \rightarrow \mathbb{R}^{3 \times 3}$ , sodass  $p^\varphi(x^\varphi, n^\varphi) = \sigma(x^\varphi) \cdot n^\varphi$  für alle  $x^\varphi \in \overline{\Omega}^\varphi$  und alle  $n^\varphi \in S^1$  gilt.

*Beweis.* Sei  $x^\varphi \in \Omega^\varphi$  und  $n^\varphi = (a, b, c)^T \in S^1$ , d.h.  $a^2 + b^2 + c^2 = 1$ . Wir wählen  $x, y, z \in \mathbb{R}$  so, dass  $\frac{1}{2}yz = a$ ,  $\frac{1}{2}zx = b$ ,  $\frac{1}{2}xy = c$  gilt und betrachten den Tetraeder mit den Endpunkten

$$x_0^\varphi, \quad x_0^\varphi + \varepsilon x \cdot e_1, \quad x_0^\varphi + \varepsilon y \cdot e_2 \quad \text{und} \quad x_0^\varphi + \varepsilon x \cdot e_3,$$

der für hinreichend kleines  $\varepsilon$  ganz in  $\Omega^\varphi$  liegt.

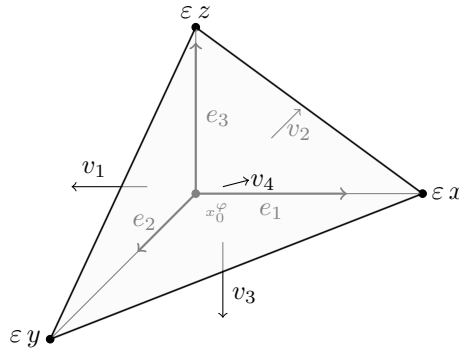


Abbildung 4.4.: Tetraeder an  $x_0^\varphi$  mit Seitenkantenlängen  $\varepsilon x$ ,  $\varepsilon y$ ,  $\varepsilon z$

Wir charakterisieren die vier Seitenflächen des Tetraeders  $A_1, A_2, A_3, A_4$  durch die auf den Seitenflächen senkrecht stehenden Normalenvektoren  $-e_1, -e_2, -e_3, n^\varphi$  multipliziert mit dem jeweiligen Flächeninhalt:

$$\begin{aligned} v_1 &= \frac{1}{2} (\varepsilon z e_3) \times (\varepsilon y e_2) = -\varepsilon^2 a e_1, \\ v_2 &= \frac{1}{2} (\varepsilon x e_1) \times (\varepsilon z e_3) = -\varepsilon^2 b e_2, \\ v_3 &= \frac{1}{2} (\varepsilon y e_2) \times (\varepsilon x e_1) = -\varepsilon^2 c e_3, \\ v_4 &= \frac{1}{2} (\varepsilon y e_2 - \varepsilon x e_1) \times (\varepsilon z e_3 - \varepsilon x e_1) = \varepsilon^2 (a e_1 + b e_2 + c e_3) = \varepsilon^2 n^\varphi. \end{aligned}$$

#### 4. Einführung in die nichtlineare Elastizitätstheorie

Die zugehörigen Flächenkräfte sind

$$\begin{aligned} p_1 &= \int_{A_1} p^\varphi(x_0^\varphi, -e_1) dS(x^\varphi), & p_2 &= \int_{A_2} p^\varphi(x_0^\varphi, -e_2) dS(x^\varphi), \\ p_3 &= \int_{A_3} p^\varphi(x_0^\varphi, -e_3) dS(x^\varphi), & p_4 &= \int_{A_4} p^\varphi(x_0^\varphi, n^\varphi) dS(x^\varphi). \end{aligned}$$

Für hinreichend kleines  $\varepsilon$  können wir die Kräfte hinreichend genau approximieren durch

$$p_1 \sim \varepsilon^2 a \cdot p^\varphi(x_0^\varphi, -e_1), \quad p_2 \sim \varepsilon^2 b \cdot p^\varphi(x_0^\varphi, -e_2), \quad p_3 \sim \varepsilon^2 c \cdot p^\varphi(x_0^\varphi, -e_3), \quad p_4 \sim \varepsilon^2 \cdot p^\varphi(x_0^\varphi, n^\varphi).$$

Mit dem aus Symmetriegründen für alle  $n \in S^1$  geltenden  $p^\varphi(x_0^\varphi, -n) = -p^\varphi(x_0^\varphi, n)$  können wir auch schreiben

$$p_1 \sim -\varepsilon^2 a \cdot p^\varphi(x_0^\varphi, e_1), \quad p_2 \sim -\varepsilon^2 b \cdot p^\varphi(x_0^\varphi, e_2), \quad p_3 \sim -\varepsilon^2 c \cdot p^\varphi(x_0^\varphi, e_3), \quad p_4 \sim \varepsilon^2 \cdot p^\varphi(x_0^\varphi, n^\varphi).$$

Aus der Formel für das Kraftgleichgewicht (4.1.4) mit dem Umstand, dass die Volumenkräfte  $f^\varphi$  kubisch für  $\varepsilon \rightarrow 0$  kubisch gegen Null streben, die Flächenkräfte hingegen nur quadratisch, die Volumenkräfte im Grenzübergang somit vernachlässigbar sind, folgt die Forderung  $p_1 + p_2 + p_3 + p_4 = 0$ , die uns, nochmal durch  $\varepsilon^2$  geteilt,

$$p^\varphi(x_0^\varphi, n^\varphi) = p^\varphi(x_0^\varphi, a e_1 + b e_2 + c e_3) = a \cdot p^\varphi(x_0^\varphi, e_1) + b \cdot p^\varphi(x_0^\varphi, e_2) + c \cdot p^\varphi(x_0^\varphi, e_3)$$

sicherstellt. ■

**Satz 4.1.6** (Satz von Cauchy). *Es gelte das wie im vorigen beschriebene Spannungsprinzip von Cauchy und Euler. Zudem gehen wir von Stetigkeit von  $f^\varphi$  in  $x^\varphi$  und  $t^\varphi$  in  $n^\varphi$  und von stetiger Differenzierbarkeit von  $t^\varphi$  in  $x^\varphi$  aus. Dann existiert eine Abbildung  $\sigma^\varphi: \overline{\Omega}^\varphi \rightarrow \text{Sym}(3)$ , für die*

$$p^\varphi(x^\varphi, n^\varphi) = \sigma^\varphi(x^\varphi) \cdot n^\varphi \quad \text{für alle } x^\varphi \in \overline{\Omega}^\varphi \text{ und alle } n^\varphi \in S^1 \quad (4.1.5)$$

gilt und die

$$\text{div } \sigma^\varphi(x^\varphi) = -f^\varphi(x^\varphi) \quad \text{für alle } x^\varphi \in \overline{\Omega}^\varphi \quad (4.1.6)$$

und

$$\sigma^\varphi(x^\varphi) \cdot n^\varphi = g^\varphi(x^\varphi) \quad \text{für alle } x^\varphi \in \Gamma^\varphi \quad (4.1.7)$$

erfüllt.

**Bemerkung 4.1.7.** Hierbei entspricht  $\text{div } \sigma^\varphi = -f$  dem Gleichgewicht der Kräfte. Das Gleichgewicht der Drehmomente schlägt sich wiederum in  $\sigma^\varphi(x^\varphi) \in \text{Sym}(3)$  nieder.

*Beweis.* Die Linearität von  $p^\varphi$  haben wir in Lemma 4.1.5 bereits gezeigt.

Unter Anwendung des Gaußschen Integralsatzes gilt für alle Teilgebiete  $M^\varphi \subset \overline{\Omega}^\varphi$  das Gleichgewicht der Kräfte als

$$0 = \int_{M^\varphi} f^\varphi(x^\varphi) dx^\varphi + \int_{\partial M^\varphi} \sigma^\varphi(x^\varphi) \cdot n^\varphi dS(x^\varphi) = \int_{M^\varphi} f^\varphi(x^\varphi) + \text{div } \sigma^\varphi(x^\varphi) dx^\varphi,$$

woraus  $f^\varphi(x^\varphi) = -\text{div } \sigma^\varphi(x^\varphi)$  für alle  $x^\varphi \in \overline{\Omega}^\varphi$  folgt.

Wir schreiben  $\sigma_{i,\cdot}^\varphi(x^\varphi)$  für die Zeilen von  $\sigma^\varphi(x^\varphi)$ , dann ist

$$p^\varphi(x^\varphi, n^\varphi) = \sigma^\varphi(x^\varphi) \cdot n^\varphi = \begin{pmatrix} \langle \sigma_{1,\cdot}^\varphi, n^\varphi \rangle \\ \langle \sigma_{2,\cdot}^\varphi, n^\varphi \rangle \\ \langle \sigma_{3,\cdot}^\varphi, n^\varphi \rangle \end{pmatrix}.$$

Zusammen mit  $f^\varphi = -\text{div } \sigma^\varphi$  erhalten wir dann

$$f_i^\varphi(x^\varphi) = \text{div } \sigma_{i,\cdot}^\varphi(x^\varphi) \quad \text{und} \quad p_i^\varphi(x^\varphi, n^\varphi) = \langle \sigma_{i,\cdot}^\varphi, n^\varphi \rangle \quad \text{für } i \in \{1, 2, 3\}.$$

#### 4. Einführung in die nichtlineare Elastizitätstheorie

Für  $a, b \in \mathbb{R}^3$  ist  $a \times b = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$ , so lässt sich das Gleichgewicht der Drehmomente für alle Teilgebiete  $M^\varphi \subset \bar{\Omega}^\varphi$  schreiben als

$$\begin{aligned} 0 &= \int_{M^\varphi} x_i^\varphi f_j^\varphi(x^\varphi) - x_j^\varphi f_i^\varphi(x^\varphi) dx^\varphi + \int_{\partial M^\varphi} \langle x_i^\varphi \sigma_{j,\cdot}^\varphi(x^\varphi), n^\varphi \rangle - \langle x_j^\varphi \sigma_{i,\cdot}^\varphi(x^\varphi), n^\varphi \rangle dS(x^\varphi) \\ &= \int_{M^\varphi} x_i^\varphi f_j^\varphi(x^\varphi) - x_j^\varphi f_i^\varphi(x^\varphi) dx^\varphi + \int_{\partial M^\varphi} \sum_{k=1}^3 (x_i \sigma_{j,\cdot}^\varphi(x^\varphi) - x_j \sigma_{i,\cdot}^\varphi(x^\varphi)) \cdot n_k^\varphi dS(x^\varphi) \end{aligned} \quad (4.1.8)$$

für alle  $(i, j) \in \{(2, 3), (3, 1), (1, 2)\}$ . Wir erinnern uns an die Greensche Formel (1.2.15) auf Seite 11:  $\int_{\partial A} u_i(x) n_i(x) dS(x) = \int_A \frac{\partial u}{\partial x_i}(x) dx$  für  $u \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ . Mit ihr ist

$$\begin{aligned} &\int_{\partial M^\varphi} \sum_{k=1}^3 (x_i \sigma_{jk}^\varphi(x^\varphi) - x_j \sigma_{ik}^\varphi(x^\varphi)) \cdot n_k^\varphi dS(x^\varphi) \\ &= \int_{M^\varphi} \sum_{k=1}^3 \frac{d}{dx_k} (x_i \sigma_{jk}^\varphi(x^\varphi) - x_j \sigma_{ik}^\varphi(x^\varphi)) dx^\varphi \\ &= \int_{M^\varphi} \sum_{k=1}^3 \left( \delta_{ik} \sigma_{jk}^\varphi(x^\varphi) - \delta_{jk} \sigma_{ik}^\varphi(x^\varphi) + x_i \frac{d}{dx_k} \sigma_{jk}^\varphi(x^\varphi) - x_j \frac{d}{dx_k} \sigma_{ik}^\varphi(x^\varphi) \right) dx^\varphi \\ &= \int_{M^\varphi} \sigma_{ij}^\varphi(x^\varphi) - \sigma_{ji}^\varphi(x^\varphi) dx^\varphi + \int_{M^\varphi} \underbrace{x_i \operatorname{div} \sigma_{j,\cdot}^\varphi(x^\varphi)}_{-f_j(x^\varphi)} - \underbrace{x_j \operatorname{div} \sigma_{i,\cdot}^\varphi(x^\varphi)}_{-f_i(x^\varphi)} dx^\varphi. \end{aligned}$$

Einsetzen in (4.1.8) ergibt

$$0 = \int_{M^\varphi} \sigma_{ij}^\varphi(x^\varphi) - \sigma_{ji}^\varphi(x^\varphi) dx^\varphi \quad \text{für alle Teilgebiete } M^\varphi \subset \bar{\Omega}^\varphi \text{ und alle } (i, j) \in \{(2, 3), (3, 1), (1, 2)\}$$

und damit  $0 = \sigma_{ij}^\varphi(x^\varphi) - \sigma_{ji}^\varphi(x^\varphi)$  für alle  $x^\varphi \in \bar{\Omega}^\varphi$  und alle  $(i, j) \in \{(2, 3), (3, 1), (1, 2)\}$ , was gleichbedeutend mit  $\sigma^\varphi(x^\varphi)^T = \sigma^\varphi(x^\varphi)$  für alle  $x^\varphi \in \bar{\Omega}^\varphi$  ist.  $\blacksquare$

Kennen wir die Abbildung  $\sigma^\varphi$ , so können wir den Vektor der Kraftdichte in deformierten Gebiet in Abhängigkeit von der Stelle  $x^\varphi \in \Omega^\varphi$  in Richtung  $n^\varphi \in S^1$  beides bezogen auf das deformierte Gebiet als

$$p^\varphi(x^\varphi, n^\varphi) = \sigma^\varphi(x^\varphi) \cdot n^\varphi$$

ausdrücken. Damit wir  $p^\varphi$  auch bezogen auf das Referenzgebiet angeben können, definieren wir die Abbildung  $S_1^\varphi : \Omega \rightarrow \operatorname{Sym}(3)$  mittels

$$S_1^\varphi(x) := \sigma^\varphi(x^\varphi) \operatorname{Cof} \nabla \varphi(x) \quad \text{für alle } x \in \Omega.$$

Sie hat nun die Eigenschaft, dass

$$S_1^\varphi(x) \cdot n = \sigma^\varphi(x^\varphi) \operatorname{Cof} \nabla \varphi(x) \cdot n = \|\operatorname{Cof} \nabla \varphi(x)\| \sigma^\varphi(x^\varphi) \cdot n^\varphi$$

gilt. Sei  $n = a \times b$ , dann ist  $(\nabla \varphi(x) a) \times (\nabla \varphi(x) b) = \operatorname{Cof} \nabla \varphi(x) \cdot n$  die Anwendung der Deformation auf  $n$ , normiert erhalten wir

$$n^\varphi = \frac{\operatorname{Cof} \nabla \varphi(x) \cdot n}{\|\operatorname{Cof} \nabla \varphi(x)\|} \in S^1.$$

Für einen Normalenvektor  $n$  soll die Anwendung auf  $S_1(x)$  also den gleichen Kraftdichtenvektor führen wie die Anwendung des deformierten (und nicht wieder normierten) Normalenvektor auf  $\sigma^\varphi(x^\varphi)$ . Wenden wir bloß  $n^\varphi$  auf  $\sigma^\varphi(x^\varphi)$  an, so müssen wir noch mit der Längenverzerrung von  $n$  multiplizieren.

Um nun unsere Gleichgewichtsgleichung  $\operatorname{div} \sigma^\varphi(x^\varphi) = -f^\varphi(x^\varphi)$  auf das undeformierte Gebiet zu übertragen, brauchen wir zunächst folgende Proposition. Im Beweis nutzen wir die auch *Piola-Identität* genannte Gleichung  $\operatorname{div}(\operatorname{Cof} \nabla \varphi(x)) = 0$  aus Proposition 2.1.9.



#### 4. Einführung in die nichtlineare Elastizitätstheorie

**Proposition 4.1.8.** Sei  $S_1: \Omega \rightarrow \mathbb{R}^{3 \times 3}$  mit  $x \mapsto S_1(x)$  ein Elastizitätsgesetz zur Deformation  $\varphi: \Omega \rightarrow \mathbb{R}^3$  mit  $S_1^\varphi(x) = \sigma^\varphi(x^\varphi) \text{Cof}(\nabla\varphi(x))$  und  $x^\varphi = \varphi(x)$ . Dann gilt

$$\text{div } S_1^\varphi(x) = \det \nabla\varphi(x) \cdot \text{div } \sigma^\varphi(x^\varphi). \quad (4.1.9)$$

*Beweis.* Die Divergenz von  $S_1^\varphi$  und  $\sigma^\varphi$  ist zeilenweise definiert, für die  $i$ -te Zeile gilt

$$\begin{aligned} \text{div}(S_1^\varphi(x))_{i,\cdot} &= \sum_{j=1}^n \frac{\partial}{\partial x_j} (S_1(x))_{ij} \\ &= \sum_{j=1}^n \frac{\partial}{\partial x_j} \left( \det \nabla\varphi(x) \cdot \sigma^\varphi(\varphi(x)) \cdot \nabla\varphi(x)^{-T} \right) \\ &= \sum_{j=1}^n \frac{\partial}{\partial x_j} \left( \det \nabla\varphi(x) \cdot \sum_{k=1}^n \sigma^\varphi(\varphi(x))_{ik} (\nabla\varphi(x)^{-T})_{kj} \right) \\ &= \sum_{k=1}^n \sum_{j=1}^n \frac{\partial}{\partial x_j} \left( \sigma^\varphi(\varphi(x))_{ik} \cdot \det \nabla\varphi(x) (\nabla\varphi(x)^{-T})_{kj} \right) \\ &= \sum_{k=1}^n \sum_{j=1}^n \frac{\partial}{\partial x_j} \sigma^\varphi(\varphi(x))_{ik} \cdot \left( \det \nabla\varphi(x) (\nabla\varphi(x)^{-1})_{jk} \right) \\ &\quad + \underbrace{\sum_{k=1}^n \sigma^\varphi(\varphi(x))_{ik} \sum_{j=1}^n \frac{\partial}{\partial x_j} \left( \det \nabla\varphi(x) (\nabla\varphi(x)^{-T})_{kj} \right)}_{= \text{div}(\text{Cof } \nabla\varphi(x))_{k,\cdot} = 0} \end{aligned} \quad (4.1.10)$$

Mit der Kettenregel folgt

$$\frac{\partial}{\partial x_j} \sigma^\varphi(\varphi(x))_{ik} = \sum_{l=1}^n \frac{\partial}{\partial x_l^\varphi} \sigma^\varphi(x^\varphi)_{ik} \frac{\partial}{\partial x_j} \varphi_l(x) = \sum_{l=1}^n \frac{\partial}{\partial x_l^\varphi} \sigma^\varphi(x^\varphi)_{ik} (\nabla\varphi(x))_{lj}.$$

Einsetzen in (4.1.10) ergibt

$$\begin{aligned} \text{div}(S_1^\varphi(x))_{i,\cdot} &= \sum_{k=1}^n \sum_{j=1}^n \sum_{l=1}^n \frac{\partial}{\partial x_l^\varphi} \sigma^\varphi(x^\varphi)_{ik} (\nabla\varphi(x))_{lj} \det \nabla\varphi(x) (\nabla\varphi(x)^{-1})_{jk} \\ &= \sum_{k=1}^n \sum_{l=1}^n \frac{\partial}{\partial x_l^\varphi} \sigma^\varphi(x^\varphi)_{ik} \det \nabla\varphi(x) \underbrace{\sum_{j=1}^n (\nabla\varphi(x))_{lj} (\nabla\varphi(x)^{-1})_{jk}}_{=(\mathbb{1})_{lk}} \\ &= \sum_{k=1}^n \sum_{l=1}^n \frac{\partial}{\partial x_l^\varphi} \sigma^\varphi(x^\varphi)_{ik} \det \nabla\varphi(x) \delta_{lk} \\ &= \sum_{k=1}^n \frac{\partial}{\partial x_k^\varphi} \sigma^\varphi(x^\varphi)_{ik} \det \nabla\varphi(x) = \text{div}(\sigma^\varphi(x^\varphi))_{i,\cdot} \det \nabla\varphi(x). \quad \blacksquare \end{aligned}$$

**Folgerung 4.1.9.** Zudem ist

$$\int_{\partial M} S_1^\varphi(x) \, dS(x) = \int_{\partial M^\varphi} \sigma^\varphi(x^\varphi) \, dS(x^\varphi). \quad (4.1.11)$$

*Beweis.* Wir können die vorige Proposition direkt zum Substituieren benutzen. Mit dem Gaußschen Inte-

gralsatz folgt

$$\begin{aligned} \int_{\partial M} S_1^\varphi(x) dS(x) &= \int_M \operatorname{div} S_1^\varphi(x) dx = \int_M \operatorname{div} \sigma^\varphi(\varphi(x)) \det \nabla \varphi(x) dx \\ &= \int_{M^\varphi} \operatorname{div} \sigma^\varphi(x^\varphi) dx^\varphi = \int_{\partial M^\varphi} \sigma^\varphi(x^\varphi) dS(x^\varphi). \end{aligned} \quad \blacksquare$$

Nach dem Satz von Cauchy gilt die Gleichgewichtsgleichung  $f^\varphi(x^\varphi) + \operatorname{div} \sigma^\varphi(x^\varphi) = 0$  für alle  $x^\varphi \in \Omega^\varphi$ . Gilt  $\det \nabla \varphi(x) \neq 0$  für alle  $x \in \Omega$ , so ist die Gleichgewichtsgleichung äquivalent zu

$$0 = \det \nabla \varphi(x) \left( f^\varphi(\varphi(x)) + \operatorname{div} \sigma^\varphi(\varphi(x)) \right) = \det \nabla \varphi(x) f^\varphi(\varphi(x)) + \operatorname{div} S_1^\varphi(x) \quad \text{für alle } x \in \Omega.$$

Fassen wir die Ergebnisse dieses Abschnittes kurz zusammen:

Gegeben seien: Ein Gebiet  $\Omega \subseteq \mathbb{R}^3$ , eine Deformationsabbildung  $\varphi: \bar{\Omega} \rightarrow \mathbb{R}^3$ , eine Abbildung der Volumenkräfte  $f^\varphi: \bar{\Omega} \rightarrow \mathbb{R}^3$  und eine Flächenkraftabbildung auf einem Teil des Randes  $g^\varphi: \Gamma \subseteq \partial\Omega \rightarrow \mathbb{R}^3$ . Wir interessieren uns für die Flächenkraftabbildung  $p^\varphi: \bar{\Omega} \rightarrow \mathbb{R}^3$ , welche das Kraft- und Drehmomentgleichgewicht, sowie die Randbedingung (4.1.4) erfüllt. Per Axiom (Spannungsprinzip von Euler und Cauchy) setzen wir die Existenz eines solchen  $p^\varphi$  voraus. Dieses lässt sich dann nach dem Satz von Cauchy schreiben als  $p^\varphi(x^\varphi, n^\varphi) = \sigma^\varphi(x^\varphi) \cdot n^\varphi$  für eine Abbildung  $\sigma^\varphi: \Omega^\varphi \rightarrow \mathbb{R}^{n \times n}$ . Alternativ definieren wir für Argumente in  $\Omega$  die Abbildung  $S_1^\varphi: \Omega \rightarrow \mathbb{R}^{3 \times 3}$  mit  $S_1^\varphi(x) := \sigma^\varphi(x^\varphi) \operatorname{Cof} \nabla \varphi(x)$ . Zudem gilt  $\sigma^\varphi(x^\varphi) \in \operatorname{Sym}(3)$  für alle  $x \in \Omega$  und die Eigenschaft  $\operatorname{div} \sigma^\varphi(x^\varphi) + f^\varphi(x^\varphi)$  bzw.  $\det \nabla \varphi(x) f^\varphi(\varphi(x)) + \operatorname{div} S_1^\varphi(x) = 0$  ist erfüllt.

Weitere Information findet sich in Ciarlet [16].

## 4.2. Modellierung über Energiefunktionen

Im vorigen Abschnitt haben wir die Verformung eines Referenzgebietes  $\Omega \subseteq \mathbb{R}^n$  modelliert, indem wir die Deformationsabbildung  $\varphi: \Omega \rightarrow \mathbb{R}^n$  vorgegeben haben und unter Vorgabe einer Volumenkraftabbildung  $f^\varphi: \Omega \rightarrow \mathbb{R}^n$  Eigenschaften einer möglichen aus den Vorgaben resultierenden Flächenkraftabbildung hergeleitet haben.

In diesem Abschnitt zeigen wir die Alternative auf, die Verformung über Potentiale zu modellieren.

Wir geben hierfür die Deformationsabbildung nur für einen Teil des Randes des Referenzgebietes  $\Gamma \subseteq \partial\Omega$  vor, dazu aber noch eine zweimal stetig differenzierbare Energiefunktion  $W: \mathbb{R}^n \times \operatorname{GL}^+(3) \rightarrow \mathbb{R}$  mit  $(x, p) \mapsto W(x, p)$  mit  $W(x, QX) = W(x, X)$  für alle  $X \in \operatorname{GL}^+(3)^2$  und alle  $Q \in \operatorname{SO}(3)$  und eine stetig differenzierbare Funktion  $P: \mathbb{R}^n \rightarrow \mathbb{R}$ .

Hierbei ist  $P$  ein Potential für das Beschleunigungsfeld  $H: \mathbb{R}^n \rightarrow \mathbb{R}^n$  mit

$$\operatorname{DP}[y] = -H(y),^3 \quad (4.2.1)$$

also  $\operatorname{DP}[x^\varphi] = -H(x^\varphi)$ . Dieses Beschleunigungsfeld ist unabhängig von  $\varphi$ . Die Volumenkraftabbildung ergibt sich dann als Produkt aus der Beschleunigung in  $x^\varphi$  mit der Dichte des verformten Materials in  $x^\varphi$ :

$$f^\varphi(x^\varphi) = -\frac{1}{\det \nabla \varphi(x)} H(x^\varphi) = \frac{1}{\det \nabla \varphi(x)} \operatorname{DP}(x^\varphi). \quad (4.2.2)$$

Weiterhin ist  $W$  das Potential für  $S_1^\varphi: \Omega \rightarrow \mathbb{R}^{3 \times 3}$  mit

$$\frac{\partial W}{\partial p}(x, p) = S_1^\varphi(x), \quad (4.2.3)$$

<sup>2</sup>Diese Eigenschaft nennt sich Objektivität und wird an späterer Stelle noch thematisiert

<sup>3</sup>Das Minuszeichen ist erforderlich, da die Beschleunigungskräfte in Richtung kleiner (und nicht größer) werdenden Potentials wirken.

#### 4. Einführung in die nichtlineare Elastizitätstheorie

also  $\frac{\partial W}{\partial p}(x, \nabla\varphi(x)) = S_1^\varphi(x)$ . Dadurch, dass  $W$  nur von  $x$  und  $\nabla\varphi(x)$  abhängt, setzen wir voraus, dass  $S_1^\varphi$  für jedes  $x$  nicht von ganz  $\varphi$  abhängen kann, sondern nur von den linearen Termen einer Taylorentwicklung von  $\varphi$  in  $x$ .

Da  $S_1^\varphi$  in  $\Omega$  „lebt“, also Argumente und Funktionswerte sich auf  $\Omega$  beziehen, verschwindet zudem die Verrechnung wie bei der Volumenkraft mit der Dichte  $(\det \nabla\varphi(x))^{-1}$  in  $\Omega^\varphi$  und somit ergibt sich  $S_1^\varphi$  direkt aus der Ableitung.

Sei  $M := \{\varphi \in C^2(\bar{\Omega} \mid \nabla\varphi \in \text{GL}^+(3) \text{ und } \varphi_\Gamma = g)\}$  nun die Menge aller zweimal stetig differenzierbaren Deformationsabbildungen ohne Selbstdurchdringung auf  $\bar{\Omega}$ . Wir definieren auf  $M$  das Energiefunktional

$$I: M \rightarrow \mathbb{R} \quad \text{mit} \quad I(\varphi) = \int_{\Omega} W(x, \nabla\varphi(x)) + P(\varphi(x)) \, dx \quad (4.2.4)$$

und setzen axiomatisch voraus, dass eine Verformung immer nach dem Zustand minimaler Energie strebt, die Deformationsabbildung  $\varphi_0$ , welche  $I$  minimiert, genau die unter den Vorgaben sich physikalisch einstellende Deformationsabbildung ist.

Wenden wir Lemma 2.10.18 mit  $L(x, y, p) := W(x, p) + P(y)$  an. Es ist

$$\nabla_y L(x, y, p) = H(y) \quad \text{und} \quad \nabla_p L(x, y, p) = S_1^\varphi(x) \quad (4.2.5)$$

und damit

$$\nabla_y L(x, x^\varphi, \nabla\varphi(x)) = -\det \nabla\varphi(x) f^\varphi(\varphi(x)) \quad \text{und} \quad \nabla_p L(x, x^\varphi, \nabla\varphi(x)) = S_1^\varphi(x) \quad (4.2.6)$$

und wir erhalten für das  $\varphi$ , welches die Energie minimiert, die *Euler-Lagrange-Gleichung*

$$\det \nabla\varphi(x) f^\varphi(\varphi(x)) + \text{div } S_1^\varphi(x) = 0 \quad \text{für alle } x \in \Omega, \quad (4.2.7)$$

in Übereinstimmung mit dem letzten Abschnitt.

Nutzen wir nun noch die Bedingung  $W(x, QX) = W(x, X)$  für alle  $X \in \text{GL}^+(n)$  und  $Q \in \text{SO}(n)$ : Mit Folgerung 2.10.26 erhalten wir  $\nabla\varphi(x)^{-1} S_1(\nabla\varphi(x)) = \nabla\varphi(x)^{-1} S_1^\varphi(x) \in \text{Sym}(n)$ , also

$$\left( \det \nabla\varphi(x) \nabla\varphi(x)^{-1} \sigma^\varphi(x^\varphi) \nabla\varphi(x)^{-T} \right)^T = \det \nabla\varphi(x) \nabla\varphi(x)^{-1} \sigma^\varphi(x^\varphi) \nabla\varphi(x)^{-T}, \quad (4.2.8)$$

was äquivalent ist zu

$$\det \nabla\varphi(x) \nabla\varphi(x)^{-1} \sigma^\varphi(x^\varphi)^T \nabla\varphi(x)^{-T} = \det \nabla\varphi(x) \nabla\varphi(x)^{-1} \sigma^\varphi(x^\varphi) \nabla\varphi(x)^{-T}, \quad (4.2.9)$$

sodass Kürzen auf  $\sigma^\varphi(x^\varphi)^T = \sigma^\varphi(x^\varphi)$  führt, welches eine weitere Eigenschaft von  $\sigma^\varphi$  aus dem letzten Abschnitt ist.

### 4.3. Nichtlineares Elastizitätsgesetz

Wir haben in Abschnitt 4.1 zu einer gegebenen Deformation die Eigenschaft der Funktion  $\sigma^\varphi: \Omega^\varphi \rightarrow \mathbb{R}^{3 \times 3}$  hergeleitet, die in jedem Punkt  $x^\varphi \in \Omega^\varphi$  des verformten Gebiets die Flächenkraftdichte auf eine Tangentialfläche mit Normalenvektor  $n^\varphi \in S^1$  mittels  $\sigma^\varphi(x^\varphi) \cdot n^\varphi$  angibt. Auch wenn wir  $\sigma^\varphi$  nicht explizit angeben konnten, so konnten wir zumindest  $\sigma^\varphi(x^\varphi) \in \text{Sym}(3)$  und  $\text{div } \sigma^\varphi(x^\varphi) + f^\varphi(x^\varphi) = 0$  nachweisen. Wir haben auch die Funktion  $S_1: \Omega \rightarrow \mathbb{R}^n$  über den Zusammenhang  $S_1^\varphi(x) := \sigma^\varphi(x^\varphi) \text{Cof } \nabla\varphi(x)$  eingeführt, welche die Flächenkräfte mittels  $S_1^\varphi(x) \cdot n$  angibt, wobei der Zusammenhang  $\|\text{Cof } \nabla\varphi(x)\| \cdot n^\varphi = \text{Cof } \nabla\varphi(x) \cdot n$  gilt.

In Abschnitt 4.2 haben wir Potentiale  $W: \Omega \times \text{GL}^+(3) \rightarrow \mathbb{R}$  und  $P: \mathbb{R}^n \rightarrow \mathbb{R}$  und damit implizit  $S_1^\varphi$  bzw.  $\sigma^\varphi$  und  $f^\varphi$  vorgegeben und über die Annahme einer die Energie

$$I(\varphi) = \int_{\Omega} W(x, \nabla\varphi(x_0)) + P(\varphi(x)) \, dx \quad (4.3.1)$$

minimierenden Deformationsabbildung dieselben Eigenschaften  $\sigma^\varphi(x^\varphi) \in \text{Sym}(3)$  und  $\text{div } \sigma^\varphi(x^\varphi) + f^\varphi(x^\varphi) = 0$  nachgewiesen. Allerdings haben wir hierfür die Eigenschaft  $W(x, QX) = W(x, X)$  für alle  $X \in \text{GL}^+(3)$  und  $Q \in \text{SO}(3)$ , die wir im Folgenden noch als Objektivität kennenlernen werden, ausgenutzt.

Desweiteren haben wir durch  $S_1^\varphi = D_p W[x, p] \Big|_{p=\nabla\varphi(x)}$  implizit eine weitere Annahme an  $S_1^\varphi$  und damit auch an  $\sigma^\varphi$  gemacht. Wir nennen sie die Lokalität der Deformation.

### 4.3.1. Lokalität der Deformation

Betrachten wir die Deformation abgegrenzt auf ein hinreichend kleines Gebiet  $\Omega^* \subseteq \Omega$  (z.B. eine Kugel oder ein Würfel) um einen inneren Punkt des Referenzgebietes  $x_0 \in \Omega$ , so sind die Terme höherer Ordnung vernachlässigbar. Unter Berücksichtigung der linearen Terme der Taylor-Entwicklung von  $\varphi$  in  $x_0$  ist die Abweichung quadratisch in  $\|h\|$ :

$$\varphi(x_0 + h) = \underbrace{x_0^\varphi + \nabla\varphi(x_0) \cdot h}_{\text{lineare Verzerrung}} + \mathcal{O}(\|h\|^2). \quad (4.3.2)$$

Wir nennen eine Deformation nun *lokal*, wenn  $\sigma^\varphi$  nur vom Verhalten von  $\varphi$  in einem so kleinen Gebiet  $\Omega^* \subseteq \Omega$  abhängt, speziell nur von  $x_0^\varphi$  und  $\nabla\varphi(x)$  (nicht von  $\varphi(y)$  mit  $y \neq x$ , nicht von höheren Ableitungen von  $\varphi$  an der Stelle  $x$ , etc.). Diese Abhängigkeit impliziert die Existenz einer Abbildung  $\tilde{\sigma}: \Omega^\varphi \times \text{GL}^+(3) \rightarrow \text{Sym}(3)$  mit  $\tilde{\sigma}(x^\varphi, \nabla\varphi(x)) = \sigma^\varphi(x^\varphi)$ .

Mit gleicher Argumentation gibt es eine Abbildung  $\tilde{S}_1: \Omega^\varphi \times \text{GL}^+(3) \rightarrow \mathbb{R}^{3 \times 3}$  mit  $\tilde{S}_1(x^\varphi, \nabla\varphi(x)) = \sigma^\varphi(x^\varphi) \text{Cof } \nabla\varphi(x)$ .

Noch einen Schritt weiter gehen wir mit der Annahme an die Homogenität der Deformation. Hierfür müssen wir aber Einschränkungen an die physischen Eigenschaften des verformten Materials hinnehmen.

### 4.3.2. Homogenität der Deformation

Wenn wir davon ausgehen, dass das deformierte Material im physikalischen Sinn homogen ist, d.h. das Material unabhängig von der Stelle auf Verformung mit Spannung reagiert, können wir den Zusammenhang zwischen  $\varphi$  und  $\sigma^\varphi$  bzw.  $S_1^\varphi$  vereinfachen:

Wir bezeichnen eine Deformation als *homogen*, wenn aus  $\nabla\varphi(x_0) = \nabla\varphi(x_1)$  für  $x_0, x_1 \in \Omega$  immer  $\sigma^\varphi(x_0^\varphi) = \sigma^\varphi(x_1^\varphi)$  folgt, d.h. bei gleichem Deformationsgradienten in  $x_0$  und  $x_1$  die gleichen Flächenkräfte herrschen.

Die ausschließliche Abhängigkeit von  $\nabla\varphi(x)$  impliziert die Existenz einer Abbildung  $\sigma: \text{GL}^+(3) \rightarrow \text{Sym}(3)$  mit  $\sigma(\nabla\varphi(x)) = \sigma^\varphi(x^\varphi)$  und damit auch die Existenz einer Abbildung  $S_1: \text{GL}^+(3) \rightarrow \mathbb{R}^{3 \times 3}$  mit  $S_1(\nabla\varphi(x)) = \sigma^\varphi(x^\varphi) \text{Cof } \nabla\varphi(x)$ .

Außerdem überträgt sich die Homogenität auch – falls vorhanden – auf das Potential zu  $S_1$ . Wir schreiben auch hier gleich  $W(p)$  anstatt  $W(x, p)$  und können zu  $DW(\nabla\varphi(x)) = S_1(\nabla\varphi(x))$  vereinfachen.

Bezeichnen wir nun  $F := \nabla\varphi(x) \in \text{GL}^+(3)$  als den *Deformationsgradienten* an der Stelle  $x \in \Omega$ , dann erhalten wir

$$F \mapsto W(F), \quad F \mapsto \sigma(F), \quad F \mapsto S_1(F) \quad \text{mit} \quad DW(F) = S_1(F) = \sigma(F) \text{Cof } F. \quad (4.3.3)$$

**Bemerkung 4.3.1.** Unter Vernachlässigung der Volumenkräfte können wir (4.2.7) nun zu  $\text{div } S_1(F) = 0$  vereinfachen; der sog. *Divergenzfreiheit des ersten Piola-Kirchhoff-Spannungstensors*.

### 4.3.3. Polarzerlegung des Deformationsgradienten

Nach Satz 2.4.2 (Satz über die Polarzerlegung) können wir den Deformationsgradienten  $F \in \text{GL}^+(3)$  eindeutig zerlegen in

$$F = V R = R U \quad \text{mit} \quad V, U \in \text{Sym}^+(3) \quad \text{und} \quad R \in \text{SO}(3). \quad (4.3.4)$$

#### 4. Einführung in die nichtlineare Elastizitätstheorie

Hierbei sind sowohl  $V$  als auch  $U$  als symmetrische Matrizen diagonalisierbar und haben aufgrund von  $V = R U R^T$  dieselben Eigenwerte  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_+$ . Diese nennen wir auch die Singulärwerte von  $F$ .

Betrachten wir nun den linearen Teil der Deformation  $\varphi(x_0 + h)$

$$\varphi(x_0 + h) - \mathcal{O}(\|h\|^2) = \varphi(x_0) + F.h, \quad (4.3.5)$$

so können wir diese linearisierte Deformation im Falle einer rechten Polarzerlegung  $F = VR$  als die Überlagerung von genau drei Bewegungen auffassen:

- i) einer Rotation um  $R \in \text{SO}(3)$ ,
- ii) einer Dehnung<sup>4</sup> (Streckung/Stauchung)  $V \in \text{Sym}^+(3)$  in drei zueinander orthogonalen Richtungen (den Hauptdehnungsachsen),
- iii) einer Verschiebung um  $\varphi(x_0) \in \mathbb{R}^3$ .

Verwenden wir anstatt der rechten die linke Polarzerlegung  $F = RU$  mit  $U \in \text{Sym}^+(3)$  und dem gleichen  $R \in \text{SO}(3)$  wie oben, so besteht  $\varphi(x_0 + h) = \varphi(x_0) + RU.h$  auch aus drei Bewegungen, nur i) mit ii) vertauscht.

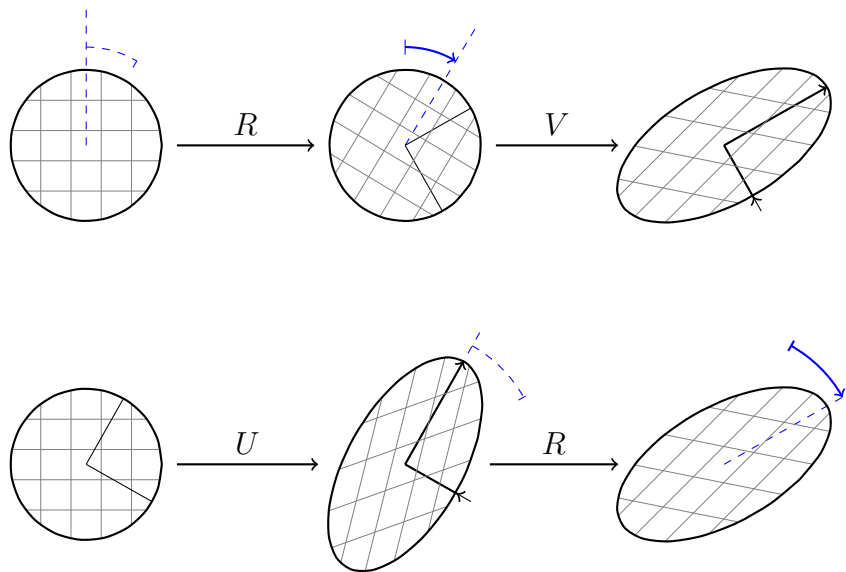


Abbildung 4.5.: rechte und linke Polarzerlegung des Deformationsgradienten

#### 4.3.4. Konstitutives Gesetz

Unter Voraussetzung der Homogenität haben wir den Zusammenhang zwischen Deformation und resultierender Spannung als  $F \mapsto \sigma(F)$  bzw.  $F \mapsto S_1(F)$  hergestellt. Für jedes  $F$  sind  $\sigma(F)$  und  $S_1(F)$  Spannungstensoren, wir werden in diesem Abschnitt noch  $S_2(F)$  und  $T^{\text{Biot}}(F)$  vorstellen. Alle haben allerdings die Eigenschaft, ineinander überführbar (so wie  $S_1(F) = \sigma(F) \text{Cof } F$ ) zu sein. Die Abbildung von  $F$  in die Äquivalenzklasse der zugehörigen Spannungstensoren nennen wir das *Elastizitätsgesetz* oder auch das *konstitutive Gesetz*.

Es sind unterschiedliche konstitutive Gesetze denkbar, manche passen besser auf gewisse physikalische Beobachtungen bei Deformationsexperimenten als andere. Eine physikalisch motivierte Eigenschaft, die wir von jedem Elastizitätsgesetz fordern, ist  $\sigma(\mathbb{1}) = 0$ . Wir nennen dies die Eigenschaft der spannungsfreien Referenzkonfiguration (kurz gesagt: „ohne Deformation keine Spannung“). Eine andere Eigenschaft, die wir bereits (als  $W(QX) = W(X)$ ) kennengelernt haben und die im Folgenden noch näher erläutert wird, ist die Objektivität. Diese setzen wir auch stets voraus.

<sup>4</sup>auch Dilatation genannt

Mögliche konstitutive Gesetze sind  $\sigma(F) = V - \mathbb{1}$  oder  $\sigma(F) = \log V$ .

Im eindimensionalen Fall entspricht  $F \mapsto \sigma(F) = V - \mathbb{1}$  dem Hooke'schen Gesetz: Wird eine Strecke der Länge 1 auf die Länge  $\lambda$  gestaucht/gestreckt, ist  $F = V = \lambda \cdot \mathbb{1}$ . Die wirkenden Kräfte sind nun proportional zur Auslenkung, also  $\lambda - 1$ .

### 4.3.5. Spannungstensoren

Die in diesem Unterabschnitt vorgestellten Spannungstensoren haben allesamt die Eigenschaft, auf einen Normalenvektor angewandt, die in einem Sinne zugehörige Flächenkraft auszudrücken. Wir haben hierbei verschiedene Möglichkeiten der Wahl des Bezugssystems.

Sei  $n \in S^1 := \{v \in \mathbb{R}^3 \mid \|v\| = 1\}$  ein Normalenvektor, senkrecht stehend auf der Tangentialebene in einem Punkt des Randes des Referenzgebietes  $\Omega$ . Der Vektor  $n$  repräsentiert somit eine Teilfläche dieser Ebene mit Flächenmaß eins. Wir haben bereits gesehen, dass diese Fläche bei einer Verformung des Referenzgebietes mittels  $F \in GL^+(3)$  in das deformierte Gebiet zu der Fläche transformiert wird, auf der  $(\text{Cof } F) \cdot n$  senkrecht steht, und die das Flächenmaß  $\|(\text{Cof } F) \cdot n\|$  besitzt. Die normierte zu  $n$  korrespondierende Fläche im deformierten Gebiet bezeichnen wir von nun an mit  $\hat{n} = \frac{(\text{Cof } F) \cdot n}{\|(\text{Cof } F) \cdot n\|}$ .

Wir können es auch so ausdrücken:  $\|(\text{Cof } F) \cdot n\|$  entspricht genau der Flächenverzerrung durch  $F$ : Der Flächeninhalt einer Fläche mit Normalenvektor  $n$  verändert sich durch die Deformation um den Faktor  $\|(\text{Cof } F) \cdot n\|$ .

Die Anwendung von  $\hat{n}$  auf  $\sigma$  und  $n$  auf  $S_1$  haben wir bereits kennengelernt, hinzu kommt jetzt noch die Anwendung von  $n$  auf  $S_2$  und auf  $T^{\text{Biot}}(F)$ .

Als Übersicht soll das folgende Beispiel dienen:

**Beispiel 4.3.2.** In Abbildung 4.6 deformieren wir einen Einheitswürfel mittels  $F = \begin{pmatrix} 4/5 & 1 & 0 \\ -1/5 & 6/5 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ .

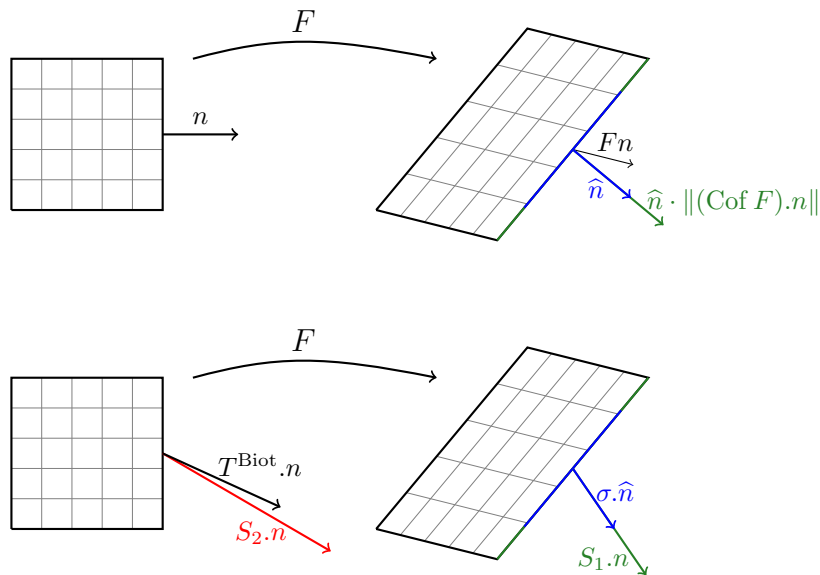


Abbildung 4.6.: Seitenansicht der Deformation eines Einheitswürfels und wirkende Kräfte

### Cauchy-Spannungstensor

Der *Cauchy-Spannungstensor*  $\sigma$  drückt die Spannung im verformten Material bezogen auf Flächen im verformten Material aus (kurz gesagt: Kraft pro deformierter Fläche). Hierbei beschreiben die Einträge

#### 4. Einführung in die nichtlineare Elastizitätstheorie

der Matrix  $\sigma(F)$  mit ihren neun Einträgen die jeweils drei Komponenten der drei Flächenkräfte zu  $\hat{n} = e_1, e_2, e_3$ . Die Symmetrie von  $\sigma(F)$  entspricht der Drehmomentfreiheit des Zusammenwirkens der drei Kräfte.

So entspricht  $\sigma.n^\varphi$  der Kraft im verformten Material auf eine Fläche mit Flächenmaß 1, auf der  $n$  senkrecht steht.

##### Erster Piola-Kirchhoff-Spannungstensor

Der *erste Piola-Kirchhoff-Spannungstensor*  $S_1$  drückt die Spannung im verformten Material bezogen auf Flächen im unverformten Material aus (kurz gesagt: Kraft pro ursprünglicher Fläche).

Hierzu ist  $S_1$  definiert als

$$S_1(F) := \sigma(F) \text{Cof } F \quad (4.3.6)$$

und hat die Eigenschaft

$$S_1(F).n = \sigma(F)(\text{Cof } F).n = \|(\text{Cof } F).n\| \cdot \sigma(F).n^\varphi.$$

Wir sehen,  $\|(\text{Cof } F).n\|$  bildet den Verzerrungsfaktor der Ebene von der Ausgangs- zur deformierten Konfiguration.

Natürlich lässt sich der Cauchy-Spannungstensor aus dem erstem Piola-Kirchhoff-Spannungstensor mittels  $\sigma(F) = \frac{1}{\det F} S_1(F) F^T$  berechnen.

##### Zweiter Piola-Kirchhoff-Spannungstensor

Der *zweite Piola-Kirchhoff-Spannungstensor*  $S_2$  transformiert die Spannung des verformten Materials bezogen auf Flächen im unverformten Material wieder zurück auf das undeformierte Bezugssystem und ist definiert als:

$$S_2(F) := F^{-1} S_1(F). \quad (4.3.7)$$

##### Biot-Spannungstensor

Der *Biot-Spannungstensor*  $T^{\text{Biot}}$  dreht die Spannung des verformten Materials bezogen auf Flächen im unverformten Material wieder zurück auf das undeformierte Bezugssystem und ist definiert als:

$$T^{\text{Biot}}(F) := R^T S_1(F). \quad (4.3.8)$$

Hierbei ist  $R \in \text{SO}(3)$  der orthogonale Faktor der Polarzerlegung  $F = VR = RU$ .

##### Umrechnungsdiagramm

In Abbildung 4.7 sehen wir ein Umrechnungsdiagramm für die verschiedenen Spannungstensoren.

### 4.3.6. Hyperelastizität

**Definition 4.3.3.** Wir bezeichnen ein Elastizitätsgesetz als *hyperelastisch*, wenn es durch eine elastische Energiefunktion<sup>5</sup>  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  induziert wird. Dies ist genau dann der Fall, wenn der erste Piola-Kirchhoff-Spannungstensor  $S(F)$  ein Gradientenfeld ist, also eine Funktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  existiert mit  $DW(F) = S_1(F)$ .

---

<sup>5</sup>welche als Modellierung eines physikalischen Zusammenhangs die physikalisch immer vorauszusetzende Eigenschaft der Objektivität immer in sich trägt

#### 4. Einführung in die nichtlineare Elastizitätstheorie

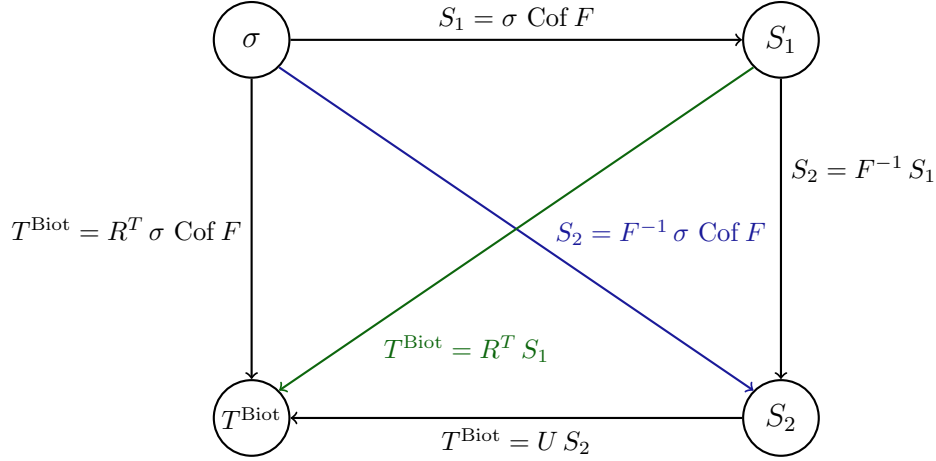


Abbildung 4.7.: Umrechnung der verschiedenen Spannungstensoren

**Bemerkung 4.3.4.** Beispiele für hyperelastische Energiefunktionen sind

$$W_{\text{SVK}}(F) = \|U^2 - \mathbb{1}\|_{\mu, \kappa}^2, \quad (4.3.9)$$

$$W_{\text{H}}(F) = \|\log U\|_{\mu, \kappa}^2. \quad (4.3.10)$$

Hierbei ist  $U = \sqrt{F^T F}$  und  $\|\cdot\|_{\mu, \kappa}$  die gewichtete Matrixnorm, definiert als

$$\|X\|_{\mu, \kappa}^2 := \mu \|\text{dev}_3 X\|^2 + \frac{\kappa}{2} (\text{tr}(X))^2 \quad (4.3.11)$$

mit dem deviatorischen (spurfreien) Anteil  $\text{dev}_n X = X - \frac{1}{n} \text{tr}(X) \mathbb{1}$  und der Frobeniusnorm  $\|X\| := \sqrt{\text{tr}(X^T X)} = \sqrt{\langle X, X \rangle}$ .

#### 4.3.7. Objektivität

Wir haben bereits in Abschnitt 4.2 die Eigenschaft  $W(x, QX) = W(x, X)$  für alle  $Q \in \text{SO}(3)$  und  $X \in \text{GL}^+(3)$  kennengelernt, die sich unter Einfluss der Homogenität zu  $W(QX) = W(X)$  vereinfacht. Bevor wir uns mit der Objektivität von Energiefunktionen beschäftigen, behandeln wir zunächst die Objektivität von Spannungsabbildungen.

Dazu sei  $\sigma: \text{GL}^+(3) \rightarrow \text{Sym}(3)$  ein Elastizitätsgesetz, gegeben durch die Abbildung von  $F \in \text{GL}^+(3)$  auf den Cauchy-Spannungstensor  $\sigma(F) \in \text{Sym}(3)$ . Und sei  $n \in \mathbb{R}^3$  ein Normalenvektor auf dem Rand des mit  $F$  verzerrten Gebietes. Zur Erinnerung: Die Kraft pro Flächeneinheit, die auf die Fläche mit Normalenvektor  $n$  wirkt, wird durch  $\sigma(F).n$  gegeben.

**Definition 4.3.5.** Wir nennen  $\sigma$  im physikalischen Sinne *objektiv*, wenn diese Kraft unabhängig vom Bezugssystem ist. Drehen wir das Bezugssystem um  $Q \in \text{SO}(3)$  nach der Deformation, d.h. wenden  $QF$  anstatt  $F$  an, so dreht sich auch der Normalenvektor zu  $Qn$  und es soll  $\sigma(QF).Qn$  die gleiche Kraft ergeben, wie  $Q\sigma(F).n$ , was auf

$$\sigma(QF) = Q\sigma(F)Q^T \quad \text{für alle } F \in \text{GL}^+(3) \text{ und } \text{SO}(3) \quad (4.3.12)$$

führt.

**Bemerkung 4.3.6.** Aufgrund von physikalischen Überlegungen muss bei (physikalisch plausiblen) Elastizitätsgesetzen immer die Eigenschaft der Objektivität gelten! So werden wir im Folgenden nicht immer explizit erwähnen, wenn ein Elastizitätsgesetz objektiv ist.



**Lemma 4.3.7.** *Ein Elastizitätsgesetz ist genau dann objektiv, wenn eine der folgenden zueinander äquivalenten Bedingungen erfüllt ist:*

- i)  $\sigma(QF) = Q\sigma(F)Q^T$  für alle  $F \in \text{GL}^+(3)$  und  $Q \in \text{SO}(3)$ .
- ii)  $S_1(QF) = QS_1(F)$  für alle  $F \in \text{GL}^+(3)$  und  $Q \in \text{SO}(3)$ .
- iii)  $S_2(QF) = S_2(F)$  für alle  $F \in \text{GL}^+(3)$  und  $Q \in \text{SO}(3)$ .
- iv)  $T^{\text{Biot}}(QF) = T^{\text{Biot}}(F)$  für alle  $F \in \text{GL}^+(3)$  und  $Q \in \text{SO}(3)$ .

*Beweis.* Die Äquivalenz der Objektivität zu i) ergibt sich direkt aus der Definition. Wir zeigen noch, dass i), ii), iii) und iv) zueinander äquivalent sind:

Sei  $Q \in \text{SO}(3)$  und  $F \in \text{GL}^+(3)$  mit Polarzerlegung  $F = UR$  und  $U \in \text{Sym}^+(3)$  und  $R \in \text{SO}(3)$ , somit hat  $QF \in \text{GL}^+(3)$  die Polarzerlegung  $QF = \hat{R}$  mit gleichem  $U$  und  $\hat{R} = QR \in \text{SO}(3)$ . Zudem gilt  $\text{Cof}(QF) = Q \text{Cof}(F)$  nach Folgerung 2.1.29. Nun können wir folgern

i)  $\Rightarrow$  ii):

$$S_1(QF) = \sigma(QF) \text{Cof}(QF) = Q\sigma(F)Q^T Q \text{Cof} F = QS_1(F)$$

ii)  $\Rightarrow$  iii):

$$S_2(QF) = (QF)^{-1} S_1(QF) = F^{-1} Q^T Q S_1(F) = S_2(F)$$

iii)  $\Rightarrow$  iv):

$$T^{\text{Biot}}(QF) = (QR)^T (QF) S_2(QF) = R^T Q^T Q F S_2(F) = T^{\text{Biot}}(F)$$

iv)  $\Rightarrow$  i):

$$\sigma(QF) = (QR) T^{\text{Biot}}(QF) \text{Cof}(QF)^{-1} = QRT^{\text{Biot}}(F) \text{Cof}(F)^{-1} Q^T = Q\sigma(F)Q^T \quad \blacksquare$$

**Lemma 4.3.8.** *Sei  $T^{\text{Biot}}: \text{GL}^+(3) \rightarrow \text{Sym}(3)$  ein objektives Elastizitätsgesetz. Dann ist  $T^{\text{Biot}}$  allein schon durch die Abbildungsvorschrift auf  $\text{Sym}^+(3)$  bestimmt, d.h. es existiert eine Abbildung  $\tilde{T}^{\text{Biot}}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  mit  $\tilde{T}^{\text{Biot}} = T^{\text{Biot}}|_{\text{Sym}^+(3)}$  und  $T^{\text{Biot}}(F) = \tilde{T}^{\text{Biot}}(U)$  für alle  $F = RU \in \text{GL}^+(3)$  mit  $U \in \text{Sym}^+(3)$  und  $Q \in \text{SO}(3)$ .*

*Beweis.* Sei  $F \in \text{GL}^+(3)$  mit Polarzerlegung  $F = RU$  mit  $U \in \text{Sym}^+(3)$  und  $R \in \text{SO}(3)$  gegeben. Ist  $T^{\text{Biot}}(U)$  bekannt, so gilt aufgrund der Objektivität  $T^{\text{Biot}}(F) = T^{\text{Biot}}(RU) = T^{\text{Biot}}(U)$ , also ist auch  $T^{\text{Biot}}(F)$  bekannt.  $\blacksquare$

**Bemerkung 4.3.9.** Die Einschränkung auf  $\text{Sym}^+(3)$  werden wir im Folgenden anstatt mit  $\tilde{T}^{\text{Biot}}$  auch mit  $T^{\text{Biot}}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  als Elastizitätsgesetz in  $U$  ausdrücken.

Zudem erhalten wir mit  $\hat{T}^{\text{Biot}}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  und  $C \mapsto \hat{T}^{\text{Biot}}(C) := T^{\text{Biot}}(\sqrt{C})$  ein  $T^{\text{Biot}}$  zugehöriges Elastizitätsgesetz in  $C$ .

### Objektivität von Energiefunktionen

Wir führen nun auch formal den bereits schon verwendeten Begriff der Objektivität von Energiefunktionen ein:

**Definition 4.3.10.** Wir bezeichnen eine Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  als objektiv, wenn

$$W(QF) = W(F) \quad \text{für alle } F \in \text{GL}^+(3) \text{ und } Q \in \text{SO}(3) \quad (4.3.13)$$

gilt.

#### 4. Einführung in die nichtlineare Elastizitätstheorie

**Bemerkung 4.3.11.** Ist  $F = RU$  mit  $R \in \text{SO}(3)$  und  $U \in \text{Sym}^+(3)$  die (linke) Polarzerlegung des Deformationsgradienten  $F \in \text{GL}^+(3)$ , dann gilt mit Objektivität  $W(F) = W(U)$ . So ist  $W$  allein schon durch die Einschränkung auf die Menge der positiv definiten symmetrischen Matrizen  $W|_{\text{Sym}^+(3)}$  festgelegt. So schreiben wir gleichwertig auch  $W: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit  $U \mapsto W(U)$  als Energiefunktion.

**Proposition 4.3.12.** Die Objektivität einer Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  impliziert die Existenz einer Energiefunktion  $\widehat{W}: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$\widehat{W}(F^T F) = W(F) \quad \text{für alle } F \in \text{GL}^+(3). \quad (4.3.14)$$

**Bemerkung 4.3.13.** Das Argument von  $\widehat{W}$  bezeichnen wir im Allgemeinen mit  $C := F^T F \in \text{Sym}^+(3)$ , dem rechten Cauchy-Green-Deformationstensor [32].

*Beweis.* Für alle  $X \in \text{Sym}^+(3)$  ist die Abbildung  $X \mapsto \sqrt{X}$  wohldefiniert. Mit  $\widehat{W}(X) := W(\sqrt{X})$  erhalten wir

$$\widehat{W}(F^T F) = W(\sqrt{F^T F}) = W(U) = W(F),$$

wobei  $F = RU$  die (linke) Polarzerlegung von  $F \in \text{GL}^+(3)$  ist. ■

**Lemma 4.3.14.** Sei  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  eine Energiefunktion. Ist  $W$  objektiv, so ist das durch  $W$  induzierte Elastizitätsgesetz objektiv.

*Beweis.* Wir entwickeln zunächst  $W$  an  $F \in \text{GL}^+(3)$  mit der Taylorformel zu

$$W(F + H) = W(F) + \langle S_1(F), H \rangle + \mathcal{O}(\|H\|^2) \quad \text{für alle } H \in \text{GL}^+(3). \quad (4.3.15)$$

Nehmen wir nun einerseits Objektivität an, dann gilt

$$W(QF) = W(F) \quad \text{und} \quad W(Q(F + H)) = W(F + H) \quad \text{für alle } F, H \in \text{GL}^+(3) \text{ und } Q \in \text{SO}(3).$$

Entwickeln wir  $W$  in  $QF \in \text{GL}^+(3)$

$$W(QF + \widehat{H}) = W(QF) + \langle S_1(QF), \widehat{H} \rangle + \mathcal{O}(\|\widehat{H}\|^2)$$

und setzen  $\widehat{H} = QH$ , erhalten wir  $\mathcal{O}(\|\widehat{H}\|^2) = \mathcal{O}(\|H\|^2)$  und

$$\begin{aligned} W(F + H) &= W(Q(F + H)) = W(QF) + \langle S_1(QF), QH \rangle + \mathcal{O}(\|H\|^2) \\ &= W(F) + \langle Q^T S_1(QF), H \rangle + \mathcal{O}(\|H\|^2), \end{aligned}$$

was mit (4.3.15) auf

$$\langle S_1(F) - Q^T S_1(QF), H \rangle = \mathcal{O}(\|H\|^2) \quad \text{für alle } H \in \text{GL}^+(3)$$

und damit auf  $S_1(F) = Q^T S_1(QF)$  führt. ■

Überführen wir Proposition 2.10.25 auf die hier verwendete Notation.

**Proposition 4.3.15.** Sei  $W: \text{GL}^+(n) \rightarrow \mathbb{R}$  eine objektive Energiefunktion und  $\widehat{W}: \text{Sym}^+(n) \rightarrow \mathbb{R}$  mit  $\widehat{W}(C) = W(F)$  für alle  $F \in \text{GL}^+(n)$  die hieraus resultierende Energiefunktion in  $C := F^T F$ . Mit  $S_1(F) = DW(F)$  und  $S_2(F) = F^{-1} S_1(F)$  gilt für alle  $F, H \in \text{GL}^+(n)$

$$D_C \widehat{W}(C) = \frac{1}{2} S_2(F) = \frac{1}{2} F^{-1} S_1(F) \in \text{Sym}(n) \quad (4.3.16)$$

und

$$\langle S_1(F), H \rangle = \langle D_C \widehat{W}(C), F^T H + H^T F \rangle. \quad (4.3.17)$$

*Beweis.* Siehe Beweis zu Proposition 2.10.25 auf Seite 56. ■

### 4.3.8. Isotropie

Anders als die Objektivität impliziert die in diesem Abschnitt vorgestellte Isotropie eine Einschränkung an die physikalischen Eigenschaften des deformierten Materials. So sind sowohl isotrope, als auch anisotrope Elastizitätsgesetze denkbar. Wir bezeichnen mit Isotropie zunächst einen physikalischen Begriff, der sich unter der Annahme von Objektivität als zur klassischen (mathematischen) Isotropie<sup>6</sup> als äquivalent erweisen wird.

**Definition 4.3.16.** Wir nennen ein Elastizitätsgesetz  $\sigma: \text{GL}^+(3) \rightarrow \text{Sym}(3)$  im physikalischen Sinne *isotrop*, wenn das verzerrte Material keine Vorzugsrichtung<sup>7</sup> aufweisen soll. Wenn wir beispielsweise unseren Würfel vor der Verzerrung durch eine Vierteldrehung in einen Würfel gleicher Form überführen, soll dieser sich unter  $F$  mit der gleichen Kraft verformen, es soll also  $\sigma(FQ).n = \sigma(F).n$  gelten, was auf

$$\sigma(FQ) = \sigma(F) \quad \text{für alle } F \in \text{GL}^+(3) \text{ und } Q \in \text{SO}(3) \quad (4.3.18)$$

führt.

Beispiele für isotrope Materialien sind Gummi und Stahl, Beispiele für nicht isotrope Materialien sind Carbon und Holz.

**Bemerkung 4.3.17.** Wir erinnern uns: Für ein gegebenes Elastizitätsgesetz sind die Abbildungen

$$S_1, S_2, T^{\text{Biot}}, \sigma: \text{GL}^+(3) \rightarrow \text{Sym}(3), \quad (4.3.19)$$

welche den Deformationsgradienten  $F \in \text{GL}^+(3)$  auf einen der äquivalenten Spannungstensoren  $S_1(F)$ ,  $S_2(F)$ ,  $T^{\text{Biot}}(F)$ ,  $\sigma(F)$  abbilden, ineinander umrechenbar (vgl. Abbildung 4.7). Sei  $F = VR$  die Polarzerlegung mit  $V \in \text{Sym}^+(3)$  und  $R \in \text{SO}(3)$ , dann gelten die zyklischen Umrechnungsformeln, welche wir in den folgenden beiden Lemmata verwenden werden:

$$\begin{aligned} S_1(F) &= \sigma(F) \text{Cof}(F), \\ S_2(F) &= F^{-1} S_1(F), \\ T^{\text{Biot}}(F) &= R^T \sigma(F) \text{Cof}(F), \\ \sigma(F) &= R T^{\text{Biot}}(F) \text{Cof}(F)^{-1}. \end{aligned} \quad (4.3.20)$$

**Lemma 4.3.18.** Ein Elastizitätsgesetz ist genau dann isotrop, wenn eine der folgenden zueinander äquivalenten Bedingungen erfüllt ist:

- i)  $\sigma(FQ) = \sigma(F)$  für alle  $F \in \text{GL}^+(3)$  und  $Q \in \text{SO}(3)$ .
- ii)  $S_1(FQ) = S_1(F)Q^T$  für alle  $F \in \text{GL}^+(3)$  und  $Q \in \text{SO}(3)$ .
- iii)  $S_2(FQ) = Q^T S_2(F)Q$  für alle  $F \in \text{GL}^+(3)$  und  $Q \in \text{SO}(3)$ .
- iv)  $T^{\text{Biot}}(FQ) = Q^T T^{\text{Biot}}(F)Q$  für alle  $F \in \text{GL}^+(3)$  und  $Q \in \text{SO}(3)$ .

*Beweis.* Die Äquivalenz der Isotropie zu i) ergibt sich direkt aus der Definition. Wir zeigen noch, dass i), ii), iii) und iv) zueinander äquivalent sind:

Sei  $Q \in \text{SO}(3)$  und  $F \in \text{GL}^+(3)$  mit Polarzerlegung  $F = VR$  und  $V \in \text{Sym}^+(3)$  und  $R \in \text{SO}(3)$ , somit hat  $FQ \in \text{GL}^+(3)$  die Polarzerlegung  $FQ = V\tilde{R}$  mit gleichem  $V$  und  $\tilde{R} = RQ \in \text{SO}(3)$ . Zudem gilt  $\text{Cof}(FQ) = \text{Cof}(F)Q$  nach Folgerung 2.1.29. Nun können wir folgern

i)  $\Rightarrow$  ii):

$$S_1(FQ) = \sigma(FQ) \text{Cof}(FQ) = \sigma(F) \text{Cof}(F)Q = S_1(F)Q.$$

<sup>6</sup> $\Phi(Q^T X Q) = Q^T \Phi(X) Q$  für alle  $X \in \text{GL}^+(3)$  und alle  $Q \in \text{SO}(3)$

<sup>7</sup>So hat beispielsweise gemasertes Holz eine Vorzugsrichtung, Gummi oder Stahl hingegen nicht.

#### 4. Einführung in die nichtlineare Elastizitätstheorie

ii)  $\Rightarrow$  iii):

$$S_2(FQ) = (FQ)^{-1}S_1(FQ) = Q^T F^{-1}S_1(F)Q = Q^T S_2(F)Q.$$

iii)  $\Rightarrow$  iv):

$$T^{\text{Biot}}(FQ) = (RQ)^T(FQ)S_2(FQ) = Q^T R^T F S_2(F)Q = Q^T T^{\text{Biot}}(F)Q.$$

iv)  $\Rightarrow$  i):

$$\sigma(FQ) = (RQ)T^{\text{Biot}}(FQ)\text{Cof}(FQ)^{-1} = RT^{\text{Biot}}(F)\text{Cof}(F)^{-1} = \sigma(F). \quad \blacksquare$$

**Folgerung 4.3.19.** *Somit gilt für ein isotropes Elastizitätsgesetz, wenn wir zudem von Objektivität ausgehen, dass für die Spannungstensorabbildungen die mathematische Definition der Isotropie einer tensorwertigen Tensorfunktion gilt:*

$$\begin{aligned} \sigma(Q^T F Q) &= Q^T \sigma(F) Q, \\ S_1(Q^T F Q) &= Q^T S_1(F) Q, \\ S_2(Q^T F Q) &= Q^T S_2(F) Q, \\ T^{\text{Biot}}(Q^T F Q) &= Q^T T^{\text{Biot}}(F) Q \end{aligned} \quad (4.3.21)$$

für alle  $F \in \text{GL}^+(3)$  und  $Q \in \text{SO}(3)$ .

**Lemma 4.3.20.** *Sei  $\sigma: \text{GL}^+(3) \rightarrow \text{Sym}(3)$  ein isotropes Elastizitätsgesetz. Dann ist  $\sigma$  allein schon durch die Abbildungsvorschrift auf  $\text{Sym}^+(3)$  bestimmt, d.h. es existiert eine Abbildung  $\tilde{\sigma}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  mit  $\tilde{\sigma} = \sigma|_{\text{Sym}^+(3)}$  und  $\sigma(F) = \tilde{\sigma}(V)$  für alle  $F = VR \in \text{GL}^+(3)$  mit  $V \in \text{Sym}^+(3)$  und  $Q \in \text{SO}(3)$ .*

*Zudem ist  $\sigma$  koaxial und es existiert eine Abbildung  $\varrho: \mathbb{R}_+^3 \rightarrow \mathbb{R}_+^3$  mit  $(\lambda_1, \lambda_2, \lambda_3) \mapsto (\sigma_1, \sigma_2, \sigma_3)$ , welche die Eigenwerte von  $V$  auf die Eigenwerte von  $\sigma(V)$  abbildet. Ist  $V = Q \text{diag}(\lambda_1, \lambda_2, \lambda_3)Q^T$ , so ist*

$$\sigma(F) = \sigma(V) = Q \text{diag}(\sigma_1, \sigma_2, \sigma_3)Q^T. \quad (4.3.22)$$

*Damit sind die Hauptdehnungsachsen (die Eigenvektoren von  $V$  zu  $\lambda_1, \lambda_2, \lambda_3$  als Spalten in  $Q$ ) auch die Hauptspannungsachsen (die Eigenvektoren von  $\sigma(V)$  zu  $\sigma_1, \sigma_2, \sigma_3$  als Spalten in  $Q$ ). So werden  $\sigma_1, \sigma_2, \sigma_3$  auch Hauptspannungen (principal stresses) genannt.*

*Beweis.* Sei  $F \in \text{GL}^+(3)$  mit Polarzerlegung  $F = VR$  mit  $V \in \text{Sym}^+(3)$  und  $R \in \text{SO}(3)$  gegeben. Ist  $\sigma(V)$  bekannt, so gilt aufgrund der Isotropie  $\sigma(F) = \sigma(VR) = \sigma(V)$ , also ist auch  $\sigma(F)$  bekannt.

Zusammen mit der Objektivität folgt die mathematische Isotropie  $\sigma(QVQ^T) = Q\sigma(V)Q^T$  für alle  $V \in \text{Sym}^+(3)$  und alle  $Q \in \text{SO}(3)$  und Proposition 2.5.12 garantiert die Aussage.  $\blacksquare$

**Bemerkung 4.3.21.** Die Einschränkung auf  $\text{Sym}^+(3)$  werden wir im Folgenden anstatt mit  $\tilde{\sigma}$  auch mit  $\sigma: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  als Elastizitätsgesetz in  $V$  ausdrücken.

Zudem erhalten wir mit  $\hat{\sigma}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  und  $B \mapsto \hat{\sigma}(B) := \sigma(\sqrt{B})$  ein  $\sigma$  zugehöriges Elastizitätsgesetz in  $B = FF^T$ .

Zusammen mit Bemerkung 4.3.9 erhalten wir somit für das Elastizitätsgesetz Funktionen  $\hat{\sigma}, \hat{T}^{\text{Biot}}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  mit

$$\hat{\sigma}(Q^T B Q) = Q^T \hat{\sigma}(B) Q, \quad (4.3.23)$$

$$\hat{T}^{\text{Biot}}(Q^T C Q) = Q^T \hat{T}^{\text{Biot}}(C) Q, \quad (4.3.24)$$

d.h. mathematisch isotrope Elastizitätsgesetze in  $B$  und  $C$ . Nach Folgerung 2.5.13 kommutiert jede isotrope Tensorfunktion mit ihrem Argument, so können wir eine wichtige Eigenschaft sofort folgern.

#### 4. Einführung in die nichtlineare Elastizitätstheorie

**Folgerung 4.3.22.** Gegeben seien die Darstellungen eines isotropen Elastizitätsgesetzes  $\sigma, \hat{\sigma}, T^{\text{Biot}}, \hat{T}^{\text{Biot}} : \text{Sym}^+(3) \rightarrow \text{Sym}(3)$ . Dann kommutieren alle vier Abbildungen mit ihrem Argument, d.h.

$$\begin{aligned}\sigma(V) \cdot V &= V \cdot \sigma(V), \\ \hat{\sigma}(B) \cdot B &= B \cdot \hat{\sigma}(V), \\ T^{\text{Biot}}(U) \cdot U &= U \cdot T^{\text{Biot}}(U), \\ \hat{T}^{\text{Biot}}(C) \cdot C &= C \cdot \hat{T}^{\text{Biot}}(C)\end{aligned}\tag{4.3.25}$$

für alle  $V, U, C, B \in \text{Sym}^+(3)$ .

#### Isotropie von Energiefunktionen

Betrachten wir nun noch den Begriff der Isotropie von Energiefunktionen:

**Definition 4.3.23.** Wir bezeichnen eine Energiefunktion  $W : \text{GL}^+(3) \rightarrow \mathbb{R}$  als isotrop, wenn

$$W(FQ) = W(F) \quad \text{für alle } F \in \text{GL}^+(3) \text{ und } Q \in \text{SO}(3)\tag{4.3.26}$$

gilt.

**Bemerkung 4.3.24.** Ist  $F = VR = UR$  mit  $R \in \text{SO}(3)$  und  $U, V \in \text{Sym}^+(3)$  die Polarzerlegung des Deformationsgradienten  $F \in \text{GL}^+(3)$ , dann gilt mit Isotropie  $W(F) = W(V)$ . Da wir aber für gewöhnlich immer von Objektivität ausgehen können, ist es sehr ungewöhnlich, eine Energiefunktion in Abhängigkeit von  $V$  anzugeben, da ja bereits  $W(F) = W(U)$  gilt und eine Abbildungsvorschrift in  $V$ , falls sie denn aufgrund von Isotropie existiert, identisch mit der Vorschrift in  $U$  wäre.

**Bemerkung 4.3.25.** Eine objektive, isotrope Energiefunktion  $W : \text{GL}^+(3) \rightarrow \mathbb{R}$  erfüllt die klassische (mathematische) Definition der Isotropie einer skalarwertigen Tensorfunktion:

$$W(Q^T F Q) = W(F) \quad \text{für alle } F \in \text{GL}^+(3) \text{ und } Q \in \text{SO}(3).$$

**Lemma 4.3.26.** Sei  $W : \text{GL}^+(3) \rightarrow \mathbb{R}$  eine Energiefunktion. Ist  $W$  isotrop (und in diesem Fall nicht zwingend objektiv), so ist das durch  $W$  induzierte Elastizitätsgesetz isotrop.

*Beweis.* Der Beweis verläuft ähnlich wie der Beweis zu Lemma 4.3.14:

Unter Annahme von Isotropie gilt

$$W(FQ) = W(F) \quad \text{und} \quad W((F+H)Q) = W(F+H) \quad \text{für alle } F, H \in \text{GL}^+(3) \text{ und } Q \in \text{SO}(3).$$

Entwickeln wir  $W$  in  $FQ \in \text{GL}^+(3)$

$$W(FQ + \tilde{H}) = W(FQ) + \langle S_1(FQ), \tilde{H} \rangle + \mathcal{O}(\|\tilde{H}\|^2)$$

und setzen  $\tilde{H} = HQ$ , erhalten wir  $\mathcal{O}(\|\tilde{H}\|^2) = \mathcal{O}(\|H\|^2)$  und

$$\begin{aligned}W(F+H) &= W((F+H)Q) = W(FQ) + \langle S_1(FQ), HQ \rangle + \mathcal{O}(\|H\|^2) \\ &= W(F) + \langle S_1(QF)Q^T, H \rangle + \mathcal{O}(\|H\|^2),\end{aligned}$$

was mit (4.3.15) auf

$$\langle S_1(F) - S_1(FQ)Q^T, H \rangle = \mathcal{O}(\|H\|^2) \quad \text{für alle } H \in \text{GL}^+(3)$$

und damit auf  $S_1(F) = S_1(FQ)Q^T$  führt. ■

#### 4. Einführung in die nichtlineare Elastizitätstheorie

**Lemma 4.3.27.** Für alle  $F \in \text{GL}^+(3)$  mit  $B = FF^T = V^2$  gilt

$$\sigma(F) = \sigma(V) = \frac{1}{\det V} DW(V) V \quad (4.3.27)$$

$$= \hat{\sigma}(B) = \frac{2}{\sqrt{\det B}} D\widehat{W}(B) B. \quad (4.3.28)$$

*Beweis.* Wir erhalten (4.3.27) direkt über die Definition  $S_1(F) = DW(F)$  und  $\sigma(F) = \sigma(V)$ . Weiterhin ist mit  $W(V) = \widehat{W}(B)$  auch

$$\begin{aligned} \sigma(V) &= D_V \widehat{W}(V^2) (\det V)^{-1} V \\ &= D_B \widehat{W}(V^2) 2V (\det V)^{-1} V \\ &= \frac{2}{\sqrt{\det V^2}} D\widehat{W}(V^2) V^2, \end{aligned}$$

sodass mit  $\hat{\sigma}(B) = \sigma(V)$  auch (4.3.28) folgt. ■

**Lemma 4.3.28.** Sei  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  mit  $F \mapsto W(F)$  eine isotrope Energiefunktion. Dann existiert eine nur von den Singulärwerten von  $F$  abhängende Energiefunktion  $g: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  und eine nur von den Eigenwerten von  $B$  abhängende Energiefunktion  $\hat{g}: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  mit

$$W(F) = g(\lambda_1, \lambda_2, \lambda_3) = \hat{g}(\lambda_1^2, \lambda_2^2, \lambda_3^2). \quad (4.3.29)$$

Hierbei sind  $g$  und  $\hat{g}$  invariant gegenüber Permutation ihrer Parameter. Die Hauptspannungen gemäß Lemma 4.3.20, d.h. die Eigenwerte von  $\sigma(F)$ , lassen sich mittels  $g$  bzw.  $\hat{g}$  bestimmen. Für  $i \in \{1, 2, 3\}$  ist

$$\sigma_i = \frac{\lambda_i}{\lambda_1 \lambda_2 \lambda_3} \cdot \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) \quad (4.3.30)$$

$$= \frac{2 \lambda_i^2}{\sqrt{\lambda_1^2 \lambda_2^2 \lambda_3^2}} \cdot \frac{\partial \hat{g}}{\partial \lambda_i^2}(\lambda_1^2, \lambda_2^2, \lambda_3^2). \quad (4.3.31)$$

Die Singulärwerte von  $S_1(F)$  lassen sich ebenso mittels  $g$  bzw.  $\hat{g}$  bestimmen. Für  $i \in \{1, 2, 2\}$  ist

$$t_i = \frac{\lambda_1, \lambda_2, \lambda_3}{\lambda_i} \sigma_i = \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) \quad (4.3.32)$$

$$= 2 \lambda_i \frac{\partial \hat{g}}{\partial \lambda_i^2}(\lambda_1^2, \lambda_2^2, \lambda_3^2). \quad (4.3.33)$$

*Beweis.* Mit Lemma 2.5.26 erhalten wir aus der Isotropie von  $W$  die Existenz von  $g: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  und mit  $\hat{g}(x, y, z) = g(\sqrt{x}, \sqrt{y}, \sqrt{z})$  auch die Existenz von  $\hat{g}: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  mit  $W(F) = g(\lambda_1, \lambda_2, \lambda_3) = \hat{g}(\lambda_1^2, \lambda_2^2, \lambda_3^2)$  einschließlich der Invarianz gegenüber Argumentpermutation.

Seien nun  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_+$  beliebig gewählt. Wählen wir nun für den Deformationsgradienten die Diagonalmatrix  $F = (\text{diag}(\lambda_1, \lambda_2, \lambda_3))$ , dann folgt mit Lemma 4.3.27

$$\sigma(\text{diag}(\lambda_1, \lambda_2, \lambda_3)) = \frac{1}{\lambda_1 \lambda_2 \lambda_3} DW(\text{diag}(\lambda_1, \lambda_2, \lambda_3)) \text{diag}(\lambda_1, \lambda_2, \lambda_3).$$

Mit Lemma 4.3.20 ist  $\sigma(\text{diag}(\lambda_1, \lambda_2, \lambda_3)) = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$ , also ist auch  $DW(\text{diag}(\lambda_1, \lambda_2, \lambda_3))$  eine Diagonalmatrix und es gilt

$$DW(\text{diag}(\lambda_1, \lambda_2, \lambda_3)) = \text{diag}\left(\frac{\partial g}{\partial \lambda_1}(\lambda_1, \lambda_2, \lambda_3), \frac{\partial g}{\partial \lambda_2}(\lambda_1, \lambda_2, \lambda_3), \frac{\partial g}{\partial \lambda_3}(\lambda_1, \lambda_2, \lambda_3)\right).$$

Betrachten wir die Diagonalelemente der zusammengesetzten Gleichung, so erhalten wir  $\sigma_i = \frac{\lambda_i}{\lambda_1 \lambda_2 \lambda_3} \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3)$ .

Analog ist

$$D\widehat{W}(\text{diag}(\lambda_1^2, \lambda_2^2, \lambda_3^2)) = \text{diag}\left(\frac{\partial \widehat{g}}{\partial \lambda_1^2}(\lambda_1^2, \lambda_2^2, \lambda_3^2), \frac{\partial \widehat{g}}{\partial \lambda_2^2}(\lambda_1^2, \lambda_2^2, \lambda_3^2), \frac{\partial \widehat{g}}{\partial \lambda_3^2}(\lambda_1^2, \lambda_2^2, \lambda_3^2)\right).$$

und wir erhalten

$$\text{diag}(\sigma_1, \sigma_2, \sigma_3) = \frac{2}{\sqrt{\lambda_1^2 \lambda_2^2 \lambda_3^2}} D\widehat{W}(\text{diag}(\lambda_1^2, \lambda_2^2, \lambda_3^2)) \text{diag}(\lambda_1^2, \lambda_2^2, \lambda_3^2)$$

und somit bei Betrachtung der Diagonalelemente  $\sigma_i = \frac{2\lambda_i^2}{\sqrt{\lambda_1^2 \lambda_2^2 \lambda_3^2}} \cdot \frac{\partial \widehat{g}}{\partial \lambda_i^2}(\lambda_1^2, \lambda_2^2, \lambda_3^2)$ .

Zudem ist  $S_1(\text{diag}(\lambda_1, \lambda_2, \lambda_3)) = \sigma(\text{diag}(\lambda_1, \lambda_2, \lambda_3)) \cdot \text{Cof}(\text{diag}(\lambda_1, \lambda_2, \lambda_3))$ , und damit  $t_i = \sigma_i \frac{\lambda_1 \lambda_2 \lambda_3}{\lambda_i}$ . Die  $t_i$  sind die Singulärwerte von  $S_1(F)$ , denn für  $F = RU = VR$  und  $R \in \text{SO}(3)$  und  $V, U \in \text{Sym}^+(3)$  gilt

$$\begin{aligned} S_1(F) &= S_1(VR) = S_1(V) R^T \\ &= S_1(RU) = R S_1(U). \end{aligned} \quad \blacksquare$$

Verwenden wir die Nebenbedingung  $\det F = 1$ , gehen also von inkompressiblem Material aus, so nehmen  $\sigma_i$  und  $t_i$  aus Lemma 4.3.28 die Form

$$\sigma_i = \lambda_i \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) + p, \quad (4.3.34)$$

$$t_i = \frac{1}{\lambda_i} \sigma_i = \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) + p \frac{1}{\lambda_i} \quad (4.3.35)$$

an. Hierbei ist  $p$  ein Lagrange-Multiplikator, der auch *hydrostatischer Druck* genannt wird.

### 4.3.9. Einige einfache Beispieldeformationen für inkompressible Materialien

Im Folgenden betrachten wir für den Fall, dass das Elastizitätsgesetz durch eine isotrope Energiefunktion  $W(F) = g(\lambda_1, \lambda_2, \lambda_3)$  induziert wird, die Hauptspannungen von drei Typen von volumenerhaltenden Deformationen: Im uniaxialen Fall wird ein Referenzwürfel in  $e_1$ -Richtung gedehnt (siehe Abbildung 4.8), im equibiaxialen Fall wird er zugleich in  $e_1$ -Richtung und in  $e_2$ -Richtung gedehnt (siehe Abbildung 4.9), im Pure-Shear-Fall wird er in  $e_1$ -Richtung gedehnt und in  $e_2$ -Richtung konstant gehalten (siehe Abbildung 4.10). In allen drei Fällen gleichen die übrigen Richtungen die Deformation volumenerhaltend aus.

#### Uniaxiale Deformation

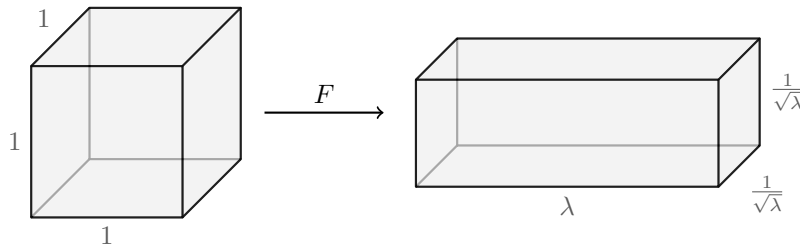


Abbildung 4.8.: Uniaxiale Deformation

**Proposition 4.3.29** (Uniaxiale Deformation). Für  $\lambda \in \mathbb{R}_+$  seien  $\lambda_1 = \lambda$  und  $\lambda_2 = \lambda_3 = \frac{1}{\sqrt{\lambda}}$ . Zudem werde  $\sigma_2 = \sigma_3 = 0$  bzw.  $t_2 = t_3 = 0$  angenommen. Dann gilt

$$\sigma_1 = \lambda \frac{dg}{d\lambda}\left(\lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}\right) \quad \text{und} \quad t_1 = \frac{dg}{d\lambda}\left(\lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}\right). \quad (4.3.36)$$

#### 4. Einführung in die nichtlineare Elastizitätstheorie

*Beweis.* Im Fall von  $i \in \{2, 3\}$  ergibt (4.3.35)

$$\frac{\partial g}{\partial \lambda_2} \left( \lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}} \right) = \frac{\partial g}{\partial \lambda_3} \left( \lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}} \right) = -p\sqrt{\lambda},$$

also

$$p = -\frac{1}{2\sqrt{\lambda}} \frac{\partial g}{\partial \lambda_2} \left( \lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}} \right) - \frac{1}{2\sqrt{\lambda}} \frac{\partial g}{\partial \lambda_3} \left( \lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}} \right).$$

Wenden wir die Kettenregel an, so ist (4.3.35) für  $i = 3$

$$\begin{aligned} \frac{dg}{d\lambda} \left( \lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}} \right) &= \frac{\partial g}{\partial \lambda_1} \left( \lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}} \right) - \frac{1}{2\lambda^{3/2}} \frac{\partial g}{\partial \lambda_2} \left( \lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}} \right) - \frac{1}{2\lambda^{3/2}} \frac{\partial g}{\partial \lambda_3} \left( \lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}} \right) \\ &= \frac{\partial g}{\partial \lambda_1} \left( \lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}} \right) + \frac{1}{\lambda} p = t_3. \end{aligned}$$

Mit  $\sigma_i = \lambda_i t_i$  gilt zudem die Aussage für die erste Hauptspannung. ■

#### Equibiaxiale Deformation

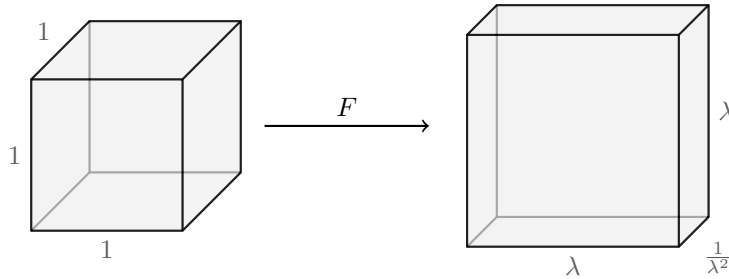


Abbildung 4.9.: Equibiaxiale Deformation

**Proposition 4.3.30** (Equibiaxiale Deformation). Für  $\lambda \in \mathbb{R}_+$  seien  $\lambda_1 = \lambda_2 = \lambda$  und  $\lambda_3 = \frac{1}{\lambda^2}$ . Zudem werde  $\sigma_3 = 0$  bzw.  $t_3 = 0$  angenommen. Dann gilt

$$\sigma_1 = \sigma_2 = \frac{\lambda}{2} \frac{dg}{d\lambda} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) \quad \text{und} \quad t_1 = t_2 = \frac{1}{2} \frac{dg}{d\lambda} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right). \quad (4.3.37)$$

*Beweis.* Im Fall von  $i = 3$  ergibt (4.3.35)

$$\frac{\partial g}{\partial \lambda_3} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) = -p\lambda^2.$$

Zudem gilt aus Symmetriegründen<sup>8</sup>

$$\frac{\partial g}{\partial \lambda_1} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) = \frac{\partial g}{\partial \lambda_2} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right)$$

annehmen. Wenden wir die Kettenregel an, so ist (4.3.35) für  $i \in \{1, 2\}$

$$\begin{aligned} \frac{dg}{d\lambda} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) &= \frac{\partial g}{\partial \lambda_1} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) + \frac{\partial g}{\partial \lambda_2} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) - \frac{\partial g}{\partial \lambda_3} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) \cdot \frac{2}{\lambda^3} \\ &= 2 \frac{\partial g}{\partial \lambda_1} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) + p \frac{2}{\lambda} = 2t_1 = 2t_2. \end{aligned}$$

Mit  $\sigma_i = \lambda_i t_i$  gilt zudem die Aussage für die erste Hauptspannung. ■

<sup>8</sup>genauer: aufgrund der Invarianz von  $g$  gegenüber Argumentpermutation



**Pure-Shear-Deformation**

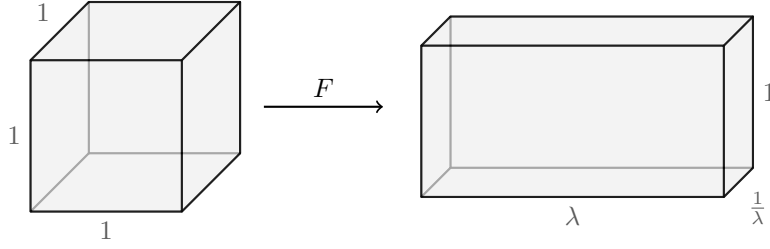


Abbildung 4.10.: Pure-Shear-Deformation

**Proposition 4.3.31** (Pure-Shear-Deformation). Für  $\lambda \in \mathbb{R}_+$  seien  $\lambda_1 = \lambda$ ,  $\lambda_2 = 1$  und  $\lambda_3 = \frac{1}{\lambda}$ . Zudem werde  $\sigma_3 = 0$  bzw.  $t_3 = 0$  angenommen. Dann gilt

$$\sigma_1 = \lambda \frac{dg}{d\lambda} \left( \lambda, 1, \frac{1}{\lambda} \right) \quad \text{und} \quad t_1 = \frac{dg}{d\lambda} \left( \lambda, 1, \frac{1}{\lambda} \right). \quad (4.3.38)$$

*Beweis.* Im Fall von  $i = 3$  ergibt (4.3.35)

$$\frac{\partial g}{\partial \lambda_3} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) = -p \lambda.$$

Wenden wir die Kettenregel an, so ist (4.3.35) für  $i = 1$

$$\begin{aligned} \frac{dg}{d\lambda} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) &= \frac{\partial g}{\partial \lambda_1} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) - \frac{1}{\lambda^2} \frac{\partial g}{\partial \lambda_3} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) \\ &= \frac{\partial g}{\partial \lambda_1} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) + p \frac{1}{\lambda} = t_1. \end{aligned}$$

Mit  $\sigma_i = \lambda_i t_i$  gilt zudem die Aussage für die erste Hauptspannung. ■

**4.3.10. Verschiedene Energiefunktionen und ihr Einfluss auf uniaxiale, biaxiale und Pure-Shear-Deformation**

Wie in Abschnitt 4.3.6 erläutert, wird jedes hyperelastische isotrope Elastizitätsgesetz durch eine isotrope Energiefunktion bestimmt. Wir wollen nun isotrope Energiefunktionen darauf untersuchen, wie gut sich der durch sie hergestellte Dehnungs-Spannungs-Zusammenhang mit realen, physikalischen Messdaten deckt. Hierzu verwenden wir Daten, die Treloar [77] 1944 bei Experimenten zur uniaxialen Streckung und Stauchung (vgl. Proposition 4.3.29), zur equibiaxialen Streckung (vgl. Proposition 4.3.30) und zur Pure-Shear-Dehnung (vgl. Proposition 4.3.31) ermittelt hat.

Zum Vergleich verwenden wir als Energiefunktionen die *Saint-Venant-Kirchhoff-Energie*

$$\begin{aligned} W_{\text{SVK}}(F) &= \frac{\mu}{4} \|F^T F - \mathbb{1}\|^2 + \frac{\Lambda}{8} (\text{tr}(F^T F - \mathbb{1}))^2 \\ &= g_{\text{NH}}(\lambda_1, \lambda_2, \lambda_3) = \frac{\mu}{4} ((\lambda_1^2 - 1)^2 + (\lambda_2^2 - 1)^2 + (\lambda_3^2 - 1)^2) + \frac{\Lambda}{8} (\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3)^2 \end{aligned} \quad (4.3.39)$$

mit  $\mu \in \mathbb{R}_+$ ,  $\Lambda \in \mathbb{R}$  mit  $3\Lambda + 2\mu > 0$ , die inkompressible<sup>9</sup> *Neo-Hooke-Energie*

$$W_{\text{NH}}(F) = \frac{\mu}{2} (\|F\|^2 - 3) = g_{\text{NH}}(\lambda_1, \lambda_2, \lambda_3) = \frac{\mu}{2} (\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3) \quad (4.3.40)$$

<sup>9</sup>Inkompressibel heißt, dass  $\det F = 1$ , also  $F \in \text{SL}(3)$  vorausgesetzt wird.

#### 4. Einführung in die nichtlineare Elastizitätstheorie

mit  $\mu \in \mathbb{R}_+$ , sowie die inkompressible quadratische *Hencky-Energie*

$$W_{\text{H}}(F) = \mu \|\log U\|^2 = g_{\text{H}}(\lambda_1, \lambda_2, \lambda_3) = \mu((\log \lambda_1)^2 + (\log \lambda_2)^2 + (\log \lambda_3)^2) \quad (4.3.41)$$

mit  $\mu \in \mathbb{R}_+$ .

Jones und Treloar [40] benutzen selbst eine inkompressible Form einer Valanis-Landel-Energie (siehe auch Definition 5.2.1 auf Seite 127)

$$W_{\text{JT}}(F) = g_{\text{JT}}(\lambda_1, \lambda_2, \lambda_3) = w(\lambda_1) + w(\lambda_2) + w(\lambda_3) \quad (4.3.42)$$

mit

$$w'(\lambda) = 0.69\left(\lambda^{0.3} - \frac{1}{\lambda}\right) + 0.01\left(\lambda^3 - \frac{1}{\lambda}\right) - 0.0122\left(\lambda^{-3} - \frac{1}{\lambda}\right) + \frac{c}{\lambda} \quad (4.3.43)$$

und dem Parameter  $c = w'(1)$ . Hierbei sei angemerkt, dass diese *Jones-Treloar-Energie* darauf angepasst ist, die Messwerten von Treloar [77] möglichst gut zu approximieren und auf anderen Messdaten im Allgemeinen keine guten Ergebnisse liefert. Für beliebige Messdaten lässt sich die von Ogden [61] vorgeschlagene Energie

$$W_{\text{OG}}(F) = g_{\text{OG}}(\lambda_1, \lambda_2, \lambda_3) = \sum_{i=1}^M \frac{\mu_i}{\alpha_i} (\lambda_1^{\alpha_i} + \lambda_2^{\alpha_i} + \lambda_3^{\alpha_i} - 3) \quad (4.3.44)$$

mit Parametern  $\alpha_1, \dots, \alpha_M$  und  $\mu_1, \dots, \mu_M$  verwenden. Mit ihr lässt sich eine beliebig gute Annäherung<sup>10</sup> bei hinreichend großem  $M$  erreichen. Allerdings lernt auch die *Ogden-Energie* vielmehr die Testdaten, als dass sie einen physikalischen Dehnungs-Spannungs-Zusammenhang im Sinne der Elastizität aufbaut.<sup>11</sup>

Die folgenden Diagramme basieren auf denen von Martin [46], der ebenso verschiedene Energien bezüglich der Messdaten von Treloar vergleicht. Wir stauchen/strecken in allen drei Fällen den Einheitswürfel zu einem Quader, dabei in  $e_1$ -Richtung auf Breite  $\lambda$ , unter Einwirkung von Kräften auf alle sechs Würfelflächen, insbesondere auf die Fläche mit Normale  $e_1$ , d.h. die rechte Seitenfläche, mit Kraft  $t_1$ .<sup>12</sup>

#### Uniaxiale Deformation

Für eine uniaxiale Deformation  $F \in \text{GL}^+(3)$  mit Singulärwerten  $(\lambda_1, \lambda_2, \lambda_3) = \left(\lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}\right)$ , wie in Abbildung 4.8 zu sehen, verhält sich die Kraft  $t_1$  in Abhängigkeit vom Dehnungsfaktor in  $e_1$ -Richtung  $\lambda \in \mathbb{R}_+$  ( $\lambda < 1$  für Stauchung,  $\lambda > 1$  für Streckung) wie folgt:

Ogden:

$$t_1 = \frac{d}{d\lambda} g_{\text{O}}\left(\lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}\right) = \frac{d}{d\lambda} \sum_{i=1}^M \frac{\mu_i}{\alpha_i} (\lambda^{\alpha_i} + 2\lambda^{-\frac{\alpha_i}{2}} - 3) = \sum_{i=1}^M \mu_i (\lambda^{\alpha_i-1} - \lambda^{-\frac{\alpha_i}{2}-1}), \quad (4.3.45)$$

Jones-Treloar:

$$\begin{aligned} t_1 &= \frac{d}{d\lambda} g_{\text{JT}}\left(\lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}\right) = \frac{d}{d\lambda} (w(\lambda) + 2w(\frac{1}{\sqrt{\lambda}})) = w'(\lambda) - \frac{1}{\lambda^{3/2}} w'(\frac{1}{\sqrt{\lambda}}) \\ &= 0.69(\lambda^{0.3} - \lambda^{-1.65}) + 0.01(\lambda^3 - \lambda^{-3}) + 0.0122(1 - \lambda^{-3}), \end{aligned} \quad (4.3.46)$$

Saint-Venant-Kirchhoff:

$$t_1 = \frac{d}{d\lambda} g_{\text{SVK}}\left(\lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}\right) = \frac{d}{d\lambda} \frac{\mu}{4} \left( (\lambda^2 - 1)^2 + 2\left(\frac{1}{\lambda} - 1\right)^2 \right) + \frac{\Lambda}{8} \left( \lambda^2 + \frac{2}{\lambda} - 3 \right)^2$$

<sup>10</sup>ähnlich wie bei der Approximation einer analytischen Funktion durch ein Taylorpolynom hinreichend großer Ordnung

<sup>11</sup>Das typische Szenario ist: Mit Testdaten für einfache Beispieldeformationen werden die Parameter einer Energie bestimmt, um dann für beliebige Deformationen realistische Aussagen zum Dehnungs-Spannungs-Zusammenhang zu treffen. Hierfür eignet sich die Ogden-Energie weniger gut als andere verwendete Energien.

<sup>12</sup>Kraft pro Flächeneinheit im Referenzgebiet bezogen auf Fläche mit Flächeneinheit 1 im Referenzgebiet.

#### 4. Einführung in die nichtlineare Elastizitätstheorie

$$= \left(\mu + \frac{\Lambda}{2}\right)\left(\lambda^3 - \frac{1}{\lambda^3}\right) - \left(\mu + \frac{3}{2}\Lambda\right)\left(\lambda - \frac{1}{\lambda^2}\right), \quad (4.3.47)$$

Neo-Hooke:

$$t_1 = \frac{d}{d\lambda} g_{\text{NH}}\left(\lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}\right) = \frac{d}{d\lambda} \frac{\mu}{2} \left(\lambda^2 + \frac{2}{\lambda} - 3\right) = \mu \left(\lambda - \frac{1}{\lambda^2}\right), \quad (4.3.48)$$

Hencky:

$$t_1 = \frac{d}{d\lambda} g_{\text{H}}\left(\lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}\right) = \frac{d}{d\lambda} \frac{3\mu}{2} (\log \lambda)^2 = \frac{3\mu \log \lambda}{\lambda}. \quad (4.3.49)$$

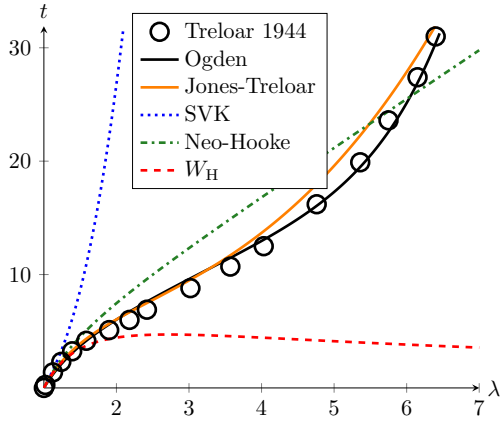


Abbildung 4.11.: Uniaxiale Streckung

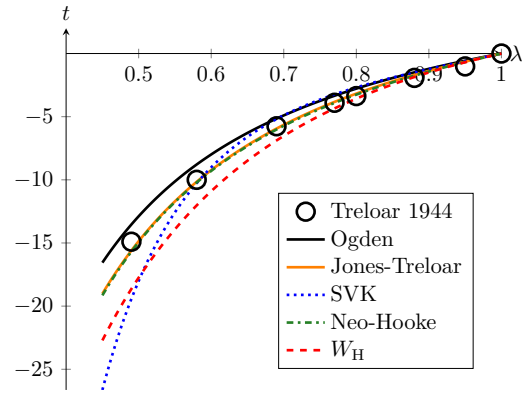


Abbildung 4.12.: Uniaxiale Stauchung

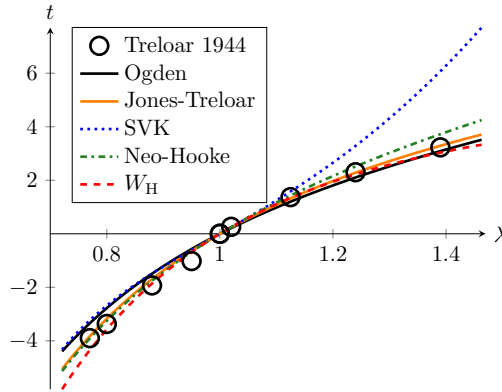


Abbildung 4.13.: Uniaxiale Dehnung im Bereich kleiner Verzerrung

#### Equibiaxiale Deformation

Für eine equibiaxiale Deformation  $F \in \text{GL}^+(3)$  mit Singulärwerten  $(\lambda_1, \lambda_2, \lambda_3) = \left(\lambda, \lambda, \frac{1}{\lambda^2}\right)$ , wie in Abbildung 4.9 zu sehen, verhält sich die Kraft  $t_1$  in Abhängigkeit vom Dehnungsfaktor in  $e_1$ -Richtung  $\lambda \in \mathbb{R}_+$  wie folgt:

Ogden:

$$t_1 = t_2 = \frac{d}{d\lambda} g_{\text{O}}\left(\lambda, \lambda, \frac{1}{\lambda^2}\right) = \frac{1}{2} \frac{d}{d\lambda} \sum_{i=1}^M \frac{\mu_i}{\alpha_i} (2\lambda^{\alpha_i} + \lambda^{-2\alpha_i} - 3) = \sum_{i=1}^M \mu_i (\lambda^{\alpha_i-1} - \lambda^{-2\alpha_i-1}), \quad (4.3.50)$$

#### 4. Einführung in die nichtlineare Elastizitätstheorie

Jones-Treloar:

$$\begin{aligned} t_1 = t_2 &= \frac{d}{d\lambda} g_{JT} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) = \frac{1}{2} \frac{d}{d\lambda} \left( 2w(\lambda) + w\left(\frac{1}{\lambda^2}\right) \right) = w'(\lambda) - \frac{1}{\lambda^3} w'\left(\frac{1}{\lambda^2}\right) \\ &= 0.69(\lambda^{0.3} - \lambda^{-3.6}) + 0.01(\lambda^3 - \lambda^{-9}) + 0.0122(\lambda^3 - \lambda^{-3}), \end{aligned} \quad (4.3.51)$$

Saint-Venant-Kirchhoff:

$$\begin{aligned} t_1 = t_2 &= \frac{d}{d\lambda} g_{SVK} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) = \frac{1}{2} \frac{d}{d\lambda} \frac{\mu}{4} \left( 2(\lambda^2 - 1)^2 + \left(\frac{1}{\lambda^4} - 1\right)^2 \right) + \frac{\Lambda}{8} \left( 2\lambda^2 + \frac{1}{\lambda^4} - 3 \right)^2 \\ &= \mu \left( \lambda^3 - \lambda + \frac{1}{\lambda^5} - \frac{1}{\lambda^9} \right) + \frac{\Lambda}{2} \left( 2\lambda^3 - 3\lambda - \frac{1}{\lambda^3} + \frac{3}{\lambda^5} - \frac{1}{\lambda^9} \right), \end{aligned} \quad (4.3.52)$$

Neo-Hooke:

$$t_1 = t_2 = \frac{d}{d\lambda} g_{NH} \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) = \frac{1}{2} \frac{d}{d\lambda} \frac{\mu}{2} \left( 2\lambda^2 + \frac{1}{\lambda^4} - 3 \right) = \mu \left( \lambda - \frac{1}{\lambda^5} \right), \quad (4.3.53)$$

Hencky:

$$t_1 = t_2 = \frac{d}{d\lambda} g_H \left( \lambda, \lambda, \frac{1}{\lambda^2} \right) = \frac{1}{2} \frac{d}{d\lambda} 6\mu (\log \lambda)^2 = \frac{6\mu \log \lambda}{\lambda}. \quad (4.3.54)$$

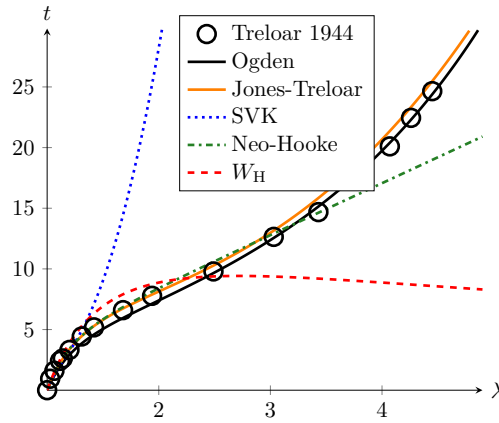


Abbildung 4.14.: Equibiaxiale Streckung

#### Pure-Shear-Deformation

Für eine Pure-Shear-Deformation  $F \in GL^+(3)$  mit Singulärwerten  $(\lambda_1, \lambda_2, \lambda_3) = \left( \lambda, \lambda, \frac{1}{\lambda^2} \right)$ , wie in Abbildung 4.10 zu sehen, verhält sich die Kraft  $t_1$  in Abhängigkeit vom Dehnungsfaktor in  $e_1$ -Richtung  $\lambda \in \mathbb{R}_+$  wie folgt:

Ogden:

$$t_1 = \frac{d}{d\lambda} g_O \left( \lambda, 1, \frac{1}{\lambda} \right) = \frac{d}{d\lambda} \sum_{i=1}^M \frac{\mu_i}{\alpha_i} (\lambda^{\alpha_i} + \lambda^{-\alpha_i} - 2) = \sum_{i=1}^M \mu_i (\lambda^{\alpha_i-1} - \lambda^{-\alpha_i-1}), \quad (4.3.55)$$

Jones-Treloar:

$$\begin{aligned} t_1 &= \frac{d}{d\lambda} g_{JT} \left( \lambda, 1, \frac{1}{\lambda} \right) = \frac{d}{d\lambda} \left( w(\lambda) + w(1) + w\left(\frac{1}{\lambda}\right) \right) = w'(\lambda) - \frac{1}{\lambda^2} w'\left(\frac{1}{\lambda}\right) \\ &= 0.69(\lambda^{0.3} - \lambda^{-2.3}) + 0.01(\lambda^3 - \lambda^{-5}) + 0.0122(\lambda^1 - \lambda^{-3}), \end{aligned} \quad (4.3.56)$$

#### 4. Einführung in die nichtlineare Elastizitätstheorie

Saint-Venant-Kirchhoff:

$$\begin{aligned} t_1 &= \frac{d}{d\lambda} g_{\text{SVK}}\left(\lambda, 1, \frac{1}{\lambda}\right) = \frac{d}{d\lambda} \frac{\mu}{4} \left( (\lambda^2 - 1)^2 + \left(\frac{1}{\lambda^2} - 1\right)^2 \right) + \frac{\Lambda}{8} \left( \lambda^2 + \frac{1}{\lambda^2} - 2 \right)^2 \\ &= \left( \mu + \frac{\Lambda}{2} \right) \left( \lambda^3 - \frac{1}{\lambda^3} \right) - (\mu + \Lambda) \left( \lambda - \frac{1}{\lambda^3} \right), \end{aligned} \quad (4.3.57)$$

Neo-Hooke:

$$t_1 = \frac{d}{d\lambda} g_{\text{NH}}\left(\lambda, 1, \frac{1}{\lambda}\right) = \frac{d}{d\lambda} \frac{\mu}{2} \left( \lambda^2 + \frac{1}{\lambda^2} - 2 \right) = \mu \left( \lambda - \frac{1}{\lambda^3} \right), \quad (4.3.58)$$

Hencky:

$$t_1 = \frac{d}{d\lambda} g_{\text{H}}\left(\lambda, 1, \frac{1}{\lambda}\right) = \frac{d}{d\lambda} 2\mu (\log \lambda)^2 = \frac{4\mu \log \lambda}{\lambda}. \quad (4.3.59)$$

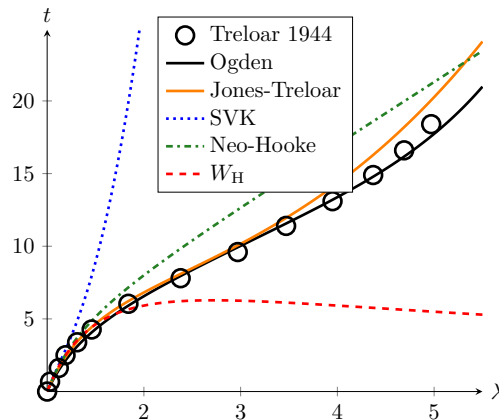


Abbildung 4.15.: Pure-Shear-Dehnung

#### 4.3.11. Darstellung isotroper Energiefunktionen als Funktion in den Invarianten

In Lemma 2.13.25 aus Seite 66 haben wir gezeigt, dass wir jede isotrope (unendlich oft differenzierbare) Funktion  $W: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit  $U \mapsto W(U)$  auch als (unendlich oft differenzierbare) Funktion  $\Psi: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  in den Invarianten  $I_1, I_2, I_3$  von  $U$  schreiben können. Diese Invarianten sind hierbei die Koeffizienten des charakteristischen Polynoms von  $U$

$$\chi_U(x) = x^3 - I_1 x^2 + I_2 x - I_3$$

und damit die elementarsymmetrischen Polynome der Eigenwerte  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_+$  von  $U$ . Wegen der Isotropie können wir hierbei auch  $W: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit  $U \mapsto W(U)$  durch  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  mit  $F \mapsto W(F)$  ersetzen, wobei  $I_1, I_2, I_3$  die Invarianten von  $U = \sqrt{F^T F}$  und  $\lambda_1, \lambda_2, \lambda_3$  die Singulärwerte von  $F$  sind. Aufgrund der Bijektivität der Abbildung  $U \mapsto C = U^2$  bzw.  $F \mapsto C = F^T F$  können wir die Isotrope Funktion  $W$  ebenso als Funktion  $\hat{\Psi}: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  in den Invarianten  $I_1, I_2, I_3$  von  $C$  schreiben. Für die Invarianten gilt

$$I_1 = \text{tr}(C), \quad I_2 = \text{tr}(\text{Cof } C), \quad I_3 = \det C.$$

Nach dem Satz von Cayley-Hamilton ist zudem

$$C^3 - I_1 C^2 + I_2 C - I_3 = 0, \quad (4.3.60)$$

woraus sich die folgenden nützlichen Formeln ergeben.

**Proposition 4.3.32.** Sei  $F \in \text{GL}^+(3)$  und  $C = F^T F$ , dann gilt für die Invarianten  $I_1, I_2, I_3$  von  $C$

$$\|F\|^2 = \text{tr}(C) = I_1, \quad (4.3.61)$$

$$\|\text{Cof } F\|^2 = \text{tr}(\text{Cof } C) = I_2, \quad (4.3.62)$$

$$\|C\|^2 = I_1^2 - 2I_2, \quad (4.3.63)$$

$$\|C^{-1}\|^2 = \frac{I_2^2}{I_3^2} - 2\frac{I_1}{I_3}. \quad (4.3.64)$$

**Bemerkung 4.3.33.** Stellen wir (4.3.63) um, erhalten wir sofort die bereits verwendete Identität  $I_2 = \frac{1}{2} \text{tr}(C)^2 - \frac{1}{2} \text{tr}(C^2)$ .

*Beweis.* (4.3.61) ist klar, (4.3.62) ergibt sich über den Zusammenhang  $\text{Cof}(F^T F) = (\text{Cof } F)^T (\text{Cof } F)$ . Multiplizieren wir (4.3.60) mit  $C^{-1}$  bzw. mit  $C^{-2}$ , dann erhalten wir

$$C^2 - I_1 C + I_2 \cdot \mathbb{1} - I_3 C^{-1} = 0 \quad \text{bzw.} \quad C - I_1 \cdot \mathbb{1} + I_2 C^{-1} - I_3 C^{-2} = 0.$$

Wenden wir den Spuroperator an, nutzen seine Linearität und den Umstand, dass  $\text{Cof } C = I_3 \cdot C^{-1}$  gilt, so erhalten wir

$$\|C\|^2 - I_1^2 + 3I_2 - I_2 = 0 \quad \text{bzw.} \quad I_1 - 3I_1 + \frac{I_2^2}{I_3} - I_3 \|C^{-1}\|^2 = 0.$$

Umstellen liefert nun (4.3.63) bzw. (4.3.64). ■

### 4.3.12. Verzerrung

#### Linearer Fall

Sei  $F \in \text{GL}^+(3)$  mit Polarzerlegung  $F = RU$  und  $R \in \text{SO}(3)$  und  $U \in \text{Sym}^+(3)$ . Zudem sei  $\nabla u = F - \mathbb{1}$ . In (4.3.74) zeigen wir (setzen wir  $m = 1$ )

$$U = \mathbb{1} + \text{sym } \nabla u + \mathcal{O}(\|\nabla u\|^2). \quad (4.3.65)$$

In Folgerung 2.3.2 auf Seite 26 haben wir gezeigt, dass es für jedes  $Q \in \text{SO}(n)$  ein  $W \in \mathfrak{so}(n)$  gibt, sodass  $Q = \mathbb{1} + W + \mathcal{O}(\|Q - \mathbb{1}\|^2)$  gilt. Aufgrund der stetigen Differenzierbarkeit der Abbildung  $F \mapsto F\sqrt{F^T F} = R$  auf  $\text{GL}^+(3)$  ist die Abbildung auch lokal lipschitz-stetig, somit erhalten wir für eine Konstante  $c \in \mathbb{R}$  die Abschätzung  $\|R - \mathbb{1}\|^2 \leq c^2 \|F - \mathbb{1}\|^2 = c^2 \|\nabla u\|^2$  und können auf

$$R = \mathbb{1} + W + \mathcal{O}(\|\nabla u\|^2) \quad \text{mit } W \in \mathfrak{so}(n) \quad (4.3.66)$$

schließen. Nun gilt einerseits

$$F = \mathbb{1} + \nabla u = \mathbb{1} + \text{sym } \nabla u + \text{skew } \nabla u, \quad (4.3.67)$$

andererseits liefert das Produkt der Polarfaktoren

$$\begin{aligned} F &= R \cdot U = (\mathbb{1} + W + \mathcal{O}(\|\nabla u\|^2)) \cdot (\mathbb{1} + \text{sym } \nabla u + \mathcal{O}(\|\nabla u\|^2)) \\ &= \mathbb{1} + \text{sym } \nabla u + W + \mathcal{O}(\|\nabla u\|^2), \end{aligned} \quad (4.3.68)$$

woraus wir  $W = \text{skew } \nabla u + \mathcal{O}(\|\nabla u\|^2)$  und mit der Linearität von skew schlussendlich  $W = \text{skew } \nabla u$  folgern können.

Wir definieren den *linearen Verzerrungstensor*  $\varepsilon := \text{sym } \nabla u$  und erhalten so

$$F = \mathbb{1} + \underbrace{\varepsilon}_{\text{lin. Verzerr.}} + \underbrace{W}_{\text{lin. Rotation}} \quad (4.3.69)$$

#### 4. Einführung in die nichtlineare Elastizitätstheorie

mit  $\varepsilon = \text{sym } \nabla u \in \text{Sym}(3)$  und  $W = \text{skew } \nabla u \in \mathfrak{so}(3)$ . Hierbei ist  $W$  die Linearisierung der Rotation<sup>13</sup>: Für infinitesimal kleine Rotationen lässt sich  $F$  somit **additiv** in Verzerrung und Rotation zerlegen. Wir vergleichen dies mit der nichtlinearen **multiplikativen** (Polar-)Zerlegung

$$F = \underbrace{R}_{\text{Rotation}} \cdot \underbrace{U}_{\text{Dehnung}} \quad (4.3.70)$$

mit Dehnung  $U \in \text{Sym}^+(3)$  und Rotation  $R \in \text{SO}(3)$ .

Zudem sei angemerkt, ist  $F \in \text{SO}(3)$  eine reine Rotation, so ist die Dehnung eins und die (lineare) Verzerrung null.

Wir untersuchen die Abweichung von  $F$  zur linearen Rotation: Für  $Q \in \text{SO}(n)$  ist  $Q = \mathbb{1} + W + O(\|Q - \mathbb{1}\|^2)$  mit  $W \in \mathfrak{so}(n)$ , so linearisiert sich  $Q$  zu  $\mathbb{1} + W$ . Wegen  $F = \mathbb{1} + \nabla u$ , ist die linearisierte Dehnung die Abweichung von  $\nabla u$  zu  $\mathfrak{so}(n)$ . Betrachten wir die Seth-Hill-Verzerrungsfamilie. Wegen  $(\mathbb{1} + H)^n = \mathbb{1} + n \cdot H + O(\|H\|^2)$  ist

$$\begin{aligned} \frac{1}{n}(U^n - \mathbb{1}) &= \frac{1}{n}((F^T F)^{\frac{n}{2}} - \mathbb{1}) \\ &= \frac{1}{n}((\mathbb{1} + 2 \text{sym } \nabla u + \nabla u^T \nabla u)^{\frac{n}{2}} - \mathbb{1}) \\ &= \frac{1}{n}(\mathbb{1} + n \text{sym } \nabla u + O(\|\nabla u\|^2) - \mathbb{1}) \\ &= \text{sym } \nabla u + O(\|\nabla u\|^2). \end{aligned} \quad (4.3.71)$$

So linearisiert sich  $U$  zu  $\mathbb{1} + \text{sym } \nabla u$ . Ebenso lässt sich zeigen, dass sich  $R$  zu  $\mathbb{1} + \text{skew } \nabla u$  linearisiert. Wir sehen,  $F$  ist genau dann eine linearisierte Rotation, wenn  $\nabla u$  schiefssymmetrisch ist.

#### Nichtlinearer Fall

Um auch im nichtlinearen Fall eine Verzerrung zu definieren, machen wir uns klar, was wir von einer Verzerrung erwarten:

- i) Sie drückt den Abstand von  $U$  zu  $\mathfrak{so}(3)$  aus.
- ii) Sie linearisiert zu  $\varepsilon = \text{sym } \nabla u$ .
- iii) Im Fall  $F \in \text{SO}(3)$  ist sie null (schon durch ii) mit abgedeckt).

Da der Abstandsbegriff nicht eindeutig festgelegt ist (und die Linearisierung auch noch Freiheiten lässt), kann es verschiedene Verzerrungsabbildungen geben.

**Beispiel 4.3.34** (Seth-Hill-Verzerrungsfamilie). Für ein  $m > 0$  definieren wir die Abbildung

$$U \mapsto \frac{1}{m}(U^m - \mathbb{1}) \quad (4.3.72)$$

als Seth-Hill-Verzerrung [72]. Für  $m \rightarrow 0$  konvergiert diese Abbildung gegen den Hencky-Dehnungstensor

$$U \mapsto \log U. \quad (4.3.73)$$

*Beweis der Linearisierungseigenschaft.* Wegen  $(\mathbb{1} + H)^n = \mathbb{1} + n \cdot H + O(\|H\|^2)$  ist

$$\begin{aligned} \frac{1}{m}(U^m - \mathbb{1}) &= \frac{1}{m}((F^T F)^{m/2} - \mathbb{1}) \\ &= \frac{1}{m}((\mathbb{1} + 2 \text{sym } \nabla u + \nabla u^T \nabla u)^{m/2} - \mathbb{1}) \\ &= \frac{1}{m}(\mathbb{1} + m \text{sym } \nabla u + O(\|\nabla u\|^2) - \mathbb{1}) \\ &= \varepsilon + O(\|\nabla u\|^2) \quad \text{mit } \varepsilon = \text{sym } \nabla u. \quad \blacksquare \end{aligned} \quad (4.3.74)$$

<sup>13</sup>Wir sehen,  $F$  ist genau dann eine linearisierte Rotation, wenn  $\nabla u$  schiefssymmetrisch ist.

### 4.3.13. Lineares Elastizitätsgesetz

Im nichtlinearen Fall lässt sich das Elastizitätsgesetz  $\sigma$  als Abbildung der Dehnung  $V \in \text{Sym}^+(3)$  auf den Spannungstensor  $\sigma(V)$  ausdrücken. Im linearen isotropen Fall verwenden wir statt der Dehnung die Linearisierung der Deformation  $\mathbb{1} + \varepsilon = \mathbb{1} + \text{sym } \nabla u$  und bilden so den linearen Verzerrungstensor  $\varepsilon$  auf die Spannung ab. Von dieser Abbildung verlangen wir zudem noch die Linearität in  $\varepsilon$ . Insgesamt erhalten wir eine isotrope lineare Abbildung  $\mathbb{C}$  mit  $\varepsilon \mapsto \sigma = \mathbb{C} \cdot \varepsilon$ .

In Lemma 2.5.40 auf Seite 38 haben wir gesehen, dass sich jede isotrope lineare Funktion  $\mathbb{C}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  als

$$\mathbb{C} \cdot X = \mu \text{sym } X + \mu_c \text{skew } X + \frac{\Lambda}{2} \text{tr}(X) \cdot \mathbb{1} \quad \text{für alle } X \in \mathbb{R}^{n \times n}, \quad (4.3.75)$$

mit konstanten Koeffizienten  $\mu, \mu_c, \Lambda \in \mathbb{R}$  schreiben lässt. Verwenden wir die Einschränkung auf symmetrische Argumente  $\mathbb{C}: \text{Sym}(3) \rightarrow \text{Sym}(3)$ , so entfällt der schiefsymmetrische Summand. Mit  $\text{sym } \varepsilon = \varepsilon$  und einer Umbenennung/Skalierung der Konstanten erhalten wir:

**Lemma 4.3.35** (Lineares Elastizitätsgesetz). *Jedes lineare Elastizitätsgesetz  $\sigma: \text{Sym}(3) \rightarrow \text{Sym}(3)$  als lineare isotrope Funktion im linearen Verzerrungstensor  $\varepsilon := \text{sym } \nabla u$  lässt sich als mit  $\mu, \Lambda \in \mathbb{R}$  bzw. mit  $\mu, \kappa \in \mathbb{R}$  gewichtete Summe*

$$\sigma(\varepsilon) = 2\mu \varepsilon + \Lambda \text{tr}(\varepsilon) \cdot \mathbb{1} = \sigma(\varepsilon) = 2\mu \text{dev}_3 \varepsilon + \kappa \text{tr}(\varepsilon) \cdot \mathbb{1} \quad (4.3.76)$$

*schreiben. Hierbei bezeichnen  $\Lambda$  und  $\mu$  die Lamé-Parameter und  $\kappa$  den Bulk-Modulus.*

**Bemerkung 4.3.36.** Die beiden Darstellungen ergeben sich über die Identität  $\varepsilon = \text{dev}_3 \varepsilon + \frac{1}{3} \text{tr}(\varepsilon) \cdot \mathbb{1}$  und den daraus resultierenden Zusammenhang  $\kappa = \Lambda + \frac{2}{3} \mu$ .



## 5. Konvexität von Energiefunktionen in $C$

Wesentliche Teile dieses Kapitels wurden bereits in [31] veröffentlicht.

In einem Beitrag von 2014 haben Lehmich et al. [44] Konvexitätsbedingungen für elastische Energien  $W(F) = \widehat{W}(C)$  im rechten Cauchy-Green-Deformationstensor  $C = F^T F$  angegeben – Bedingungen, welche auch eine Abhängigkeit vom Logarithmus der Determinante ausdrücken und zudem das notwendige Singularitätsverhalten  $\widehat{W}(C) \rightarrow \infty$  für  $\det F \rightarrow 0$  zeigen.

Spector [75] und Šilhavý [73] haben dieses Thema aufgegriffen und jeweils alternative Beweise vorgestellt. Ihr Hauptresultat, das wir schon als Satz 2.11.6 auf Seite 57 vorgestellt haben, lautet:

**Satz 5.0.37.** *Sei  $f \in C^2(\mathbb{R}_+)$ . Dann ist die Funktion  $\text{Sym}^+(n) \rightarrow \mathbb{R}$  mit  $X \mapsto f(\det X)$  genau dann konvex in  $X$ , wenn*

$$f''(s) + \frac{n-1}{ns} f'(s) \geq 0 \quad \text{und} \quad f'(s) \leq 0 \quad \text{für alle } s \in \mathbb{R}_+ \quad (5.0.1)$$

erfüllt ist.

**Bemerkung 5.0.38.** Beispielsweise erfüllt  $f(s) = -\log s$  die Bedingungen (5.0.1).

Lehmich et al. stellen die Behauptung auf, dass die Konvexität einer Energiefunktion in  $C$  eine nützliche Eigenschaft ist. Diese Behauptung wollen wir in diesem Kapitel mit einem auf ihr aufbauenden Lemma belegen.

Wir wissen, wäre die Energie polykonvex und koerziv, so könnten wir mittels direkter Methoden der Variationsrechnung die Existenz eines Minimierers zeigen. Aber wir müssen uns vor Augen führen, dass sowohl die Bedingung der *Rang-eins-Konvexität* [71], als auch die Bedingung der *Polykonvexität* (aus der Rang-eins-Konvexität folgt) [2] nicht allein durch die Konvexität der Energie in  $C$  erfüllt sind.

Natürlich haben wir immer die Möglichkeit, auf Basis des Satzes über die impliziten Funktionen einen Existenzbeweis zu führen, nur leider gilt dieser nur für kleine Umgebungen um die Identität [18].

Auch wenn Konvexität in  $C$  nicht die Polykonvexität in  $F$  impliziert, finden wir in [44] ein Beispiel einer Verzerrungsenergie  $W(F) = \widehat{W}(C)$ , die sowohl polykonvex in  $F$  als auch konvex in  $C$  ist; Zudem zeigt diese Energie auch das korrekte Verhalten für unendliche Kompression, d.h.  $\widehat{W}(C) \rightarrow \infty$  für  $\det C \rightarrow 0$ .

Eine allgemeinere Anwendung der Konvexität in  $C$  wurde durch Le Dret und Raoult [42] gegeben: Sie zeigen, dass wenn  $W(F) = \widehat{W}(C)$  in  $C$  konvex ist, dann ist die *quasikonvexe Hülle*  $QW$  von  $W$  durch eine Entsprechung der *Pipkin'schen Formel* [63] gegeben ist:

$$QW(F) = \inf_{S \in \text{Sym}^+(3)} \widehat{W}(F^T F + S). \quad (5.0.2)$$

Schließlich sei noch bemerkt, dass sich die Eigenschaft der Konvexität in  $C$  bei der Ableitung einiger Modelle der Elasto-Plastizität sinnvoll nutzen lässt, siehe dazu [33].

### 5.1. Anwendungen der Konvexität einer Energiefunktion in $C$

Wir betrachten das Problem der Minimierung der elastischen Energie  $I(\varphi) = \int_{\Omega} W(\nabla \varphi) dx$ , wobei  $\varphi: \Omega \rightarrow \mathbb{R}^3$  die Deformation eines elastischen Körpers  $\Omega \subset \mathbb{R}^3$  ist, welche einer Randbedingung  $\varphi|_{\Gamma} = \phi_0$  mit

## 5. Konvexität von Energiefunktionen in $C$

$\Gamma \subset \partial\Omega$  genügt, siehe [15]. Die zugehörige formale Euler-Lagrange-Gleichung, wie schon in Bemerkung 4.3.1 beschrieben, lautet  $\operatorname{div} S_1(F) = 0$ .

Das folgende Lemma, welches sich auch in [28, 29] finden lässt, zeigt die Anwendbarkeit unserer Konvexitätsbedingung.

**Lemma 5.1.1.** *Für ein Gebiet  $\Omega \subset \mathbb{R}^n$  mit  $C^2$ -Rand und eine auf seinem Rand definierte Funktion  $\phi: \partial\Omega \rightarrow \mathbb{R}^n$  sei  $W \in C^2(\operatorname{GL}^+(3))$  eine Energiefunktion, die garantiert, dass für jedes  $\varphi \in W^{1,p}(\Omega, \mathbb{R}^n)$ , welches die Randbedingung  $\varphi|_{\partial\Omega} = \phi$  erfüllt, das Energiefunktional*

$$I(\varphi) = \int_{\Omega} W(\nabla\varphi(x)) \, dx \quad (5.1.1)$$

wohldefiniert ist. Sei  $\varphi_0 \in C^1(\Omega)$  eine schwache Lösung der Euler-Lagrange-Gleichung

$$\operatorname{div} S_1(\nabla\varphi(x)) = 0 \quad \text{für alle } x \in \Omega$$

zum Energiefunktional (5.1.1), d.h.  $\varphi_0$  erfüllt die schwache Form der Euler-Lagrange-Gleichung<sup>1</sup>

$$\int_{\Omega} \langle S_1(\nabla\varphi(x)), \nabla\psi(x) \rangle \, dx = 0 \quad \text{für alle Testfunktionen } \psi \in C_0^\infty(\Omega). \quad (5.1.2)$$

Sei nun  $\widehat{W}$  die zugehörige Energiefunktion in  $C$ , d.h.  $\widehat{W} \in C^1(\operatorname{Sym}^+(3))$  mit  $\widehat{W}(F^T F) = W(F)$  für alle  $F \in \operatorname{GL}^+(3)$ .

Ist  $\widehat{W}$  konvex und  $S_2(\nabla\varphi_0(x))$  für alle  $x \in \Omega$  positiv semidefinit,<sup>2</sup> d.h.

$$\langle S_2(\nabla\varphi_0(x)), A \rangle \geq 0 \quad \text{für alle } A \in \mathbb{R}^{n \times n} \text{ und alle } x \in \Omega, \quad (5.1.3)$$

dann ist  $\varphi_0$  ein globaler Minimierer des Energiefunktionals (5.1.1).

**Bemerkung 5.1.2.** Es sei darauf hingewiesen, dass wir die Konvexität der Energie in  $C$  (also der Funktion  $\widehat{W}$ ) fordern. Anders als die Konvexität in  $F$  (der Funktion  $W$ ), die gegen physikalische Prinzipien verstößt und somit ausgeschlossen werden muss, steht die Konvexität in  $C$  in keinem Widerspruch zu plausiblen physikalischen Materialverhalten. Wir werden die Eigenschaft der Konvexität von  $\widehat{W}$  (in  $C$ ) als Monotonie des zweiten Piola-Kirchhoff-Spannungstensors  $S_2$

$$\langle S_2(C_1) - S_2(C_2), C_1 - C_2 \rangle \geq 0 \quad \text{für alle } C_1, C_2 \in \operatorname{Sym}^+(3) \quad (5.1.4)$$

im Beweis nutzen. Die Äquivalenz der Konvexitätsbedingungen findet sich in Proposition 2.11.3 auf Seite 57. Die zweite Forderung an  $S_2$  ist genau genommen eine Forderung an die Verkettung  $S_2 \circ \nabla\varphi_0$ .<sup>3</sup> Wir fordern die positive Semi-Definitheit von  $S_2(\nabla\varphi_0(x))$  für alle  $x \in \Omega$ . Sie drückt aus physikalischer Sicht den „tensionalen“ Charakter der (wieder auf die Referenzkonfiguration zurücktransformierten) Kräfte aus. Mit tensional bezeichnen wir den Umstand, dass die Lösung  $\varphi_0$  nur Zug- und keinerlei Kompressionskräfte aufweist.

*Beweis.* Wir haben zu zeigen

$$\begin{aligned} I(\varphi) - I(\varphi_0) &= \int_{\Omega} W(\nabla\varphi(x)) \, dx - \int_{\Omega} W(\nabla\varphi_0(x)) \, dx \\ &= \int_{\Omega} W(\nabla\varphi(x)) - W(\nabla\varphi_0(x)) \, dx \\ &= \int_{\Omega} \widehat{W}(\nabla\varphi(x)^T \nabla\varphi(x)) - \widehat{W}(\nabla\varphi_0(x)^T \nabla\varphi_0(x)) \, dx \geq 0 \quad \text{für alle } \varphi \in C^1(\Omega). \end{aligned} \quad (5.1.5)$$

<sup>1</sup>Also hat die Abbildung  $t \mapsto \int_{\Omega} W(\nabla\varphi_0(x) + t u(x)) \, dx$  für alle  $u \in C_0^\infty(\Omega)$  in  $t = 0$  eine Extremalstelle.

<sup>2</sup>Hierbei gilt  $D_C \widehat{W}(F^T F) = \frac{1}{2} S_2(F) = \frac{1}{2} \widehat{S}_2(F^T F)$ .

<sup>3</sup>Es wird nicht die Konvexität von  $S_2$  gefordert. Ganz im Gegenteil: Wäre  $S_2$  konvex, würde diese Konvexität zusammen mit der Konvexität von  $\widehat{W}$  in  $C$  umgehend die Konvexität von  $W$  in  $F$  ergeben!

## 5. Konvexität von Energiefunktionen in $\mathcal{C}$

Sei also  $\varphi \in C^1(\Omega)$  beliebig. Da  $\varphi$  und  $\varphi_0$  dieselben Randwerte haben, existiert ein  $\psi \in C_0^\infty(\Omega)$  mit  $\varphi(x) = \varphi_0(x) + \psi(x)$  für alle  $x \in \Omega$ .

Sei nun  $x \in \Omega$  beliebig. Wir schreiben unter Vernachlässigung der Arguments  $x$  verkürzt  $\nabla\varphi_0$ ,  $\nabla\varphi$ ,  $\nabla\psi$  und setzen  $C_0 := \nabla\varphi_0^T \nabla\varphi_0 \in \text{Sym}^+(3)$  und  $C := \nabla\varphi^T \nabla\varphi \in \text{Sym}^+(3)$ .

Mit (4.3.16) aus Proposition 4.3.15 und der Konvexitätscharakterisierung aus Proposition 2.11.3 ii) gilt

$$\widehat{W}(C) - \widehat{W}(C_0) \geq \langle D_C \widehat{W}(C_0), C - C_0 \rangle = \left\langle \frac{1}{2} \widehat{S}_2(C_0), C - C_0 \right\rangle.$$

Nun ist

$$C - C_0 = (\nabla\varphi_0 + \nabla\psi)^T (\nabla\varphi_0 + \nabla\psi) - \nabla\varphi_0^T \nabla\varphi_0 = \nabla\varphi_0^T \nabla\psi + \nabla\psi^T \nabla\varphi_0 + \nabla\psi^T \nabla\psi.$$

Mit (4.3.17) aus Proposition 4.3.15 ist

$$\begin{aligned} \left\langle \frac{1}{2} \widehat{S}_2(C_0), C - C_0 \right\rangle &= \left\langle \frac{1}{2} \widehat{S}_2(C_0), \nabla\varphi_0^T \nabla\psi + \nabla\psi^T \nabla\varphi_0 \right\rangle + \left\langle \frac{1}{2} \widehat{S}_2(C_0), \nabla\psi^T \nabla\psi \right\rangle \\ &= \langle S_1(\nabla\varphi_0), \nabla\psi \rangle + \left\langle \frac{1}{2} S_2(\nabla\varphi_0), \nabla\psi^T \nabla\psi \right\rangle. \end{aligned}$$

Nun ist  $\varphi_0$  eine Lösung der schwachen Form (5.1.2) und  $S_2(\nabla\varphi_0(x))$  nach Voraussetzung positiv definit, somit können wir schließen

$$I(\varphi) - I(\varphi_0) \geq \underbrace{\int_{\Omega} \langle S_1(\nabla\varphi_0(x)), \nabla\psi(x) \rangle dx}_{=0} + \underbrace{\int_{\Omega} \left\langle \frac{1}{2} S_2(\nabla\varphi_0(x)), \nabla\psi(x)^T \nabla\psi(x) \right\rangle dx}_{\geq 0} \geq 0.$$

Da  $\varphi \in C^1(\Omega)$  beliebig gewählt war, folgt (5.1.5). ■

**Bemerkung 5.1.3.** Wir haben in Lemma 5.1.1 die Konvexität von  $\widehat{W}$  gefordert, diese globale Konvexität im Beweis aber nicht benutzt. Tatsächlich wird die *globale* Konvexität von  $\widehat{W}$  nicht benötigt und es reicht die abgeschwächte Konvexitätsforderung

$$\widehat{W}(C) - \widehat{W}(C_0(x)) \geq \langle D_C \widehat{W}(C_0(x)), C - C_0(x) \rangle \quad \text{für alle } C \in \text{Sym}^+(3) \text{ und alle } x \in \Omega, \quad (5.1.6)$$

wobei  $C_0 = \nabla\varphi_0^T \nabla\varphi_0$  der Cauchy-Green-Tensor zur vorgegebenen schwachen Lösung  $\varphi_0$  der Euler-Lagrange-Gleichung ist.

Spector [75, S. 254] zeigt in seinem Beweis von Satz 5.0.37, dass diese abgeschwächte Konvexitätsbedingung (5.1.6) im Falle von

$$\widehat{W}(C) = f(\det C)$$

mit einer gegebenen Funktion  $f: \mathbb{R}_+ \rightarrow \mathbb{R}$  für alle  $C \in \text{Sym}^+(3)$  erfüllt ist, falls

$$f''(s) + \frac{n-1}{ns} f'(s) \geq 0 \quad \text{für alle } s \in \mathbb{R}_+, \quad (5.1.7)$$

$$f'(s) \leq 0 \quad \text{für alle } s \in (0, \det C_0] \quad (5.1.8)$$

gilt.

Vor allem können die Bedingungen an  $\widehat{W}$  aus Lemma 5.1.1 somit abgeschwächt werden, um Energiefunktionen der Form

$$\widehat{W}(C) = M(C) + f(\det C)$$

unter der Bedingung,  $M: \text{Sym}^+(n) \rightarrow \mathbb{R}$  ist (global) konvex und  $f: \mathbb{R}_+ \rightarrow \mathbb{R}$  erfüllt (5.1.7) und (5.1.8) für alle  $C_0 = C_0(x) = \nabla\varphi_0(x)^T \nabla\varphi_0(x)$ , mit einzuschließen.

## 5. Konvexität von Energiefunktionen in $C$

**Bemerkung 5.1.4.** Wenn wir die Konvexitätsbedingung aus Lemma 5.1.1 zur allgemeinen Konvexitätsbedingung  $D_C^2 \widehat{W}(C) \cdot (H, H) \geq c^+ \|H\|^2$  für alle symmetrischen Matrizen  $H$  mit einem  $c^+ > 0$  verschärfen und falls die Lösung  $\varphi_0$  zudem ein Diffeomorphismus ist, dann lässt sich zeigen, dass

$$\int_{\Omega} D_F^2 W(\nabla \varphi_0) \cdot (\nabla \vartheta, \nabla \vartheta) \, dx \geq \tilde{c}^+ \|\nabla \vartheta\|_{L^2(\Omega)}^2$$

für alle Testfunktionen  $\vartheta \in C_0^\infty(\Omega)$  mit einem  $\tilde{c}^+ > 0$  gilt, was seinerseits die lokale Stabilität der Lösung impliziert [54].

In Bemerkung 5.1.3 haben wir den Volumenterm  $f(\det C)$  aus der geschlossenen Darstellung von  $\widehat{W}$  herausgelöst. Dieses Vorgehen wollen wir im nächsten Abschnitt mit den Valanis-Landel-Energien und dazu sehr ähnlichen noch weiterführen.

## 5.2. Die Konvexität der Valanis-Landel-Energien in $C$

**Definition 5.2.1.** Eine Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  hat die (verallgemeinerte) *Valanis-Landel-Form*, wenn es differenzierbare Funktionen  $w, f: \mathbb{R}_+ \rightarrow \mathbb{R}$  mit  $f'(1) = w'(1) = 0$  gibt, sodass

$$W(F) = \sum_{i=1}^3 \underbrace{w(\lambda_i)}_{\text{Achsenterm}} + \underbrace{f(\det F)}_{\text{Volumenterm}} \quad \text{für alle } F \in \text{GL}^+(3). \quad (5.2.1)$$

Hierbei sind  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_+$  die Singulärwerte von  $F \in \text{GL}^+(3)$ .

**Bemerkung 5.2.2.** Der Begriff der Verallgemeinerung rührt von der Hinzunahme des Volumenterms  $f$ , um die Energie nicht nur auf inkompressible Deformationen anwenden zu können: Ohne diesen Term, d.h. für  $f \equiv 0$  erhalten wir einen klassischen Vertreter der inkompressiblen *Valanis-Landel-Energie*.

Da die Singulärwerte von  $F$  die Quadratwurzeln der Eigenwerte von  $C = F^T F$  sind und auch  $\det F = \sqrt{\det C}$  gilt, können wir die Funktion  $W$  auch als Funktion in  $C$  schreiben als

$$W(C) = \widehat{w}(\lambda_1^2) + \widehat{w}(\lambda_2^2) + \widehat{w}(\lambda_3^2) + \widehat{f}(\det \lambda_1^2 \lambda_2^2 \lambda_3^2) \quad \text{für alle } C \in \text{Sym}^+(3) \quad (5.2.2)$$

mit  $\widehat{w}(x) = w(\sqrt{x})$  und  $\widehat{f}(x) = f(\sqrt{x})$ . Hierbei sind  $\lambda_1^2, \lambda_2^2, \lambda_3^2 \in \mathbb{R}_+$  die Eigenwerte von  $C$ . Wir können nun ein Ergebnis von Davis [22], der für herkömmliche Valanis-Landel-Energien ohne den Volumenterm (d.h.  $f \equiv 0$ ) gezeigt hat, dass  $W$  genau dann konvex in  $C$  ist, wenn  $\widehat{w}$  konvex ist, erweitern.

**Proposition 5.2.3.** Sei  $\widehat{W}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  einer Energiefunktion der (verallgemeinerten) Valanis-Landel-Form (5.2.2). Dann ist  $W$  genau dann konvex in  $C$ , wenn  $\widehat{w}$  konvex ist und  $\widehat{f}$  die Bedingung (5.0.1) aus Satz 5.0.37

$$\widehat{f}''(s) + \frac{n-1}{ns} \widehat{f}'(s) \geq 0 \quad \text{und} \quad \widehat{f}'(s) \leq 0 \quad \text{für alle } s \in \mathbb{R}_+ \quad (5.2.3)$$

erfüllt.

*Beweis.* Siehe Folgerung 2.11.10 auf Seite 58. ■

**Bemerkung 5.2.4.** Mit  $f \equiv 0$  ist die Bedingung (5.0.1) ebenfalls erfüllt, wir erhalten so das Ergebnis von Davis.

### 5.3. Beispiele für die Konvexität von

$$\widehat{W}(C) = \alpha \operatorname{tr}(C) + \beta \operatorname{tr}(C)^2 + \gamma \operatorname{tr}(C^2) - \delta \log \det C$$

Mit Folgerung 2.11.11 auf Seite 58 haben wir die Konvexität der Klasse von Funktionen

$$\widehat{W}(C) = \alpha \operatorname{tr}(C) + \beta \operatorname{tr}(C)^2 + \gamma \operatorname{tr}(C^2) - \delta \log \det C + \zeta. \quad (5.3.1)$$

für  $\alpha, \zeta \in \mathbb{R}$  und  $\beta, \gamma, \delta \in \mathbb{R}_+ \cup \{0\}$  gezeigt. Wir müssen diese Parameter für physikalisch sinnvolle Energien (spannungsfreie Referenzkonfiguration) selbstverständlich so wählen, dass  $D\widehat{W}(\mathbb{1}) = 0$  erfüllt gilt. Als Beispiel hierzu betrachten wir:

**Beispiel 5.3.1.** Es seien  $\alpha, \beta \in \mathbb{R}_+$ . Die Funktionen  $\widehat{W}_1, \widehat{W}_2, \widehat{W}_3: \operatorname{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$\widehat{W}_1(C) = \alpha \operatorname{tr}(C)^2 + \beta \operatorname{tr}(C^2) - \log \det C \quad \text{für} \quad 6\alpha + 2\beta - 1 = 0, \quad (5.3.2)$$

$$\widehat{W}_2(C) = \alpha \operatorname{tr}(C) + \beta \operatorname{tr}(C^2) - \log \det C \quad \text{für} \quad \alpha + 2\beta - 1 = 0, \quad (5.3.3)$$

$$\widehat{W}_3(C) = \alpha \operatorname{tr}(C)^2 + \beta \operatorname{tr}(C) - \log \det C \quad \text{für} \quad 6\alpha + \beta - 1 = 0 \quad (5.3.4)$$

sind konvex und es gilt  $D\widehat{W}_1(\mathbb{1}) = D\widehat{W}_2(\mathbb{1}) = D\widehat{W}_3(\mathbb{1}) = 0$ .

Aufgrund von  $\operatorname{tr}(C) = \sum \lambda_i^2$  und  $\operatorname{tr}(C^2) = \sum \lambda_i^4$  sind die so gegebenen Energien sehr eng mit den Valanis-Landel-Energien verwandt.

**Bemerkung 5.3.2.** Die Ableitungen lauten (zu finden in Lemma 5.5.1 am Ende dieses Abschnitts) im ersten Fall  $D\widehat{W}_1(C) = 2\alpha \operatorname{tr}(C) \cdot \mathbb{1} + 2\beta C - C^{-1}$ , im zweiten Fall  $D\widehat{W}_2(C) = \alpha \cdot \mathbb{1} + 2\beta C - C^{-1}$  und im dritten Fall  $D\widehat{W}_3(C) = 2\alpha \operatorname{tr}(C) \cdot \mathbb{1} + \beta \cdot \mathbb{1} - C^{-1}$ . Durch direktes Einsetzen sehen wir leicht die Eigenschaft der spannungsfreien Referenzkonfiguration.

**Bemerkung 5.3.3.** Die zu  $\widehat{W}$  aus (5.3.1) gehörige Energie  $W$  in  $F$  hat die Eigenschaft, polykonvex und koerziv zu sein. Dies garantiert die Existenz eines (die Randbedingung erfüllenden) Minimierers  $\varphi_0$  des Energiefunktionals

$$I(\varphi) := \int_{\Omega} W(\nabla \varphi(x)) \, dx. \quad (5.3.5)$$

Die Energie erfüllt die Voraussetzung von Lemma 2.11.12. Aus ihr folgt die Abschätzung  $|W(\nabla \varphi(x))| \leq c_1 \|\nabla \varphi(x)\|^4 + c_2$ , welche garantiert, dass für jedes  $\varphi \in W^{1,4}(\Omega, \mathbb{R}^n)$  mit  $\det \nabla \varphi(x) \geq c^+$  für alle  $x \in \Omega$  das Energiefunktional wohldefiniert ist.

Das heißt, ist  $\varphi_0 \in W^{1,4}(\Omega, \mathbb{R}^n)$  mit  $\det \nabla \varphi_0 \geq c^+$  für alle  $x \in \Omega$  ein globaler Minimierer, so erfüllt er auch die schwache Form der Euler-Lagrange-Gleichung.

Mit numerischen Methoden lassen sich leicht Lösungen der schwachen Form der Euler-Lagrange-Gleichung finden. Aber solange wir nicht wissen, ob  $\det \nabla \varphi_0 \geq c^+$  für ein  $c^+ > 0$  für den Minimierer gilt, wissen wir nicht, ob wir den Minimierer gefunden haben, selbst wenn wir alle Lösungen der schwachen Form der Euler-Lagrange-Gleichung kennen.

Hier zeigt sich die Nützlichkeit von Lemma 5.1.1. Wir können es direkt auf die Funktionen vom Typ (5.3.1) übertragen und wissen so, ob die unter den obigen Voraussetzungen gefundene Lösung ein Minimierer ist:

**Proposition 5.3.4.** Sei  $\widehat{W}: \operatorname{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$\widehat{W}(C) = \alpha \operatorname{tr}(C) + \beta \operatorname{tr}(C)^2 + \gamma \operatorname{tr}(C^2) - \delta \log \det C + \zeta. \quad (5.3.6)$$

und  $\alpha, \beta, \gamma, \delta, \zeta \in \mathbb{R}_+ \cup \{0\}$  so gegeben, sodass  $D\widehat{W}(\mathbb{1}) = 0$  sichergestellt ist. Angenommen,  $\widehat{W}(C)$  birgt eine glatte Lösung  $\varphi_0$  der Euler-Lagrange-Gleichung, welche  $\det \nabla \varphi_0 \geq c^+ > 0$  und  $S_2(\nabla \varphi_0)$  ist positiv semidefinit erfüllt. Dann ist  $\varphi_0$  ein globaler Minimierer.

## 5. Konvexität von Energiefunktionen in $\mathcal{C}$

Wir wollen uns noch die linearen Elastizitätsgesetze der drei Energien  $\widehat{W}_1, \widehat{W}_2, \widehat{W}_3$  aus Beispiel 5.3.1 ansehen. Am Ende dieses Abschnittes in Lemma 5.5.1 bestimmen wir die bis zur zweiten Ordnung nach  $\nabla u$  in  $\mathbb{1}$  entwickelten Energieapproximationen. Es gilt für  $\nabla u \rightarrow 0$ :

$$\widehat{W}_1(C) = \frac{3}{2} + (-12\alpha + 4) \|\operatorname{dev}_3 \varepsilon\|^2 + \frac{4}{3} \operatorname{tr}(\varepsilon)^2 + \mathcal{O}(\|\nabla u\|^3),$$

$$\widehat{W}_2(C) = \frac{3}{2} \alpha + \frac{3}{2} + (-2\alpha + 4) \|\operatorname{dev}_3 \varepsilon\|^2 + \left(-\frac{2}{3} \alpha + \frac{4}{3}\right) \operatorname{tr}(\varepsilon)^2 + \mathcal{O}(\|\nabla u\|^3),$$

$$\widehat{W}_3(C) = -9\alpha + 3 + 2 \|\operatorname{dev}_3 \varepsilon\|^2 + \left(4\alpha + \frac{2}{3}\right) \operatorname{tr}(\varepsilon)^2 + \mathcal{O}(\|\nabla u\|^3).$$

Nach dem linearen Verzerrungstensor  $\varepsilon$  abgeleitet, erhalten wir unter Vernachlässigung der Terme höherer Ordnung die linearen Elastizitätsgesetze der Form

$$\sigma = 2\mu \operatorname{dev}_3 \varepsilon + \kappa \operatorname{tr}(\varepsilon) \cdot \mathbb{1} = 2\lambda \varepsilon + \kappa \operatorname{tr}(\varepsilon) \cdot \mathbb{1}.$$

So ist:<sup>4</sup>

$$\begin{aligned} \mu &= -12\alpha + 4, & \kappa &= \frac{8}{3} & \text{für} & \widehat{W}_1, \\ \mu &= -2\alpha + 4, & \kappa &= -\frac{4}{3} + \frac{8}{3} & \text{für} & \widehat{W}_2, \\ \mu &= 1, & \kappa &= 8\alpha + \frac{4}{3} & \text{für} & \widehat{W}_3. \end{aligned}$$

### 5.4. Die Konvexität der Saint-Venant-Kirchhoff-Energie in $\mathcal{C}$

Eine weitere Klasse von Energiefunktionen wird durch die sogenannte *Saint-Venant-Kirchhoff-Energie* [30, 43, 65] dargestellt.

**Definition 5.4.1.** Seien  $\mu \in \mathbb{R}_+$  und  $\Lambda \in \mathbb{R}$  mit  $3\Lambda + 2\mu > 0$ . Dann nennt  $\widehat{W}_{\text{SVK}}: \operatorname{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$\widehat{W}_{\text{SVK}}(C) = \frac{\mu}{4} \|C - \mathbb{1}\|^2 + \frac{\Lambda}{8} (\operatorname{tr}(C - \mathbb{1}))^2 \quad (5.4.1)$$

die *Saint-Venant-Kirchhoff-Energie*.

**Proposition 5.4.2.** Die Saint-Venant-Kirchhoff-Energie ist für alle  $\mu \in \mathbb{R}_+$  und  $\Lambda \in \mathbb{R}$  mit  $3\Lambda + 2\mu > 0$  von der Form (5.3.1) und somit konvex in  $\mathcal{C}$ .

**Bemerkung 5.4.3.** Für alle  $\mu, \Lambda \in \mathbb{R}$  ist  $\widehat{W}_{\text{SVK}}$  nicht rang-eins-konvex [65].

*Beweis.* Unter Berücksichtigung von  $C^T = C$  formen wir um. Somit ist

$$\begin{aligned} \widehat{W}_{\text{SVK}}(C) &= \frac{\mu}{4} \|C - \mathbb{1}\|^2 + \frac{\Lambda}{8} (\operatorname{tr}(C - \mathbb{1}))^2 \\ &= \frac{\mu}{4} \operatorname{tr}((C - \mathbb{1})^T (C - \mathbb{1})) + \frac{\Lambda}{8} (\operatorname{tr}(C) - 3)^2 \\ &= \frac{\mu}{4} (\operatorname{tr}(C^2) - 2\operatorname{tr}(C) + 3) + \frac{\Lambda}{8} ((\operatorname{tr}(C))^2 - 6\operatorname{tr}(C) + 9) \\ &= \alpha \operatorname{tr}(C) + \beta \operatorname{tr}(C)^2 + \gamma \operatorname{tr}(C^2) + \zeta, \end{aligned}$$

mit

$$\alpha = -\frac{1}{4} (3\Lambda + 2\mu), \quad \beta = \frac{\Lambda}{8} > 0, \quad \gamma = \frac{\mu}{4} > 0 \quad \text{und} \quad \zeta = \frac{3}{8} (3\Lambda + 2\mu)$$

von der Form (5.3.1). ■

<sup>4</sup>Wir erinnern uns, es gilt:  $D\|\operatorname{dev}_3 \varepsilon\|^2 = 2 \operatorname{dev}_3 \varepsilon$  und  $D \operatorname{tr}(\varepsilon)^2 = 2 \operatorname{tr}(\varepsilon) \cdot \mathbb{1}$

**Bemerkung 5.4.4.** In Beispiel 6.1.34 auf Seite 150 zeigen wir zudem, dass das durch die Saint-Venant-Kirchoff-Energie induzierte Elastizitätsgesetz  $\hat{\sigma}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  weder die empirischen Ungleichungen noch die strikten Baker-Ericksen-Ungleichungen erfüllt.

## 5.5. Bestimmung der Approximation bis zur zweiten Ordnung von $\widehat{W}_1$ , $\widehat{W}_2$ und $\widehat{W}_3$

Wir haben in Abschnitt 5.3 drei Beispielennergien der Form (5.3.1)

$$\widehat{W}(C) = \alpha \text{tr}(C) + \beta \text{tr}(C)^2 + \gamma \text{tr}(C^2) - \gamma \log \det C + \zeta.$$

auf ihre Bedeutung zum Thema Konvexität untersucht. Hierzu haben wir bereits auf das folgende, aufgrund der Rechenlastigkeit ausgelagerte, Lemma zurückgegriffen.

**Lemma 5.5.1.** *Es seien  $\alpha, \beta \in \mathbb{R}_+$ . Die Funktionen  $\widehat{W}_1, \widehat{W}_2, \widehat{W}_3: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit*

$$\widehat{W}_1(C) = \alpha \text{tr}(C)^2 + \beta \text{tr}(C^2) - \log \det C \quad \text{für} \quad 6\alpha + 2\beta - 1 = 0, \quad (5.5.1)$$

$$\widehat{W}_2(C) = \alpha \text{tr}(C) + \beta \text{tr}(C^2) - \log \det C \quad \text{für} \quad \alpha + 2\beta - 1 = 0, \quad (5.5.2)$$

$$\widehat{W}_3(C) = \alpha \text{tr}(C)^2 + \beta \text{tr}(C) - \log \det C \quad \text{für} \quad 6\alpha + \beta - 1 = 0 \quad (5.5.3)$$

haben die Ableitungen

$$D\widehat{W}_1(C) = 2\alpha \text{tr}(C) \cdot \mathbb{1} + 2\beta C - C^{-1}, \quad (5.5.4)$$

$$D\widehat{W}_2(C) = \alpha \cdot \mathbb{1} + 2\beta C - C^{-1}, \quad (5.5.5)$$

$$D\widehat{W}_3(C) = 2\alpha \text{tr}(C) \cdot \mathbb{1} + \beta \cdot \mathbb{1} - C^{-1}. \quad (5.5.6)$$

Für diese gilt  $D\widehat{W}_1(\mathbb{1}) = D\widehat{W}_2(\mathbb{1}) = D\widehat{W}_3(\mathbb{1}) = 0$ . Die bis zur zweiten Ordnung nach  $\nabla u$  in  $\mathbb{1}$  entwickelten Energieapproximationen lauten für  $\nabla u \rightarrow 0$ :

$$\begin{aligned} \widehat{W}_1(C) &= \frac{3}{2} + 4\alpha \text{tr}(\varepsilon)^2 + (-12\alpha + 4) \text{tr}(\varepsilon^2) + \mathcal{O}(\|\nabla u\|^3) \\ &= \frac{3}{2} + (-12\alpha + 4) \|\text{dev}_3 \varepsilon\|^2 + \frac{4}{3} \text{tr}(\varepsilon)^2 + \mathcal{O}(\|\nabla u\|^3), \end{aligned} \quad (5.5.7)$$

$$\begin{aligned} \widehat{W}_2(C) &= \frac{3}{2}\alpha + \frac{3}{2} + (-2\alpha + 4) \text{tr}(\varepsilon^2) + \mathcal{O}(\|\nabla u\|^3) \\ &= \frac{3}{2}\alpha + \frac{3}{2} + (-2\alpha + 4) \|\text{dev}_3 \varepsilon\|^2 + \left(-\frac{2}{3}\alpha + \frac{4}{3}\right) \text{tr}(\varepsilon)^2 + \mathcal{O}(\|\nabla u\|^3), \end{aligned} \quad (5.5.8)$$

$$\begin{aligned} \widehat{W}_3(C) &= -9\alpha + 3 + 4\alpha \text{tr}(\varepsilon)^2 + 2 \text{tr}(\varepsilon^2) + \mathcal{O}(\|\nabla u\|^3) \\ &= -9\alpha + 3 + \left(4\alpha + \frac{2}{3}\right) \text{tr}(\varepsilon)^2 + 2 \|\text{dev}_3 \varepsilon\|^2 + \mathcal{O}(\|\nabla u\|^3). \end{aligned} \quad (5.5.9)$$

Hierbei ist  $\nabla u := \nabla \varphi - \mathbb{1} \in \mathbb{R}^{3 \times 3}$ ,  $C := \nabla u^T \nabla u \in \text{Sym}^+(3)$  und  $\varepsilon := \text{sym } \nabla u$ .

Um das Lemma zu beweisen, brauchen wir die Darstellung der beteiligten Ausdrücke  $\text{tr}(C)$ ,  $\text{tr}(C)^2$ ,  $\text{tr}(C^2)$  und  $\log \det C$  im linearen Verzerrungstensor  $\varepsilon$  und im Verschiebungsgradienten  $\nabla u$ . Schlüssel für den Determinantenterm ist die folgende Proposition, welche die Entwicklung der Determinante in  $\mathbb{1}$  auf die Invarianten zurückführt.

**Proposition 5.5.2.** *Für alle  $X \in \mathbb{R}^{n \times n}$  gilt*

$$\det(\mathbb{1} + X) = 1 + I_1 + I_2 + \dots + I_n. \quad (5.5.10)$$

## 5. Konvexität von Energiefunktionen in $\mathcal{C}$

Hierbei sind

$$I_1 = \operatorname{tr}(X), \quad I_2 = \operatorname{tr}(\operatorname{Cof} X) = \frac{1}{2} (\operatorname{tr}(X)^2 - \operatorname{tr}(X^2)), \quad \dots, \quad I_n = \det X \quad (5.5.11)$$

die Invarianten von  $X$ .

*Beweis.* Sei  $X \in \mathbb{R}^{n \times n}$ , dann lautet das charakteristische Polynom von  $X$

$$\det(X - \lambda \cdot \mathbf{1}) = (-\lambda)^n + I_1 (-\lambda)^{n-1} + I_2 (-\lambda)^{n-2} + \dots + I_n.$$

Setzen wir  $\lambda = -1$  erhalten wir die Aussage. ■

Unter Einfluss der letzten Proposition können wir mittels der Definitionen  $\nabla u := \nabla \varphi - \mathbf{1} \in \mathbb{R}^{3 \times 3}$ ,  $C := \nabla u^T \nabla u \in \operatorname{Sym}^+(3)$  und  $\varepsilon := \operatorname{sym} \nabla u$  direkt berechnen:

**Proposition 5.5.3.** *Für alle  $\nabla u \in \operatorname{GL}^+(3)$  gilt für  $\nabla u \rightarrow 0$*

$$\operatorname{tr}(C) = 3 + 2 \operatorname{tr}(\varepsilon) + \|\nabla u\|^2, \quad (5.5.12)$$

$$\operatorname{tr}(C)^2 = 9 + 12 \operatorname{tr}(\varepsilon) + 4 \operatorname{tr}(\varepsilon)^2 + 6 \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3), \quad (5.5.13)$$

$$\operatorname{tr}(C^2) = 3 + 4 \operatorname{tr}(\varepsilon) + 4 \operatorname{tr}(\varepsilon^2) + 2 \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3), \quad (5.5.14)$$

$$\det C = 1 + 2 \operatorname{tr}(\varepsilon) + 2 \operatorname{tr}(\varepsilon^2) - 2 \operatorname{tr}(\varepsilon^2) + \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3), \quad (5.5.15)$$

$$\log \det C = 2 \operatorname{tr}(\varepsilon) - 2 \operatorname{tr}(\varepsilon^2) + \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3). \quad (5.5.16)$$

*Beweis.* Es ist mit  $C = (\mathbf{1} + \nabla u)^T (\mathbf{1} + \nabla u) = \mathbf{1} + 2\varepsilon + \nabla u^T \nabla u$

$$\operatorname{tr}(C) = \operatorname{tr}(\mathbf{1} + 2\varepsilon + \nabla u^T \nabla u) = 3 + 2 \operatorname{tr}(\varepsilon) + \|\nabla u\|^2,$$

$$\begin{aligned} \operatorname{tr}(C)^2 &= (3 + 2 \operatorname{tr}(\varepsilon) + \|\nabla u\|^2)^2 \\ &= 9 + 4 \operatorname{tr}(\varepsilon)^2 + \|\nabla u\|^4 + 12 \operatorname{tr}(\varepsilon) + 6 \|\nabla u\|^2 + 4 \operatorname{tr}(\varepsilon) \|\nabla u\|^2 \\ &= 9 + 12 \operatorname{tr}(\varepsilon) + 4 \operatorname{tr}(\varepsilon)^2 + 6 \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3), \end{aligned}$$

$$\begin{aligned} \operatorname{tr}(C^2) &= \operatorname{tr}((\mathbf{1} + 2\varepsilon + \nabla u^T \nabla u)^2) \\ &= \operatorname{tr}(\mathbf{1} + 4\varepsilon^2 + (\nabla u^T \nabla u)^2 + 4\varepsilon + 2\nabla u^T \nabla u + 4\varepsilon \nabla u^T \nabla u) \\ &= 3 + 4 \operatorname{tr}(\varepsilon) + 4 \operatorname{tr}(\varepsilon^2) + 2 \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3). \end{aligned}$$



## 5. Konvexität von Energiefunktionen in $C$

Mit Proposition 5.5.2 folgt

$$\begin{aligned}
\det C &= \det(\mathbf{1} + 2\varepsilon + \nabla u^T \nabla u) \\
&= 1 + \operatorname{tr}(2\varepsilon + \nabla u^T \nabla u) + \frac{1}{2} \left( \operatorname{tr}(2\varepsilon + \nabla u^T \nabla u)^2 - \operatorname{tr}((2\varepsilon + \nabla u^T \nabla u)^2) \right) + \det(2\varepsilon + \nabla u^T \nabla u) \\
&= 1 + 2 \operatorname{tr}(\varepsilon) + \operatorname{tr}(\nabla u^T \nabla u) \\
&\quad + \frac{1}{2} (2 \operatorname{tr}(\varepsilon) + \operatorname{tr}(\nabla u^T \nabla u))^2 - \frac{1}{2} \operatorname{tr}(4\varepsilon^2 + 4\varepsilon \nabla u^T \nabla u + (\nabla u^T \nabla u)^2) \\
&\quad + \det(2\varepsilon + \nabla u^T \nabla u) \\
&= 1 + 2 \operatorname{tr}(\varepsilon) + \|\nabla u\|^2 \\
&\quad + 2 \operatorname{tr}(\varepsilon)^2 + 2 \operatorname{tr}(\varepsilon) \|\nabla u\|^2 + \frac{1}{2} \|\nabla u\|^4 - 2 \operatorname{tr}(\varepsilon^2) - 2 \operatorname{tr}(\varepsilon \nabla u^T \nabla u) - \frac{1}{2} \operatorname{tr}((\nabla u^T \nabla u)^2) \\
&\quad + \det(2\varepsilon + \nabla u^T \nabla u) \\
&= 1 + 2 \operatorname{tr}(\varepsilon) + 2 \operatorname{tr}(\varepsilon)^2 - 2 \operatorname{tr}(\varepsilon^2) + \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3).
\end{aligned}$$

Es ist somit

$$\det C - 1 = 2 \operatorname{tr}(\varepsilon) + 2 \operatorname{tr}(\varepsilon)^2 - 2 \operatorname{tr}(\varepsilon^2) + \|\nabla u\|^2 + R(\nabla u),$$

Quadrieren liefert

$$(\det C - 1)^2 = 4 \operatorname{tr}(\varepsilon)^2 + \mathcal{O}(\|\nabla u\|^3).$$

Für Potenzen  $k > 2$  gilt

$$(\det C - 1)^k = \mathcal{O}(\|\nabla u\|^3).$$

Für  $x \in (-1, 1)$  liefert die Taylorentwicklung der Logarithmusfunktion in der Eins:  $\log(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \dots$ , so dass wir schlussendlich

$$\begin{aligned}
\log \det C &= (\det C - 1) - \frac{1}{2}(\det C - 1)^2 + \mathcal{O}(\|\nabla u\|^3) \\
&= 2 \operatorname{tr}(\varepsilon) - 2 \operatorname{tr}(\varepsilon^2) + \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3)
\end{aligned}$$

erhalten. ■

Nun sind wir auch in der Lage, das ursprüngliche Lemma zu zeigen.

*Beweis von Lemma 5.5.1.* Die Konvexität von  $\widehat{W}_1, \widehat{W}_2, \widehat{W}_3$  folgt direkt aus Lemma 2.11.8. Die Ableitungen folgen mit Lemma 2.5.34.

Mit Proposition 5.5.3 und  $\operatorname{tr}(\varepsilon^2) = \|\varepsilon\|^2 = \|\operatorname{dev}_3 \varepsilon\|^2 + \frac{1}{3} \operatorname{tr}(\varepsilon)^2$  folgt

$$\begin{aligned}
\widehat{W}_1(C) &= \alpha (9 + 12 \operatorname{tr}(\varepsilon) + 4 \operatorname{tr}(\varepsilon)^2 + 6 \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3)) \\
&\quad + \beta (3 + 4 \operatorname{tr}(\varepsilon) + 4 \operatorname{tr}(\varepsilon^2) + 2 \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3)) \\
&\quad - (2 \operatorname{tr}(\varepsilon) - 2 \operatorname{tr}(\varepsilon^2) + \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3)) \\
&= \underbrace{9\alpha + 3\beta}_{=3/2} + \underbrace{(12\alpha + 4\beta - 2)}_{=0} \operatorname{tr}(\varepsilon) + 4\alpha \operatorname{tr}(\varepsilon)^2 \\
&\quad + \underbrace{(4\beta + 2)}_{=12\alpha-4} \operatorname{tr}(\varepsilon^2) + \underbrace{(6\alpha + 2\beta - 1)}_{=0} \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3) \\
&= \frac{3}{2} + 4\alpha \operatorname{tr}(\varepsilon)^2 + (-12\alpha + 4) \operatorname{tr}(\varepsilon^2) + \mathcal{O}(\|\nabla u\|^3) \\
&= \frac{3}{2} + 4\alpha \operatorname{tr}(\varepsilon)^2 + (-12\alpha + 4) (\|\operatorname{dev}_3 \varepsilon\|^2 + \frac{1}{3} \operatorname{tr}(\varepsilon)^2) + \mathcal{O}(\|\nabla u\|^3)
\end{aligned}$$

5. Konvexität von Energiefunktionen in  $\mathcal{C}$

$$= \frac{3}{2} + (-12\alpha + 4) \|\operatorname{dev}_3 \varepsilon\|^2 + \frac{4}{3} \operatorname{tr}(\varepsilon)^2 + \mathcal{O}(\|\nabla u\|^3)$$

$$\begin{aligned} \widehat{W}_2(\mathcal{C}) &= \alpha (3 + 2 \operatorname{tr}(\varepsilon) + \|\nabla u\|^2) \\ &\quad + \beta (3 + 4 \operatorname{tr}(\varepsilon) + 4 \operatorname{tr}(\varepsilon^2) + 2\|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3)) \\ &\quad - (2 \operatorname{tr}(\varepsilon) - 2 \operatorname{tr}(\varepsilon^2) + \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3)) \\ &= \underbrace{(3\alpha + 3\beta)}_{=3/2\alpha+3/2} + \underbrace{(2\alpha + 4\beta - 2)}_{=0} \operatorname{tr}(\varepsilon) \\ &\quad + \underbrace{(4\beta + 2)}_{=-2\alpha+4} \operatorname{tr}(\varepsilon^2) + \underbrace{(\alpha + 2\beta - 1)}_{=0} \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3) \\ &= \frac{3}{2} \alpha + \frac{3}{2} + (-2\alpha + 4) \operatorname{tr}(\varepsilon^2) + \mathcal{O}(\|\nabla u\|^3) \\ &= \frac{3}{2} \alpha + \frac{3}{2} + (-2\alpha + 4) (\|\operatorname{dev}_3 \varepsilon\|^2 + \frac{1}{3} \operatorname{tr}(\varepsilon)^2) + \mathcal{O}(\|\nabla u\|^3) \\ &= \frac{3}{2} \alpha + \frac{3}{2} + (-2\alpha + 4) \|\operatorname{dev}_3 \varepsilon\|^2 + \left(-\frac{2}{3} \alpha + \frac{4}{3}\right) \operatorname{tr}(\varepsilon)^2 + \mathcal{O}(\|\nabla u\|^3) \end{aligned}$$

$$\begin{aligned} \widehat{W}_3(\mathcal{C}) &= \alpha (9 + 12 \operatorname{tr}(\varepsilon) + 4 \operatorname{tr}(\varepsilon)^2 + 6\|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3)) \\ &\quad + \beta (3 + 2 \operatorname{tr}(\varepsilon) + \|\nabla u\|^2) \\ &\quad - (2 \operatorname{tr}(\varepsilon) - 2 \operatorname{tr}(\varepsilon^2) + \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3)) \\ &= \underbrace{(9\alpha + 3\beta)}_{=-9\alpha+3} + \underbrace{(12\alpha + 2\beta - 2)}_{=0} \operatorname{tr}(\varepsilon) + 4\alpha \operatorname{tr}(\varepsilon)^2 \\ &\quad + 2 \operatorname{tr}(\varepsilon^2) + \underbrace{(6\alpha + \beta - 1)}_{=0} \|\nabla u\|^2 + \mathcal{O}(\|\nabla u\|^3) \\ &= -9\alpha + 3 + 4\alpha \operatorname{tr}(\varepsilon)^2 + 2 \operatorname{tr}(\varepsilon^2) + \mathcal{O}(\|\nabla u\|^3) \\ &= -9\alpha + 3 + 4\alpha \operatorname{tr}(\varepsilon)^2 + 2 (\|\operatorname{dev}_3 \varepsilon\|^2 + \frac{1}{3} \operatorname{tr}(\varepsilon)^2) + \mathcal{O}(\|\nabla u\|^3) \\ &= -9\alpha + 3 + \left(4\alpha + \frac{2}{3}\right) \operatorname{tr}(\varepsilon)^2 + 2 \|\operatorname{dev}_3 \varepsilon\|^2 + \mathcal{O}(\|\nabla u\|^3) \end{aligned}$$

■

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

In der isotropen nichtlinearen Elastizität ist der (symmetrische) Cauchy-Spannungstensor  $\sigma$  eine isotrope Funktion im Finger-Tensor  $B = FF^T \in \text{Sym}^+(3)$  (siehe auch Finger [27]). Wir haben gezeigt, dass nach dem Darstellungssatz (Lemma 2.5.28) dann Funktionen  $\beta_{-1}, \beta_0, \beta_1$  in den Invarianten von  $B$  existieren, sodass die Spannungsantwort in der Form

$$\sigma(B) = \beta_0 \mathbb{1} + \beta_1 B + \beta_{-1} B^{-1} \quad (6.0.1)$$

ausgedrückt werden kann. In Abschnitt 2.5.2 auf Seite 38 ff. haben wir bereits die sogenannten *empirischen Ungleichungen*

$$\beta_0 \leq 0, \quad \beta_1 > 0, \quad \beta_{-1} \leq 0 \quad (6.0.2)$$

für isotrope Funktionen ohne Kontext zur Elastizität kennengelernt. Konform zu den Ergebnissen von Truesdell haben wir dort gezeigt, dass die empirischen Ungleichungen die Semi-Invertierbarkeit der Abbildung  $B \mapsto \sigma(B)$  implizieren. Auch Destrade et al. beschreiben diese Implikation auf korrekte Weise in ihrem Artikel aus dem Jahr 2012 [23] über Shear-Deformationen. Sie erhalten die Semi-Invertierbarkeit aus einem Ergebnis von Johnson und Hoger [39, S. 192] und geben explizite Repräsentationsformeln<sup>1</sup> für die Funktionen  $\psi_0, \psi_1$  und  $\psi_2$  mit<sup>2</sup>

$$B = \psi_0 \mathbb{1} + \psi_1 \sigma + \psi_2 \sigma^2 \quad (6.0.3)$$

in Abhängigkeit von  $\beta_{-1}, \beta_0$  und  $\beta_1$ , an.

Sie nennen die Repräsentation von  $\sigma$ , die sie in ihren weiteren Untersuchungen des Zusammenhangs zwischen Simple-Shear-Deformationen und Simple-Shear-Spannungen, benutzen, die „Invertierbarkeit der Spannungs-Dehnungs-Beziehung“.

Diese Beschreibung ist jedoch irreführend, da die Semi-Invertierbarkeit im Allgemeinen *nicht* die Invertierbarkeit dieser Abbildung mit sich bringt, wie wir schon in Truesdells Gegenbeispiel (Beispiel 2.5.42) auf Seite 39 gesehen haben und dass sich mit der Spannungsantwort

$$\sigma(B) = B - \frac{1}{3} \text{tr}(B) \cdot \mathbb{1} = \text{dev}_3 B = \text{dev}_3(B - \mathbb{1}). \quad (6.0.4)$$

deckt. Sicherlich ist die konstitutive Beziehung (6.0.4) nicht vereinbar mit einem physikalisch brauchbaren Elastizitätsgesetz, denn es ist nicht hyperelastisch, das heißt es lässt kein elastisches Energiepotential zu und verstößt gegen die Eindeutigkeit der spannungsfreien Referenzkonfiguration (wegen  $\sigma(B) = 0$  für alle  $B = \lambda \cdot \mathbb{1}$  mit  $\lambda > 0$ ). Gleichwohl zeigt das Gegenbeispiel die Notwendigkeit von zusätzlichen Annahmen,

<sup>1</sup>Die von Johnson und Hoger [39, S. 192] vorgestellten Repräsentationsformeln lauten:

$$\begin{aligned} \psi_0 &= \frac{1}{\Delta} \left( \beta_0^2 - 2\beta_{-1}\beta_1 + I_2\beta_1^2 + I_3 \frac{\beta_0\beta_1^2}{\beta_{-1}} + I_1\beta_{-1}^2 + \frac{I_2}{I_3}\beta_0\beta_{-1} \right), & \psi_2 &= \frac{1}{\Delta}, \\ \psi_1 &= -\frac{1}{\Delta} \left( 2\beta_0 + I_3 \frac{\beta_1^2}{\beta_{-1}} + \frac{I_2}{I_3}\beta_{-1} \right), & \Delta &= I_1\beta_1^2 - I_3 \frac{\beta_1^3}{\beta_{-1}} + \frac{1}{I_3}\beta_{-1}^2 - \frac{I_2}{I_3}\beta_1\beta_{-1}. \end{aligned}$$

<sup>2</sup> Erneut sei darauf hingewiesen, dass die rechte Seite von (6.0.3) implizit von  $B$  durch  $\psi_0, \psi_1$  und  $\psi_2$ , welche Funktionen in den Invarianten

$$I_1(B) = \text{tr}(B), \quad I_2(B) = \text{tr}(\text{Cof } B) \quad \text{und} \quad I_3(B) = \det B,$$

sind, abhängt. Hierbei ist  $\text{Cof } B$  der Kofaktor von  $B$ .

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

um die Invertierbarkeit der Abbildung  $B \mapsto \sigma(B)$  sicherzustellen; die Semi-Invertierbarkeit allein ist dafür nicht ausreichend. Dennoch ist sie eine sehr erwünschte Eigenschaft.

Wir werden die Implikationskette die wir in Satz 2.5.56 auf Seite 42 beschrieben haben und mit der wir von den empirischen Ungleichungen auf die Semi-Invertierbarkeit schließen, in diesem Abschnitt auf den Kontext der Elastizität übertragen und so die Ergebnisse von Truesdell und Destrade et al. in diesem Zusammenhang verfeinern.

Weiterhin werden wir für hyperelastische Elastizitätsgesetze eine konkrete Formel für die Berechnung der  $\beta_{-1}, \beta_0, \beta_1$  aus (6.0.1) in Abhängigkeit der zugrundeliegenden Energiefunktion in ihren Invarianten vorstellen und unsere Ergebnisse über die SSLI ausnutzend zeigen, dass das durch  $W(U) = \|\log U\|^2$  induzierte Elastizitätsgesetz die empirischen Ungleichungen erfüllt und somit semi-invertierbar ist. Dies wird sich als Schlüssel erweisen für eine ganze Klasse von Energiefunktionen, welche wir als vom *Hencky-Typ* bezeichnen werden, eine einfache Bedingung an die jeweilige Funktion dieser Klasse anzugeben, mit der das von ihr induzierte Elastizitätsgesetz semi-invertierbar ist. Wir werden zeigen, dass die Bedingung für die sogenannte *quadratische Hencky-Energie*, sowie auch für die *Exponentiated-Hencky-Energie* erfüllt ist.

### 6.1. Semi-Invertierbarkeit von isotropen Elastizitätsgesetzen

Erinnern wir uns an Satz 2.5.56 von Seite 42. Zugeschnitten für das Elastizitätsgesetz  $\hat{\sigma}$  in  $B \in \text{Sym}^+(3)$  lautet er:

**Satz 6.1.1.** *Gegeben sei ein isotropes Elastizitätsgesetz  $\hat{\sigma}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  mit Darstellung*

$$\hat{\sigma}(B) = \beta_{-1} B^{-1} + \beta_0 \cdot \mathbb{1} + \beta_1 B \quad \text{für alle } B \in \text{Sym}^+(3). \quad (6.1.1)$$

Hierbei sind die  $\beta_{-1}, \beta_0, \beta_1$  reellwertige Funktionen in den Invarianten von  $B$ . Zu jedem  $B$  bezeichnen wir die Eigenwerte von  $B$  mit  $\lambda_1^2, \lambda_2^2, \lambda_3^2$  und die Eigenwerte von  $\hat{\sigma}(B)$  mit  $\sigma_1, \sigma_2, \sigma_3$ , wobei  $\hat{\sigma}(\text{diag}(\lambda_1^2, \lambda_2^2, \lambda_3^2)) = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$  gilt.

Dann gelten folgende Implikationen:<sup>3</sup>

(E-TSS)	$\Rightarrow$	(WE-TSS)	$\Rightarrow$	(BE <sup>+</sup> )	$\Rightarrow$	(bi-coax)	$\Leftrightarrow$	(semi)
					$\nRightarrow$	(invert)	$\Rightarrow$	

Hierbei bezeichnen:

(E-TSS):  $\hat{\sigma}$  erfüllt die empirischen Ungleichungen.

D.h. es gilt  $\beta_{-1} \leq 0, \beta_0 \leq 0, \beta_1 > 0$  für alle  $B$ .

(WE-TSS):  $\hat{\sigma}$  erfüllt die schwachen empirischen Ungleichungen.

D.h. es gilt  $\beta_{-1} \leq 0, \beta_1 \geq 0$  für alle  $B$ , wobei für jedes  $B$  eine der beiden Ungleichungen jeweils strikt ist.

(BE<sup>+</sup>):  $\hat{\sigma}$  erfüllt die strikten Baker-Ericksen-Ungleichungen.

D.h. es gilt  $\lambda_i > \lambda_j \Rightarrow \sigma_i > \sigma_j$  für alle  $i, j \in \{1, 2, 3\}$ .

(bi-coax):  $\hat{\sigma}$  ist bi-koaxial.

D.h. jeder Eigenvektor von  $B$  zum Eigenwert  $\lambda_i$  ist Eigenvektor von  $\hat{\sigma}(B)$  zur Hauptspannung  $\hat{\sigma}_i$  und umgekehrt.

---

<sup>3</sup>siehe auch Dunn [25]

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

(semi):  $\hat{\sigma}$  ist semi-invertierbar.

D.h. es gilt  $B = \psi_0 \cdot \mathbb{1} + \psi_1 \hat{\sigma}(B) + \psi_2 \hat{\sigma}(B)^2$  mit reellen Funktionen  $\psi_0, \psi_1, \psi_2$  in den Invarianten von  $B$ .

(invert):  $\hat{\sigma}$  ist invertierbar.

D.h. es gibt eine Umkehrfunktion  $\hat{\sigma}^{-1}: \text{Sym}(3) \rightarrow \text{Sym}^+(3)$  mit  $B = \hat{\sigma}^{-1}(\hat{\sigma}(B))$  für alle  $X \in \mathfrak{S}$  und  $Y = \hat{\sigma}(\hat{\sigma}^{-1}(Y))$  für alle  $Y \in \text{Sym}(n)$ .

**Bemerkung 6.1.2.** Wir können sämtliche der obigen Bedingungen und damit auch sämtliche der Implikationen auch **nur auf isotropen Teilmengen von  $\text{Sym}^+(3)$  definieren**. Wegen der Isotropie ist klar: Wenn eine Bedingung für zwei isotrope Teilmengen erfüllt ist, dann ist sie auch für deren Vereinigung erfüllt. Dieses so mögliche Zerlegen in Teilmengen hat den Vorteil, dass wenn wir eine Bedingung nur für eine Teilmenge von  $\text{Sym}^+(3)$  nachweisen können und die aus dieser Bedingung resultierende Folgerung für den übrigbleibenden Rest erfüllt ist, die Folgerung auf der ganzen Menge gilt! So bereiten uns in diesem Kapitel Vielfache der Einheitsmatrix an einigen Stellen Probleme, und das, obwohl für diese isotrope Teilmenge jedes isotrope Elastizitätsgesetz semi-invertierbar ist, wir also für das Schließen auf Semi-Invertierbarkeit diese Teilmenge gar nicht benötigen.

Denn gilt  $\lambda_1 = \lambda_2 = \lambda_3$ , so sind die Baker-Ericksen-Ungleichungen trivialerweise erfüllt:

**Proposition 6.1.3.** Jedes isotrope Elastizitätsgesetz  $\hat{\sigma}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  erfüllt auf der Teilmenge aller Vielfachen der Einheitsmatrix  $\{s \cdot \mathbb{1} \mid s \in \mathbb{R}_+\}$  die strikten Baker-Ericksen-Ungleichungen und ist damit auf dieser Teilmenge semi-invertierbar.

*Beweis.* Erfolgt direkt aus der Definition. ■

**Wollen wir also Semi-Invertierbarkeit zeigen, brauchen wir uns nur auf alle Matrizen zu konzentrieren, die nicht drei gleiche Eigenwerte besitzen!**

Die empirischen Ungleichungen und die schwachen empirischen Ungleichungen beziehen sich auf  $\beta_{-1}, \beta_0, \beta_1$  der Darstellung eines isotropen Elastizitätsgesetzes. Im hyperelastischen Fall haben wir anstatt dieser Darstellung die Energiefunktion gegeben. Mit dem folgenden Lemma, dass sich auch bei Saccomandi im Buch von Dorfmann und Ogden [24, S. 134] findet, können wir aus dieser direkt  $\beta_{-1}, \beta_0, \beta_1$  bestimmen.

**Lemma 6.1.4.** Gegeben sei eine isotrope elastische Energiefunktion  $\widehat{W}: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit  $B \mapsto \widehat{W}(B)$ . Das durch  $\widehat{W}$  induzierte isotrope Elastizitätsgesetz  $\hat{\sigma}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  ist von der Form

$$\hat{\sigma}(B) = \beta_{-1} B^{-1} + \beta_0 \cdot \mathbb{1} + \beta_1 B, \quad (6.1.2)$$

mit Funktionen  $\beta_{-1}, \beta_0, \beta_1: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  in den Invarianten von  $B$ , die eindeutig festgelegt sind durch

$$\beta_{-1} = -2\sqrt{I_3} \frac{\partial \widehat{\Psi}}{\partial I_2}, \quad \beta_0 = \frac{2}{\sqrt{I_3}} \left( I_2 \frac{\partial \widehat{\Psi}}{\partial I_2} + I_3 \frac{\partial \widehat{\Psi}}{\partial I_3} \right) \quad \text{und} \quad \beta_1 = \frac{2}{\sqrt{I_3}} \frac{\partial \widehat{\Psi}}{\partial I_1}. \quad (6.1.3)$$

Hierbei ist  $\widehat{\Psi}: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  mit  $\widehat{\Psi}(I_1, I_2, I_3) = \widehat{W}(B)$  die Energiefunktion in Abhängigkeit von den Invarianten von  $B$ .

*Beweis.* Wir erinnern uns: Für  $B \in \text{Sym}^+(3)$  sind die Invarianten  $I_1(B) = \text{tr}(B)$ ,  $I_2(B) = \text{tr}(\text{Cof } B) = \frac{1}{2} \text{tr}(B)^2 - \frac{1}{2} \text{tr}(B^2)$  und  $I_3(B) = \det B$ . Für ihre Ableitungen gilt:

$$DI_1(B) = \mathbb{1}, \quad DI_2(B) = I_1 \cdot \mathbb{1} - B \quad \text{und} \quad DI_3(B) = \text{Cof } B = I_3 B^{-1}.$$

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

Wir verwenden die Kettenregel, um einen Ausdruck für  $\hat{\sigma}(B)$  in den Invarianten von  $B$  zu erhalten:

$$\begin{aligned}\hat{\sigma}(B) &= \frac{2}{\sqrt{\det B}} D\widehat{W}(B) \cdot B \\ &= \frac{2}{\sqrt{I_3}} \sum_{k=1}^3 \frac{\partial \widehat{\Psi}}{\partial I_k}(I_1, I_2, I_3) \cdot DI_k(B) \cdot B \\ &= \frac{2}{\sqrt{I_3}} \left( \frac{\partial \widehat{\Psi}}{\partial I_1}(I_1, I_2, I_3) B + \frac{\partial \widehat{\Psi}}{\partial I_2}(I_1, I_2, I_3) (I_1 B - B^2) + \frac{\partial \widehat{\Psi}}{\partial I_3}(I_1, I_2, I_3) I_3 \cdot \mathbb{1} \right)\end{aligned}$$

Nach dem Satz von Cayley-Hamilton ist  $B$  Nullstelle seines charakteristischen Polynoms  $P_B(t) = t^3 - I_1 t^2 + I_2 t - I_3$ , also  $0 = B^3 - I_1 B^2 + I_2 B - I_3 \cdot \mathbb{1}$ . Umformen ergibt

$$I_1 B - B^2 = I_2 \cdot \mathbb{1} - I_3 B^{-1},$$

was wir direkt einsetzen

$$\begin{aligned}\hat{\sigma}(B) &= \frac{2}{\sqrt{I_3}} \left( \frac{\partial \widehat{\Psi}}{\partial I_1}(I_1, I_2, I_3) B + \frac{\partial \widehat{\Psi}}{\partial I_2}(I_1, I_2, I_3) (I_2 \cdot \mathbb{1} - I_3 B^{-1}) + \frac{\partial \widehat{\Psi}}{\partial I_3}(I_1, I_2, I_3) I_3 \cdot \mathbb{1} \right) \\ &= \left( -2\sqrt{I_3} \frac{\partial \widehat{\Psi}}{\partial I_2}(I_1, I_2, I_3) \right) B^{-1} + \frac{2}{\sqrt{I_3}} \left( I_2 \frac{\partial \widehat{\Psi}}{\partial I_2}(I_1, I_2, I_3) + I_3 \frac{\partial \widehat{\Psi}}{\partial I_3}(I_1, I_2, I_3) \right) \cdot \mathbb{1} \\ &\quad + \frac{2}{\sqrt{I_3}} \frac{\partial \widehat{\Psi}}{\partial I_1}(I_1, I_2, I_3) B \\ &= \beta_{-1} B^{-1} + \beta_0 \cdot \mathbb{1} + \beta_1 B\end{aligned}$$

mit  $\beta_{-1}, \beta_0, \beta_1$  wie in (6.1.3). ■

Mit der Kenntnis, wie wir  $\beta_{-1}$  und  $\beta_1$  direkt mittels  $I_3$  und  $\frac{\partial \widehat{\Psi}}{\partial I_1}, \frac{\partial \widehat{\Psi}}{\partial I_2}$  bestimmen können, bzw. noch spezieller: Mit der darauffolgenden Erkenntnis, dass  $\beta_1$  stets dasselbe Vorzeichen wie  $\frac{\partial \widehat{\Psi}}{\partial I_1}$  und  $\beta_{-1}$  stets das umgekehrte Vorzeichen von  $\frac{\partial \widehat{\Psi}}{\partial I_2}$  hat, können wir die (E2)-Bedingung  $\beta_{-1} \leq 0$  und  $\beta_1 \geq 0$  inklusive der Striktheitseigenschaft einfacher direkt als Eigenschaft der Energie in den Invarianten ausdrücken. Auch können wir die „unbequeme“ Konsequenz aus Bedingung  $\beta_0 \leq 0$  ziehen:

**Folgerung 6.1.5.** *Gegeben sei eine isotrope Energiefunktion  $\widehat{W}: \text{Sym}^+(3) \rightarrow \mathbb{R}$ ,  $B \mapsto \widehat{W}(B)$  mit Darstellung  $\widehat{\Psi}: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  in den Invarianten von  $B \in \text{Sym}^+(3)$ , d.h.  $\widehat{W}(B) = \widehat{\Psi}(I_1, I_2, I_3)$ . Dann erfüllt das aus  $\widehat{W}$  resultierende Elastizitätsgesetz  $\hat{\sigma}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  mit  $B \mapsto \hat{\sigma}(B)$  die schwachen empirischen Ungleichungen genau dann, wenn*

$$\frac{\partial \widehat{\Psi}}{\partial I_1}(I_1, I_2, I_3) \geq 0 \quad \text{und} \quad \frac{\partial \widehat{\Psi}}{\partial I_2}(I_1, I_2, I_3) \geq 0 \quad \text{für alle } B \in \text{Sym}^+(3) \quad (6.1.4)$$

*gilt und dabei für jedes  $B \in \text{Sym}^+(3)$  jeweils eine der Ungleichungen strikt ist. Erfüllt  $\hat{\sigma}$  die empirischen Ungleichungen, so folgt aus der  $\beta_0 \leq 0$ -Bedingung, dass*

$$\frac{\partial \widehat{\Psi}}{\partial I_3}(I_1, I_2, I_3) \leq 0 \quad \text{für alle } B \in \text{Sym}^+(3). \quad (6.1.5)$$

*gelten muss.*

*Beweis.* Der erste Teil ist aufgrund der Nichtnegativität der Invarianten klar. Zum zweiten Teil: Gelten die empirischen Ungleichungen, so ist einerseits nach Definition  $\beta_0 \leq 0$ , somit  $\frac{2}{\sqrt{I_3}} \left( I_2 \frac{\partial \widehat{\Psi}}{\partial I_2} + I_3 \frac{\partial \widehat{\Psi}}{\partial I_3} \right) \leq 0$ , als auch nach Teil eins  $I_2 \frac{\partial \widehat{\Psi}}{\partial I_2} \geq 0$ , woraus  $I_3 \frac{\partial \widehat{\Psi}}{\partial I_3} \leq 0$  folgt. ■

**Bemerkung 6.1.6.** Die „ $\beta_0 \leq 0$ “-Bedingung ist für die Gültigkeit der *schwachen empirischen Ungleichungen* (WE-TSS) mathematisch irrelevant. Mehr noch: für physikalisch sinnvolle Energien, welche die (WE-TSS) erfüllen und aus diesem Grund semi-invertierbar sind, ist die Eigenschaft  $\frac{\partial \hat{\Psi}}{\partial I_3} \leq 0$  als notwendige Eigenschaft für Semi-Invertierbarkeit physikalisch nicht zu rechtfertigen. Im Gegenteil: Wir erwarten, dass mit zunehmender volumetrischer Verzerrung auch die Energie wächst. Ein Beispiel, an der wir diese Eigenschaft beobachten können, ist die kompressible Mooney-Rivlin-Energie

$$W_{\text{MR}}(F) = \frac{\mu_1}{2} \left( \left\| \frac{F}{\det F^{1/3}} \right\|^2 - 3 \right) + \frac{\mu_2}{2} \left( \left\| \text{Cof} \frac{F}{\det F^{1/3}} \right\|^2 - 3 \right) + \frac{\gamma}{2} (\det F - 1)^2, \quad (6.1.6)$$

mit  $\mu_1, \mu_2, \gamma \in \mathbb{R}_+$ . Der volumetrische Term  $\frac{\gamma}{2}(\det F - 1)^2$  ist konvex in  $I_3$  mit Minimum bei  $I_3 = 1$  und dominiert bei großen Volumenverzerrungen den gesamten Energieausdruck. Wir zeigen in Beispiel 6.1.29 auf Seite 148: Die kompressible Mooney-Rivlin-Energie erfüllt die schwachen empirischen Ungleichungen und ist somit semi-invertierbar, aber zu jedem  $\mu_1, \mu_2, \gamma$  gibt es ein  $C \in \text{Sym}^+(3)$  mit  $\beta_0 > 0!$

Um nun zu zeigen, dass das durch  $W(U) = \|\log U\|^2$  induzierte Elastizitätsgesetz die (schwachen) empirischen Ungleichungen erfüllt, brauchen wir noch etwas Vorarbeit.

Zunächst einmal müssen wir uns klarmachen, dass die Darstellbarkeit von  $W$  in den Invarianten von  $U$  überhaupt gegeben ist. Hierzu erinnern wir uns an Lemma 2.13.25 auf Seite 66 und bringen dieses Lemma in den Kontext der nichtlinearen Elastizitätstheorie. Es sichert uns für hinreichend glatte Energiefunktionen die Darstellbarkeit in den Invarianten.

**Lemma 6.1.7.** *Sei  $W: \text{Sym}^+(3) \rightarrow \mathbb{R}$  eine isotrope, unendlich oft differenzierbare Energiefunktion, dann existiert eine unendlich oft differenzierbare Funktion  $\Psi: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  mit  $W(U) = \Psi(I_1, I_2, I_3)$  für alle  $U \in \text{Sym}^+(3)$ , wobei  $I_1 = \text{tr}(U)$ ,  $I_2 = \text{tr}(\text{Cof } U)$  und  $I_3 = \det U$  die Invarianten von  $U$  sind.*

**Bemerkung 6.1.8.** Es sollte klar sein, dass die in der Literatur kommentarlos beschriebene „Ableitung der Energie nach ihren Invarianten“ eine Eigenschaft von Lemma 6.1.7 ist. Andererseits legitimiert Lemma 6.1.7 auch diese Sorglosigkeit bei genügend glatten Energiefunktionen.

**Bemerkung 6.1.9.** Im **inkompressiblen Fall**, also bei Deformationen  $F \in \text{GL}^+(3)$  mit  $\det F = 1$ , erhalten wir mittels Lemma 6.1.4 die Koeffizientenfunktionen

$$\beta_{-1} = -2 \frac{\partial \hat{\Psi}}{\partial I_2}, \quad \beta_0 = 2 I_2 \frac{\partial \hat{\Psi}}{\partial I_2} \quad \text{und} \quad \beta_1 = 2 \frac{\partial \hat{\Psi}}{\partial I_1} \quad (6.1.7)$$

für das aus  $\widetilde{W}$  induziertes Elastizitätsgesetz  $\hat{\sigma}(B) = \beta_{-1} B^{-1} + \beta_0 \cdot \mathbb{1} + \beta_1 B$ . Zudem haben wir in Folgerung 2.13.35 auf Seite 69 gezeigt, dass für  $B \in \text{Sym}^+(3)$  mit  $I_3 = \det B = 1$  stets  $I_1 = \text{tr}(B) \geq 3$  und  $I_2 = \text{tr}(\text{Cof } B) \geq 3$  gilt, also für  $k \in \{2, 3\}$  eine Monotonie in  $I_k$  stets gleichbedeutend mit einer Monotonie in  $|I_k - 3|$  ist.<sup>4</sup>

Welche Bedeutung hat eine Monotonie in  $|I_k - 3|$ ? Geometrisch betrachtet bemisst  $4 \cdot |I_1(U) - 3| = 4 \cdot |\lambda_1 + \lambda_2 + \lambda_3 - 3|$  die Gesamtkantenverzerrung und  $2 \cdot |I_2(U) - 3| = 2 \cdot |\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3 - 3|$  die Gesamtflächenverzerrung eines mittels  $U$  verzerrten Einheitswürfels (siehe auch Folgerung 3.1.4). So entspricht monotonen Wachstum von  $\hat{\Psi}$  in  $I_1$  genau monotonem Wachstum in der Gesamtkantenverzerrung und monotonen Wachstum in  $I_2$  genau monotonem Wachstum in der Gesamtflächenverzerrung. Damit ist im inkompressiblen Fall die Bedingung  $\beta_{-1} \leq 0$  gleichbedeutend mit der Eigenschaft, dass die Energie  $\hat{\Psi}$  mit zunehmender Flächenverzerrung wächst und die Bedingung  $\beta_1 \geq 0$  gleichbedeutend der Eigenschaft, dass sie mit zunehmender Kantenverzerrung wächst.

### 6.1.1. Semi-Invertierbarkeit bei $W(U) = \|\log U\|^2$

Nun arbeiten wir darauf hin, zu zeigen, dass  $W(U) = \|\log U\|^2$  die schwachen empirischen Ungleichungen erfüllt. Wir werden sehen, dass wir die Menge aller Vielfachen der Einheitsmatrix aus der Menge  $\text{Sym}^+(3)$

<sup>4</sup>Da für  $\det U = 1$  die Ungleichungen  $I_1(U) \geq 3$  und  $I_2(U) \geq 3$  gelten, folgt  $I_k(U) \leq I_k(\hat{U}) \Leftrightarrow |I_k(U) - 3| \leq |I_k(\hat{U}) - 3|$  für alle  $k \in \{1, 2\}$  und alle  $U \in \text{Sym}^+(3)$  mit  $\det U = 1$ .

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

herausnehmen müssen, um das zu zeigen. Aber wie in Bemerkung 6.1.2 und Proposition 6.1.3 beschrieben, reicht uns dies, um damit die Semi-Invertierbarkeit von  $W(U) = \|\log U\|^2$  auf ganz  $\text{Sym}^+(3)$  zu beweisen.

Mit Folgerung 6.1.5 haben wir ein Kriterium herausgearbeitet, dass die Gültigkeit der schwachen empirischen Ungleichungen unter der Positivität der ersten beiden partiellen Ableitungen der Energie in den Invarianten<sup>5</sup> von  $C = U^2$  garantiert. Dass das Lemma sich auf die Invarianten von  $U$  (und nicht auf die von  $C$ ) bezieht, spielt mit Bemerkung 6.1.11 keine Rolle.

**Lemma 6.1.10.** *Sei  $W: \text{Sym}^+(3) \rightarrow \mathbb{R}$  eine Energiefunktion mit*

$$W(U) = \|\log U\|^2 \quad \text{für alle } U \in \text{Sym}^+(3) \quad (6.1.8)$$

und  $\Psi: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  die zugehörige Energiefunktion in den Invarianten von  $U$ , d.h.  $\Psi(I_1, I_2, I_3) = W(U)$ . Dann ist  $\Psi$  unendlich oft differenzierbar und für alle  $U$  gilt

$$\frac{\partial \Psi}{\partial I_1}(I_1, I_2, I_3) \geq 0 \quad \text{und} \quad \frac{\partial \Psi}{\partial I_2}(I_1, I_2, I_3) \geq 0, \quad (6.1.9)$$

wobei für alle  $U$  nicht beidseitig Gleichheit gelten kann, außer  $U$  ist Vielfaches der Einheitsmatrix. Beide Ungleichungen sind sogar strikt, wenn  $U$  nur paarweise unterschiedliche Eigenwerte hat.

**Bemerkung 6.1.11.** Mit gleicher Argumentation ist  $C \mapsto \|\log C\|^2$  für alle  $C \in \text{Sym}^+(3)$  als Funktion  $\widehat{\Psi}$  in den Invarianten von  $C$  mit positiven partiellen Ableitungen nach den Invarianten  $I_1$  und  $I_2$  von  $C$  darstellbar. Wegen  $\log C = 2 \log U$  ist

$$W(U) = \|\log U\|^2 = \frac{1}{4} \|\log C\|^2 \quad (6.1.10)$$

ebenso als Funktion  $W(U) = \widehat{\Psi}(I_1, I_2, I_3)$  in den **Invarianten von  $C$**  darstellbar, diese Darstellung unendlich oft differenzierbar und für alle  $U$  gilt

$$\frac{\partial \widehat{\Psi}}{\partial I_1}(I_1, I_2, I_3) \geq 0 \quad \text{und} \quad \frac{\partial \widehat{\Psi}}{\partial I_2}(I_1, I_2, I_3) \geq 0, \quad (6.1.11)$$

wobei für alle  $U$  nicht beidseitig Gleichheit gelten kann, außer  $U$  ist Vielfaches der Einheitsmatrix. Beide Ungleichungen sind sogar strikt, wenn  $U$  nur paarweise unterschiedliche Eigenwerte hat.

Um das Lemma zu beweisen, formulieren wir es zunächst um. Dabei beleuchten wir auch noch einmal die Darstellbarkeit in und Differenzierbarkeit nach den Invarianten:  $U \mapsto W(U)$  ist offenkundig isotrop, so können wir es auch als symmetrische Funktion  $f$  in den Eigenwerten von  $U$  schreiben, das heißt

$$W(U) = f(\lambda_1, \lambda_2, \lambda_3), \quad (6.1.12)$$

mit  $f(z_1, z_2, z_3) = \sum_{i=1}^3 \phi(z_i)$  und  $\phi(x) = (\log x)^2$ . Folgerung 2.13.24 von Seite 66 garantiert uns nun, dass es zu dieser symmetrischen Funktion  $f$  eine Darstellung in den elementarsymmetrischen Polynomen

$$\mathcal{F}: T \rightarrow \mathbb{R} \quad \text{mit} \quad f(z_1, \dots, z_n) = \mathcal{F}(e_1, \dots, e_n) \quad (6.1.13)$$

gibt, wobei  $T := \{(e_1, \dots, e_n) \in \mathbb{R}_+^n \mid (z_1, \dots, z_n) \in \mathbb{R}_+^n\}$  ist. Und sie sichert, dass wenn  $\phi$  unendlich oft differenzierbar ist, dass dann auch  $\mathcal{F}$  unendlich oft differenzierbar ist. Da nun aber  $\phi(x) = (\log x)^2$  eine glatte Funktion ist, haben wir einerseits die Differenzierbarkeit noch einmal bestätigt, aber auch den Weg geebnet für die folgende äquivalente Darstellung von Lemma 6.1.10.

**Lemma 6.1.12.** *Sei  $f: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  mit  $f(z_1, z_2, z_3) = \sum_{i=1}^3 (\log z_i)^2$ . Zudem sei*

$$\mathcal{F}: T \rightarrow \mathbb{R} \quad \text{mit} \quad f(z_1, z_2, z_3) = \mathcal{F}(e_1, e_2, e_3) \quad \text{und} \quad T := \{(e_1, e_2, e_3) \mid (z_1, z_2, z_3) \in \mathbb{R}_+^3\}$$

---

<sup>5</sup>Die Invarianten von  $B$  und von  $C$  sind identisch, daher spielt es keine Rolle, wenn wir in diesem Abschnitt die Darstellungen auf  $C$  beziehen.



## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

gegeben. Dann ist  $\mathcal{F}$  unendlich oft differenzierbar und es gilt für die partiellen Ableitungen

$$\frac{\partial \mathcal{F}}{\partial e_1}(e_1, e_2, e_3) \geq 0 \quad \text{und} \quad \frac{\partial \mathcal{F}}{\partial e_2}(e_1, e_2, e_3) \geq 0 \quad \text{für alle } (e_1, e_2, e_3) \in T, \quad (6.1.14)$$

sogar mit Striktheit für paarweise unterschiedliche  $(z_1, z_2, z_3)$ . Außerdem folgt aus

$$\frac{\partial \mathcal{F}}{\partial e_1}(e_1, e_2, e_3) = \frac{\partial \mathcal{F}}{\partial e_2}(e_1, e_2, e_3) = 0 \quad (6.1.15)$$

die Gleichheit aller drei Nullstellen  $z_1 = z_2 = z_3$ , d.h. es gibt kein  $(z_1, z_2, z_3) \neq (a, a, a)$ , das (6.1.15) erfüllt.

*Beweis von Lemma 6.1.10 und Lemma 6.1.12.* Mit der Differenzierbarkeit der reellen Logarithmusfunktion erhalten wir mit Folgerung 2.13.24 die Differenzierbarkeit der Funktion  $\mathcal{F}$ . Zusammen mit der Monotonie von  $\mathcal{F}$ , die der Aussage der SSLI entspricht, erhalten wir (6.1.14). Dass sogar die Striktheit für paarweise unterschiedliche Nullstellen gilt, war Teil des Beweises der SSLI.

Nun angenommen es gelte (6.1.15) und nicht  $z_1 = z_2 = z_3$ . Dazu sei o.B.d.A.  $(z_1, z_2, z_3) = (a, a, b)$  und  $a \neq b$ . Setzen wir

$$\gamma: \mathbb{R} \rightarrow \mathbb{R}_+^3 \quad \text{mit} \quad \gamma(t) = \left( a(1+t), a, \frac{b}{1+t} \right),$$

dann ist  $\gamma(0) = (a, a, b)$ . Sei nun  $\Gamma := e \circ \gamma$ , d.h.

$$\Gamma: (-1, 1) \rightarrow T \quad \text{mit} \quad \Gamma(t) = \left( e_1(\gamma(t)), e_2(\gamma(t)), e_3(\gamma(t)) \right),$$

also

$$\begin{aligned} \Gamma_1(t) &= a(1+t) + a + \frac{b}{1+t}, & \Gamma'_1(t) &= a - \frac{b}{(1+t)^2}, \\ \Gamma_2(t) &= a^2(1+t) + ab \frac{2+t}{1+t}, & \Gamma'_2(t) &= a^2 - \frac{ab}{(1+t)^2}, \\ \Gamma_3(t) &= a^2b, & \Gamma'_3(t) &= 0, \end{aligned}$$

und damit

$$\begin{aligned} \Gamma_1(0) &= 2a + b = e_1(a, a, b), \\ \Gamma_2(0) &= a^2 + 2ab = e_2(a, a, b), \\ \Gamma_3(0) &= a^2b = e_3(a, a, b) \quad \text{für alle } t \in \mathbb{R}. \end{aligned}$$

Nach unserer Definition ist  $\mathcal{F}(e_1, \dots, e_n) = f(z_1, \dots, z_n)$  also  $\mathcal{F}(\Gamma(t)) = f(\gamma(t))$  und damit

$$(\mathcal{F} \circ \Gamma)'(t) = (f \circ \gamma)'(t) \quad \text{für alle } t \in \mathbb{R}. \quad (6.1.16)$$

Es ist  $(\mathcal{F} \circ \Gamma)'(t) = \sum_{i=1}^n \frac{\partial \mathcal{F}}{\partial e_i}(\Gamma(t)) \cdot \Gamma'_i(t)$  und damit

$$(\mathcal{F} \circ \Gamma)'(0) = \underbrace{\frac{\partial \mathcal{F}}{\partial e_1}(a, a, b)}_{=0} \cdot \Gamma'_1(0) + \underbrace{\frac{\partial \mathcal{F}}{\partial e_2}(a, a, b)}_{=0} \cdot \Gamma'_2(0) + \frac{\partial \mathcal{F}}{\partial e_3}(a, a, b) \cdot \underbrace{\Gamma'_3(0)}_{=0} = 0.$$

Auf der anderen Seite ist

$$\begin{aligned} f(\gamma(t)) &= (\log a(1+t))^2 + (\log a)^2 + \left( \log \frac{b}{1+t} \right)^2 \\ &= (\log a + \log(1+t))^2 + (\log a)^2 + (\log b - \log(1+t))^2 \\ &= 2(\log a)^2 + (\log b)^2 + 2(\log a - \log b) \log(1+t) + 2(\log(1+t))^2 \end{aligned}$$

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

und damit

$$(f \circ \gamma)'(t) = 2(\log a - \log b) \frac{1}{1+t} + 4 \log(1+t) \frac{1}{1+t},$$

was auf  $(f \circ \gamma)'(0) = 2(\log a - \log b) \neq 0$  führt, ein Widerspruch zu (6.1.16). ■

Wir wissen nun also, die die Darstellung der Abbildung  $U \mapsto \|\log U\|^2$  in den Invarianten von  $C$  hat für alle  $U \in \text{Sym}^+(3)$ , welche nicht Vielfache der Einheitsmatrix sind, positive partielle Ableitungen nach  $I_1$  und  $I_2$ . Folgerung 6.1.5 garantiert uns nun die Gültigkeit der schwachen empirischen Ungleichungen. Bemerkung 6.1.2 und Proposition 6.1.3 sichern, dass dies den Schluss auf die Semi-Invertierbarkeit auf ganz  $\text{Sym}^+(3)$  zulässt. Insgesamt erhalten wir:

**Satz 6.1.13.** Die Energiefunktion  $W: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$W(U) = \|\log U\|^2 \quad \text{für alle } U \in \text{Sym}^+(3) \quad (6.1.17)$$

induziert ein Elastizitätsgesetz, dass für alle  $U \in \text{Sym}^+(3)$ , die nicht Vielfaches der Einheitsmatrix sind, die schwachen empirischen Ungleichungen (WE-TSS) erfüllt, für alle  $U \in \text{Sym}^+(3)$  die Baker-Ericksen-Inequalities (BE<sup>+</sup>) erfüllt und somit semi-invertierbar ist.

Wenn wir  $\|\log U\|^2$  noch mit einer monoton wachsenden Funktion  $f$  verketteten, erhalten wir immer noch ein Elastizitätsgesetz, welche die Baker-Ericksen-Inequalities erfüllt:

**Beispiel 6.1.14.** Sei  $f: \mathbb{R}_+ \rightarrow \mathbb{R}$  eine differenzierbare, streng monoton wachsende Funktion. Die Energiefunktion  $W: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$W(U) = f(\|\log U\|^2) \quad \text{für alle } U \in \text{Sym}^+(3) \quad (6.1.18)$$

induziert ein Elastizitätsgesetz, das ebenso wie Satz 6.1.13 (WE-TSS) und (BE<sup>+</sup>) erfüllt und semi-invertierbar ist.

*Beweis.* Für  $k \in \{1, 2\}$  und alle  $U \in \text{Sym}^+(3) \setminus \mathbb{R} \cdot \mathbb{1}$  ist

$$\frac{d}{dI_k} f(\|\log U\|^2) = \frac{d}{dI_k} f\left(\frac{1}{4} \|\log C\|^2\right) = \underbrace{f'\left(\frac{1}{4} \|\log C\|^2\right)}_{>0} \cdot \frac{d}{dI_k} \|\log C\|^2 \geq 0,$$

mit ansonsten analoger Argumentation wie bei Satz 6.1.13. ■

### 6.1.2. Energiefunktionen vom Hencky-Typ

Wir haben in diesem Kapitel bereits ausgenutzt, dass wir isotrope Energiefunktionen  $F \mapsto W(F)$  als Funktionen in den Invarianten von  $C = F^T F$  darstellen können. In diesem Abschnitt wollen wir nun Energiefunktionen untersuchen, die sich als Funktion in der Norm des spurfreien Anteils und im Betrag der Spur des Hencky-Dehnungstensors  $\log U = \log \sqrt{F^T F}$  darstellen lassen.

**Definition 6.1.15.** Wir bezeichnen eine elastische Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  als vom *Hencky-Typ*, wenn sich  $W$  als Funktion in  $\|\text{dev}_3 \log U\|^2$  und  $|\text{tr} \log U|^2$  schreiben lässt, d.h. wenn  $W$  von der Form

$$W(F) = \mathcal{W}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) \quad (6.1.19)$$

mit einer differenzierbaren Funktion  $\mathcal{W}: \mathbb{R}_+^2 \rightarrow \mathbb{R}$  ist.

**Bemerkung 6.1.16.** Offenkundig sind elastische Energiefunktionen vom *Hencky-Typ* isotrop, da  $\log$ ,  $\text{dev}_n$ ,  $\text{tr}$  und die Frobeniusnorm isotrop sind. Und aus diesem Grund können wir, wie wir es von Energiefunktionen gewohnt sind, gleichwertig  $W(F) = \mathcal{W}(\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2)$  mit  $\log V = \log \sqrt{F F^T}$  schreiben.

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

So, wie wir Darstellungssätze von isotropen Funktionen als Ausdrücke in den Ableitungen nach den Invarianten ausdrücken können, ist es uns möglich, für Funktionen vom Hencky-Typ eine Darstellung in den partiellen Ableitungen von  $\mathcal{W}$  anzugeben:

**Lemma 6.1.17.** *Sei  $\sigma: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  ein Elastizitätsgesetz, welches von einer Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  vom Hencky-Typ induziert wird. Dann gilt:*

$$\begin{aligned} \sigma(V) = \frac{2}{\det V} & \left( \frac{\partial \mathcal{W}}{\partial x_1} (\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \text{dev}_3 \log V \right. \\ & \left. + \frac{\partial \mathcal{W}}{\partial x_2} (\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \text{tr} \log V \cdot \mathbf{1} \right). \end{aligned} \quad (6.1.20)$$

*Beweis.* Wir setzen  $g: \mathbb{R}^3 \rightarrow \mathbb{R}$ , sodass  $W(F) = W(\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) = g(\lambda_1, \lambda_2, \lambda_3)$  gilt und bestimmen mit der Kettenregel:

$$\begin{aligned} \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) &= \frac{\partial \mathcal{W}}{\partial x_1} (\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \frac{d}{d\lambda_i} \|\text{dev}_3 \log V\|^2 \\ &+ \frac{\partial \mathcal{W}}{\partial x_2} (\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \frac{d}{d\lambda_i} |\text{tr} \log V|^2 \\ &= \frac{\partial \mathcal{W}}{\partial x_1} (\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \frac{2}{\lambda_i} \log \frac{\lambda_i}{(\det U)^{1/n}} \\ &+ \frac{\partial \mathcal{W}}{\partial x_2} (\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \frac{2}{\lambda_i} \text{tr} \log V. \end{aligned}$$

So lauten nach Lemma 4.3.28 auf Seite 113 die Hauptspannungen (principal stresses)

$$\begin{aligned} \sigma_i = \frac{\lambda_i}{\det U} \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) &= \frac{2}{\det U} \left( \frac{\partial \mathcal{W}}{\partial x_1} (\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \log \frac{\lambda_i}{(\det U)^{1/n}} \right. \\ & \left. + \frac{\partial \mathcal{W}}{\partial x_2} (\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \text{tr} \log V \right). \end{aligned}$$

Mit  $\sigma(V) = Q \text{diag}(\sigma_1, \sigma_2, \sigma_3)Q^T$  für  $U = Q \text{diag}(\lambda_1, \lambda_2, \lambda_3)Q^T$  erhalten wir dann das zu Zeigende. ■

Der partiellen Ableitung nach dem ersten Argument von  $\mathcal{W}$  kommt eine besondere Bedeutung zu. Wir werden mit den beiden folgenden Lemmata zeigen, dass die Semi-Invertierbarkeit des Elastizitätsgesetzes einer Energiefunktion vom Hencky-Typ nur von dieser ersten Ableitung abhängt. Dies ist der Schlüssel, um für die typischen Vertreter der Energiefunktionen vom Hencky-Typ die Semi-Invertierbarkeit der korrespondierenden Elastizitätsgesetze nachzuweisen.

Wir zeigen nun, dass eine Energiefunktion vom Hencky-Typ unter positiver partieller Ableitung von  $\mathcal{W}$  nach der ersten Komponente das Kriterium aus Folgerung 6.1.5 für die Gültigkeit der schwachen empirischen Ungleichungen erfüllt. Wie auch schon im letzten Abschnitt können wir alle  $U$  mit drei gleichen Eigenwerten bei der Argumentation auslassen.

**Lemma 6.1.18.** *Gegeben sei eine isotrope elastische Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  vom Hencky-Typ mit positiver partieller Ableitung nach  $\|\text{dev}_3 \log U\|^2$ , d.h.*

$$\frac{\partial \mathcal{W}}{\partial x_1} (\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) > 0 \quad \text{für alle } U \in \text{Sym}^+(n), \quad (6.1.21)$$

dann gilt für alle  $U \in \text{Sym}^+(n)$ , wobei  $U$  kein Vielfaches der Einheitsmatrix ist

$$\frac{\partial \widehat{\Psi}}{\partial I_1}(I_1, I_2, I_3) \geq 0 \quad \text{und} \quad \frac{\partial \widehat{\Psi}}{\partial I_2}(I_1, I_2, I_3) \geq 0. \quad (6.1.22)$$

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

Hierbei ist  $\widehat{\Psi}$  die Darstellung von  $W$  in den Invarianten  $I_1, I_2, I_3 \in \mathbb{R}_+$  von  $C = U^2$ . Es kann nicht beidseitig Gleichheit gelten. Beide Ungleichungen sind strikt, wenn  $U$  nur paarweise unterschiedliche Eigenwerte hat.

Insbesondere erfüllt dann  $W$  die schwachen empirischen Ungleichungen für alle  $U \in \text{Sym}^+(3)$ , die nicht Vielfache der Einheitsmatrix sind.

*Beweis.* Wir leiten  $\widehat{\Psi}$  nach den Invarianten  $I_k$  von  $C$  ab, aufgrund von  $\log U = \frac{1}{2} \log C$  können wir leicht umrechnen. Mit  $\|\log C\|^2 = \|\text{dev}_n \log C\|^2 + \frac{1}{n} |\text{tr} \log C|^2$  und der Kettenregel gilt

$$\begin{aligned}
 \frac{\partial \widehat{\Psi}}{\partial I_k}(U) &= \frac{d\mathcal{W}}{dI_k}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) \\
 &= \frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) \cdot \frac{d}{dI_k} \|\text{dev}_3 \log U\|^2 \\
 &\quad + \frac{\partial \mathcal{W}}{\partial x_2}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) \cdot \frac{d}{dI_k} |\text{tr} \log U|^2 \\
 &= \frac{1}{4} \cdot \frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) \cdot \frac{d}{dI_k} (\|\log C\|^2 - \frac{1}{3} |\text{tr} \log C|^2) \\
 &\quad + \frac{1}{4} \cdot \frac{\partial \mathcal{W}}{\partial x_2}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) \cdot \frac{d}{dI_k} |\text{tr} \log C|^2 \\
 &= \frac{1}{4} \cdot \frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) \cdot \frac{d}{dI_k} \|\log C\|^2 \\
 &\quad + \frac{1}{4} \cdot \left( -\frac{1}{3} \frac{\partial \mathcal{W}}{\partial x_1} + \frac{\partial \mathcal{W}}{\partial x_2} \right) (\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) \cdot \frac{d}{dI_k} |\text{tr} \log C|^2.
 \end{aligned}$$

Für  $k \in \{1, 2\}$  ist  $\frac{d}{dI_k} |\text{tr} \log C|^2 = \frac{d}{dI_k} |\log I_3|^2 = 0$ . Weiterhin ist nach Voraussetzung die erste partielle Ableitung  $\partial \mathcal{W} / \partial x_1$  positiv. Wir erhalten

$$\frac{d\mathcal{W}}{dI_k}(U) = \underbrace{\frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2)}_{>0} \cdot \frac{d}{dI_k} \|\log C\|^2 \geq 0 \quad \text{für } k \in \{1, 2\}.$$

Die Nichtnegativität des Ausdrucks erben wir aus der Ungleichung  $\frac{d}{dI_k} \|\log C\|^2 \geq 0$  für  $k \in \{1, 2\}$  aus Lemma 6.1.10 des letzten Abschnitts. Wir haben dort sogar gezeigt, dass die Ungleichung im Falle paarweise verschiedener Eigenwerte von  $C$  für  $k = 1$  und für  $k = 2$  strikt ist und ansonsten für eins der beiden  $k \in \{1, 2\}$  strikt ist. ■

Die weitere Argumentation erfolgt nun analog wie im Satz 6.1.13: Wir haben gezeigt, dass die Darstellung eines Elastizitätsgesetz vom Hencky-Typ in den Invarianten von  $C$  für alle  $U \in \text{Sym}^+(3) \setminus \mathbb{R} \cdot \mathbf{1}$  positive partielle Ableitungen nach  $I_1$  und  $I_2$  unter der Voraussetzung hat, dass die zugehörige Energiefunktion vom Hencky-Typ  $\mathcal{W}$  über eine positive erste Ableitung verfügt. Bemerkung 6.1.2 und Proposition 6.1.3 sichern uns, dass dies den Schluss auf die Semi-Invertierbarkeit auf ganz  $\text{Sym}^+(3)$  zulässt. Insgesamt erhalten wir:

**Satz 6.1.19.** *Gegeben sei eine isotrope elastische Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  vom Hencky-Typ mit positiver partieller Ableitung nach  $\|\text{dev}_3 \log U\|^2$ , d.h.*

$$\frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) > 0 \quad \text{für alle } U \in \text{Sym}^+(3). \tag{6.1.23}$$

Dann erfüllt das aus  $W$  resultierende Elastizitätsgesetz  $\widehat{\sigma}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  für alle  $B \in \text{Sym}^+(3)$ , die nicht Vielfaches der Einheitsmatrix sind, die schwachen empirischen Ungleichungen (WE-TSS), für alle  $B \in \text{Sym}^+(3)$  die Baker-Ericksen-Inequalities (BE<sup>+</sup>) und ist somit semi-invertierbar.

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

Als erstes Beispiel einer Energiefunktion von Hencky-Typ wollen wir die *quadratische Hencky-Energie* nennen. Sie ist sozusagen der Ursprung der Hencky-Typ-Energien und wurde 1929 von H. Hencky [37] im Kontext der nichtlinearen Elastizitätstheorie eingeführt. Jüngst wurde sie durch eine eindeutige geometrische Eigenschaft charakterisiert [51, 56, 57]: Sie misst die quadratische geometrische Distanz von  $F$  zur speziellen orthogonalen Gruppe  $\text{SO}(3)$  in einer links  $\text{GL}^+(n)$ -invarianten, rechts  $\text{SO}(n)$ -invarianten Riemann'schen Metrik auf  $\text{GL}^+(n)$ .

**Beispiel 6.1.20.** Die *quadratische Hencky-Energie*  $W_{\text{H}}: \text{GL}^+(3) \rightarrow \mathbb{R}$  mit

$$W_{\text{H}}(F) = \mu \|\text{dev}_3 \log U\|^2 + \frac{\kappa}{2} (\text{tr}(\log U))^2 = \mu \|\log U\|^2 + \frac{\Lambda}{2} (\text{tr}(\log U))^2. \quad (6.1.24)$$

wobei die Materialparameter  $\mu, \Lambda \in \mathbb{R}$  mit  $\mu > 0$  und  $3\Lambda + 2\mu \geq 0$  *Lamé-Parameter* und  $\kappa \in \mathbb{R}_+$  *Bulk-Modulus* genannt werden, erfüllt für alle  $U = \sqrt{F^T F}$ , die nicht Vielfache der Einheitsmatrix sind, die schwachen empirischen Ungleichungen und für alle  $U \in \text{Sym}^+(3)$  die Baker-Ericksen-Inequalities und ist semi-invertierbar. Das von  $W_{\text{H}}$  induzierte Elastizitätsgesetz  $\sigma: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  lautet

$$\sigma(V) = \frac{2}{\det V} \left( \mu \text{dev}_3 \log V + \frac{\kappa}{2} \text{tr} \log V \cdot \mathbf{1} \right). \quad (6.1.25)$$

*Beweis.* Die quadratische Hencky-Energie  $W_{\text{eH}}(U)$  ist vom Hencky-Typ mit

$$\mathcal{W}(x_1, x_2) = \mu x_1 + \frac{\kappa}{2} x_2,$$

und hat die partiellen Ableitungen

$$\frac{\partial \mathcal{W}}{\partial x_1}(x_1, x_2) = \mu \quad \text{und} \quad \frac{\partial \mathcal{W}}{\partial x_2}(x_1, x_2) = \frac{\kappa}{2}.$$

Insbesondere lässt sich Satz 6.1.19 wegen der positiven partiellen Ableitung

$$\frac{dW}{d\|\text{dev}_3 \log U\|^2}(U) = \frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) = \mu > 0$$

anwenden. Unter Zuhilfenahme von Lemma 6.1.17 lautet das Elastizitätsgesetz

$$\begin{aligned} \sigma(V) &= \frac{2}{\det V} \left( \frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \text{dev}_3 \log V \right. \\ &\quad \left. + \frac{\partial \mathcal{W}}{\partial x_2}(\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \text{tr} \log V \cdot \mathbf{1} \right) \\ &= \frac{2}{\det V} \left( \mu \text{dev}_3 \log V + \frac{\kappa}{2} \text{tr} \log V \cdot \mathbf{1} \right). \quad \blacksquare \end{aligned}$$

**Bemerkung 6.1.21.** Das von der quadratischen Hencky-Energie induzierte Elastizitätsgesetz ist zwar semi-invertierbar, aber nicht invertierbar. Dies lässt sich leicht anhand der Nichtinjektivität der Abbildung

$$x \mapsto \sigma(x^{1/n} \cdot \mathbf{1}) = \underbrace{\frac{2}{\det(x^{1/n} \cdot \mathbf{1})}}_{=x} \left( \underbrace{\mu \text{dev}_3 \log(x^{1/n} \cdot \mathbf{1})}_{=0} \right) + \frac{\kappa}{2} \left( \underbrace{\log \det(x^{1/n} \cdot \mathbf{1}) \cdot \mathbf{1}}_{=x} \right) = \kappa \frac{\log x}{x} \quad (6.1.26)$$

für  $x \in \mathbb{R}_+$  erkennen.<sup>6</sup>

Als nächstes Beispiel präsentieren wir die von Neff, Ghiba und Lankeit [57] eingeführte *Exponentiated-Hencky-Energie*.

<sup>6</sup>Die Abbildung ist stetig und es gilt  $\frac{\log 1}{1} = 0$ ,  $\frac{\log e}{e} = \frac{1}{e}$  und  $\lim_{x \rightarrow \infty} \frac{\log x}{x} = 0$ .

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

**Beispiel 6.1.22.** Die *Exponentiated-Hencky-Energie*  $W_{\text{eH}}: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$W_{\text{eH}}(U) = \frac{\mu}{k} e^{k \|\text{dev}_3 \log U\|^2} + \frac{\kappa}{2\widehat{k}} e^{\widehat{k} |\text{tr} \log U|^2}, \quad \mu, k, \kappa, \widehat{k} \in \mathbb{R}_+ \quad (6.1.27)$$

erfüllt für alle  $U \in \text{Sym}^+(3) \setminus \mathbb{R} \cdot \mathbb{1}$  die schwachen empirischen Ungleichungen und für alle  $U \in \text{Sym}^+(3)$  die strikten Baker-Ericksen-Inequalities und ist semi-invertierbar. Das von  $W_{\text{eH}}$  induzierte Elastizitätsgesetz  $\sigma: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  lautet

$$\sigma(V) = \frac{2}{\det V} \left( \mu e^{k \|\text{dev}_3 \log V\|^2} \text{dev}_3 \log V + \frac{\kappa}{2} e^{\widehat{k} |\text{tr} \log V|^2} \text{tr} \log V \cdot \mathbb{1} \right). \quad (6.1.28)$$

*Beweis.* Die Exponentiated-Hencky-Energie  $W_{\text{eH}}(U)$  ist vom Hencky-Typ mit

$$\mathcal{W}(x_1, x_2) = \frac{\mu}{k} e^{k x_1} + \frac{\kappa}{2\widehat{k}} e^{\widehat{k} x_2}$$

und hat die partielle Ableitungen

$$\frac{\partial \mathcal{W}}{\partial x_1}(x_1, x_2) = \mu e^{k x_1} \quad \text{und} \quad \frac{\partial \mathcal{W}}{\partial x_2}(x_1, x_2) = \frac{\kappa}{2} e^{\widehat{k} x_2}.$$

Insbesondere lässt sich Satz 6.1.19 wegen der positiven partiellen Ableitung

$$\frac{dW}{d\|\text{dev}_3 \log U\|^2}(U) = \frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) = \mu e^{k \|\text{dev}_3 \log U\|^2} > 0.$$

anwenden. Unter Zuhilfenahme von Lemma 6.1.17 lautet das Elastizitätsgesetz

$$\begin{aligned} \sigma(V) &= \frac{2}{\det V} \left( \frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \text{dev}_3 \log V \right. \\ &\quad \left. + \frac{\partial \mathcal{W}}{\partial x_2}(\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \text{tr} \log V \cdot \mathbb{1} \right) \\ &= \frac{2}{\det V} \left( \mu e^{k \|\text{dev}_3 \log V\|^2} \text{dev}_3 \log V + \frac{\kappa}{2} e^{\widehat{k} |\text{tr} \log V|^2} \text{tr} \log V \cdot \mathbb{1} \right). \quad \blacksquare \end{aligned}$$

**Bemerkung 6.1.23.** Das Ergebnis der Semi-Invertierbarkeit des aus der Exponentiated-Hencky-Energie hervorgehenden Elastizitätsgesetzes erhalten wir auch als Konsequenz aus der Invertierbarkeit des Elastizitätsgesetzes. Den Beweis dazu finden wir in Neff, Ghiba und Lankeit [57, S. 25 f].

**Beispiel 6.1.24.** Eine *Variation der Exponentiated-Hencky-Energie*  $W_{\text{eH}}^\sim: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$W_{\text{eH}}^\sim(U) = \frac{\mu}{k} \exp \left( \frac{k}{\mu} \cdot (\mu \|\text{dev}_3 \log U\|^2 + \frac{\kappa}{2} |\text{tr} \log U|^2) \right) \quad (6.1.29)$$

erfüllt für alle  $U \in \text{Sym}^+(3) \setminus \mathbb{R} \cdot \mathbb{1}$  die schwachen empirischen Ungleichungen und für alle  $U \in \text{Sym}^+(3)$  die Baker-Ericksen-Ungleichungen und ist semi-invertierbar. Das von  $W_{\text{eH}}$  induzierte Elastizitätsgesetz  $\sigma: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  lautet

$$\sigma(V) = \frac{1}{\det V} \frac{2k}{\mu} \cdot W_{\text{eH}}(V) \cdot \left( \mu \text{dev}_3 \log V + \frac{\kappa}{2} \text{tr} \log V \cdot \mathbb{1} \right). \quad (6.1.30)$$

*Beweis.* Diese Variation der Exponentiated-Hencky-Energie  $W_{\text{eH}}^\sim(U)$  ist vom Hencky-Typ mit

$$\mathcal{W}(x_1, x_2) = \frac{\mu}{k} \exp \left( \frac{k}{\mu} (\mu x_1 + \frac{\kappa}{2} x_2) \right)$$

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

und hat die partielle Ableitungen

$$\frac{\partial \mathcal{W}}{\partial x_1}(x_1, x_2) = k \cdot \mathcal{W}(x_1, x_2) \quad \text{und} \quad \frac{\partial \mathcal{W}}{\partial x_2}(x_1, x_2) = \frac{k\kappa}{2\mu} \cdot \mathcal{W}(x_1, x_2).$$

Insbesondere lässt sich Satz 6.1.19 wegen der positiven partiellen Ableitung

$$\frac{dW}{d\|\text{dev}_3 \log U\|^2}(U) = \frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) = k \cdot W_{\text{eH}}(U) > 0$$

anwenden. Unter Zuhilfenahme von Lemma 6.1.17 lautet das Elastizitätsgesetz

$$\begin{aligned} \sigma(V) &= \frac{2}{\det V} \left( \frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \text{dev}_3 \log V \right. \\ &\quad \left. + \frac{\partial \mathcal{W}}{\partial x_2}(\|\text{dev}_3 \log V\|^2, |\text{tr} \log V|^2) \cdot \text{tr} \log V \cdot \mathbf{1} \right) \\ &= \frac{2}{\det V} (k W_{\text{eH}}(V) \cdot \text{dev}_3 \log V + \frac{k\kappa}{2\mu} W_{\text{eH}}(V) \cdot \text{tr} \log V \cdot \mathbf{1}) \\ &= \frac{1}{\det V} \frac{2k}{\mu} \cdot W_{\text{eH}}(V) \cdot \left( \mu \text{dev}_3 \log V + \frac{\kappa}{2} \text{tr} \log V \cdot \mathbf{1} \right). \quad \blacksquare \end{aligned}$$

### 6.1.3. Energien mit additiv isochor-volumetrischem Split

**Definition 6.1.25.** Eine Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  weist einen *additiv isochor-volumetrischen Split* auf, wenn es stetig differenzierbare Funktionen  $W_{\text{iso}}: \text{SL}(3) \rightarrow \mathbb{R}$  und  $f: \mathbb{R}_+ \rightarrow \mathbb{R}$  gibt, sodass

$$W(F) = W_{\text{iso}} \left( \frac{F}{(\det F)^{1/3}} \right) + f(\det F) \quad \text{für alle } F \in \text{GL}^+(3). \quad (6.1.31)$$

Siehe auch Richter [66].

**Proposition 6.1.26.** Für jede Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$ , die einen *additiv isochor-volumetrischen Split* aufweist, lässt sich auch die Darstellung  $W(F) = \widehat{\Psi}(I_1, I_2, I_3)$  in den Invarianten von  $C = F^T F$  *additiv isochor-volumetrisch aufteilen*, d.h. es existiert eine Funktion  $\widehat{\Psi}_{\text{iso}}: \mathbb{R}^2 \rightarrow \mathbb{R}$  in den Invarianten von  $\widetilde{C} = \frac{C}{(\det C)^{1/3}}$ , dem *isochoren Anteil* von  $C = F^T F$ , mit

$$W(F) = \widehat{\Psi}_{\text{iso}}(\widetilde{I}_1, \widetilde{I}_2) + f(I_3^{1/2}) \quad \text{für alle } F \in \text{GL}^+(3). \quad (6.1.32)$$

Hierbei sind

$$I_1 = \text{tr}(C), \quad I_2 = \text{tr}(\text{Cof } C), \quad I_3 = \det C \quad (6.1.33)$$

die Invarianten von  $C$  und

$$\widetilde{I}_1 = \text{tr}(\widetilde{C}) = \frac{I_1}{I_3^{1/3}}, \quad \widetilde{I}_2 = \text{tr}(\text{Cof } \widetilde{C}) = \frac{I_2}{I_3^{2/3}}, \quad \widetilde{I}_3 = \det \widetilde{C} = 1 \quad (6.1.34)$$

die Invarianten von  $\widetilde{C}$  mit  $\widetilde{F} = \frac{F}{\det F^{1/3}}$  und  $\widetilde{C} = \widetilde{F}^T \widetilde{F}$ .

*Beweis.* Die Funktion  $W_{\text{iso}}$  ist isotrop und hat für jedes Argument  $\widetilde{F} \in \text{SL}(3)$  eine Darstellung in den Invarianten von  $\widetilde{C}$

$$W_{\text{iso}}(\widetilde{F}) = \widehat{\Psi}_{\text{iso}}(\widetilde{I}_1, \widetilde{I}_2, \widetilde{I}_3).$$

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

Die Invarianten von  $\tilde{C} = \tilde{F}^T \tilde{F} = \frac{F^T F}{(\det F^T)^{1/3} (\det F)^{1/3}} = \frac{C}{\det C^{1/3}}$  ergeben sich als

$$\begin{aligned}\tilde{I}_1 &= \operatorname{tr}(\tilde{C}) = \operatorname{tr}\left(\frac{C}{\det C^{1/3}}\right) = \frac{1}{\det C^{1/3}} \operatorname{tr}(C) = \frac{I_1}{I_3^{1/3}} \\ \tilde{I}_2 &= \operatorname{tr}(\operatorname{Cof} \tilde{C}) = \operatorname{tr}\left(\frac{1}{\det C^{2/3}} \operatorname{Cof} C\right) = \frac{I_2}{I_3^{2/3}} \\ \tilde{I}_3 &= \det \tilde{C} = \det \frac{C}{\det C^{1/3}} = 1.\end{aligned}$$

Bei  $\hat{I}_2$  verwenden wir die Umformung  $\operatorname{Cof} \tilde{C} = \det \tilde{C} \cdot \tilde{C}^{-1} = \frac{1}{\det C} \det C \frac{C^{-1}}{\det C^{-1/3}} = \frac{1}{\det C^{2/3}} \operatorname{Cof} C$ . Da  $\tilde{I}_3 = 1$  konstant ist, hängt  $\hat{\Psi}_{\text{iso}}$  nicht von  $\tilde{I}_3$  ab. Somit können wir ebenso  $W_{\text{iso}}(F) = \hat{\Psi}_{\text{iso}}(\tilde{I}_1, \tilde{I}_2)$  schreiben. ■

**Bemerkung 6.1.27.** Die Abbildung  $F \mapsto W_{\text{iso}}\left(\frac{F}{(\det F)^{1/3}}\right)$  lässt sich zwar als Funktion in den Invarianten  $\tilde{I}_1, \tilde{I}_2$ , aber im Allgemeinen **nicht** als Funktion in den Invarianten  $I_1, I_2$  (und Ausschluss von  $I_3$ !) ausdrücken. Machen wir uns klar: Die Abbildung ist invariant gegenüber rein volumetrischen Änderungen von  $F$ . Diese Änderung bedeutet allerdings eine Veränderung aller drei Invarianten  $I_1, I_2$  und  $I_3$ .<sup>7</sup> Halten wir  $I_1$  und  $I_2$  fest und verändern bloß  $I_3$ , so bedeutet dies auch eine Formänderung von  $F$  und damit eine Änderung der isochoren Energie  $W_{\text{iso}}\left(\frac{F}{(\det F)^{1/3}}\right)$ ! Und machen wir uns insbesondere klar, die Ableitung einer Energie mit additiv isochor-volumetrischem Split nach der dritten Invarianten  $I_3 = \det C$  hängt im Allgemeinen nicht nur vom volumetrischen Teil  $f(\det F)$ , sondern auch sehr wohl vom isochoren Teil  $W_{\text{iso}}\left(\frac{F}{(\det F)^{1/3}}\right)$  ab!

**Proposition 6.1.28.** Sei  $W: \operatorname{GL}^+(3) \rightarrow \mathbb{R}$  eine Energiefunktion, die einen additiv isochor-volumetrischen Split aufweist und eine Darstellung in den Invarianten der Form (6.1.32) aufweist. Sind die partiellen Ableitungen von  $\hat{\Psi}_{\text{iso}}$  positiv, d.h.

$$\frac{\partial \hat{\Psi}_{\text{iso}}}{\partial \tilde{I}_1}(\tilde{I}_1, \tilde{I}_2) \geq 0 \quad \text{und} \quad \frac{\partial \hat{\Psi}_{\text{iso}}}{\partial \tilde{I}_2}(\tilde{I}_1, \tilde{I}_2) \geq 0 \quad \text{für alle } C \in \operatorname{Sym}^+(3), \quad (6.1.35)$$

wobei für jedes  $C \in \operatorname{Sym}^+(3)$  jeweils eine der beiden Ungleichungen strikt ist, genau dann erfüllt  $\hat{W}$  die schwachen empirischen Ungleichungen (WE-TSS).

*Beweis.* Wie in der vorigen Proposition hergeleitet, habe nun  $W$  die Darstellung

$$W(F) = \hat{\Psi}_{\text{iso}}(\tilde{I}_1, \tilde{I}_2) + f(I_3^{1/2}) \quad \text{mit} \quad \tilde{I}_1 = \frac{I_1}{I_3^{1/3}}, \quad \tilde{I}_2 = \frac{I_2}{I_3^{2/3}}.$$

Ableitung mit der Kettenregel ergibt

$$\frac{\partial \hat{\Psi}}{\partial I_1}(I_1, I_2, I_3) = \frac{1}{I_3^{1/3}} \frac{\partial \hat{\Psi}_{\text{iso}}}{\partial \tilde{I}_1}(\tilde{I}_1, \tilde{I}_2) \quad \text{und} \quad \frac{\partial \hat{\Psi}}{\partial I_2}(I_1, I_2, I_3) = \frac{1}{I_3^{2/3}} \frac{\partial \hat{\Psi}_{\text{iso}}}{\partial \tilde{I}_2}(\tilde{I}_1, \tilde{I}_2).$$

Nach Folgerung 6.1.5 erfüllt  $W$  genau dann die schwachen empirischen Ungleichungen, wenn

$$\frac{\partial \hat{\Psi}}{\partial I_1}(I_1, I_2, I_3) \geq 0 \quad \text{und} \quad \frac{\partial \hat{\Psi}}{\partial I_2}(I_1, I_2, I_3) \geq 0 \quad \text{für alle } C \in \operatorname{Sym}^+(3)$$

gilt, wobei für jedes  $C \in \operatorname{Sym}^+(3)$  jeweils eine der Ungleichungen strikt sein muss. Mit der Positivität von  $\frac{1}{I_3^{1/3}}$  und von  $\frac{1}{I_3^{2/3}}$  folgt die Aussage. ■

<sup>7</sup>Für eine reine Volumenänderung  $F \mapsto cF$  gilt:  $I_1(cF) = cI_1(F)$ ,  $I_2(cF) = c^2 I_2(F)$ ,  $I_3(cF) = c^3 I_3(F)$ , es ändern sich alle drei Invarianten!



## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

Führen wir uns vor Augen: Sei  $F \in \text{GL}^+(3)$  und damit  $\tilde{F} := \frac{F}{\det F^{1/3}} \in \text{SL}(3)$ . Dann lauten die Invarianten von  $\tilde{C} = \frac{C}{\det C^{1/3}} = \tilde{F}^T \tilde{F}$

$$\begin{aligned}\tilde{I}_1 &= \text{tr}(\tilde{C}) = \text{tr}(\tilde{F}^T \tilde{F}) = \|\tilde{F}\|^2, \\ \tilde{I}_2 &= \text{tr}(\text{Cof } \tilde{C}) = \text{tr}(\text{Cof } \tilde{F}^T \tilde{F}) = \text{tr}(\text{Cof } \tilde{F})^T (\text{Cof } \tilde{F}) = \|\text{Cof } \tilde{F}\|^2, \\ \tilde{I}_3 &= \det \tilde{C} = 1.\end{aligned}$$

Vor diesem Hintergrund ist es uns ein Leichtes, die kompressible *Mooney-Rivlin-Energie* [70] und auch die kompressible *Neo-Hooke-Energie* als Beispiele für Energien mit additiv isochor-volumetrischem Split zu erkennen und direkt die Darstellung in den Invarianten von  $\tilde{C}$  anzugeben<sup>8</sup>.

**Beispiel 6.1.29.** Die kompressible Mooney-Rivlin-Energie  $W_{\text{MR}}: \text{GL}^+(3) \rightarrow \text{Sym}(3)$  mit

$$W_{\text{MR}}(F) = \frac{\mu_1}{2} \left( \left\| \frac{F}{\det F^{1/3}} \right\|^2 - 3 \right) + \frac{\mu_2}{2} \left( \left\| \text{Cof } \frac{F}{\det F^{1/3}} \right\|^2 - 3 \right) + \frac{\gamma}{2} (\det F - 1)^2, \quad \mu_1, \mu_2, \gamma \in \mathbb{R}_+ \quad (6.1.36)$$

weist einen additiv isochor-volumetrischen Split auf und hat die Darstellung in den Invarianten von  $C = F^T F$

$$W_{\text{MR}}(F) = \hat{\Psi}(I_1, I_2, I_3) = \hat{\Psi}_{\text{iso}}(\tilde{I}_1, \tilde{I}_2) + f(I_3^{1/2}) \quad (6.1.37)$$

mit

$$\hat{\Psi}_{\text{iso}}(\tilde{I}_1, \tilde{I}_2) = \frac{\mu_1}{2}(\tilde{I}_1 - 3) + \frac{\mu_2}{2}(\tilde{I}_2 - 3) \quad \text{und} \quad f(I_3^{1/2}) = \frac{\gamma}{2}(I_3^{1/2} - 1)^2. \quad (6.1.38)$$

Hierbei sind  $\tilde{I}_1 = \frac{I_1}{I_3^{1/3}}$ ,  $\tilde{I}_2 = \frac{I_2}{I_3^{2/3}}$  und  $I_1, I_2, I_3$  die Invarianten von  $C = F^T F$ . Die kompressible Mooney-Rivlin-Energie erfüllt für alle  $\mu_1, \mu_2, \gamma \in \mathbb{R}_+$  die **schwachen empirischen Ungleichungen**, aber **nicht Truesdells empirische Ungleichungen**, da die „ $\beta_0 \leq 0$ “-Bedingung verletzt ist.

*Beweis.* Mit  $\left\| \frac{F}{\det F^{1/3}} \right\|^2 = \|\tilde{F}\|^2 = \tilde{I}_1$  und  $\left\| \text{Cof } \frac{F}{\det F^{1/3}} \right\|^2 = \|\text{Cof } \tilde{F}\|^2 = \tilde{I}_2$  erhalten wir leicht die Darstellung  $\hat{\Psi}_{\text{iso}}(\tilde{I}_1, \tilde{I}_2)$ .

Der Nachweis für die schwachen empirischen Ungleichungen wird uns mit Proposition 6.1.28 leicht gemacht: Es ist  $\frac{\partial \hat{\Psi}}{\partial I_1} = \frac{\mu_1}{2} > 0$  und  $\frac{\partial \hat{\Psi}}{\partial I_2} = \frac{\mu_2}{2} > 0$ .

Bestimmen wir nun  $\beta_0$ . Hier verwenden wir die Repräsentationsformel aus Lemma 6.1.4 und leiten dann per Kettenregel ab

$$\begin{aligned}\beta_0 &= \frac{2}{\sqrt{I_3}} \left( I_2 \frac{\partial \hat{\Psi}}{\partial I_2} + I_3 \frac{\partial \hat{\Psi}}{\partial I_3} \right) \\ &= \frac{2}{\sqrt{I_3}} \left( I_2 \frac{d}{dI_2} \hat{\Psi}_{\text{iso}} \left( \frac{I_1}{I_3^{1/3}}, \frac{I_2}{I_3^{2/3}} \right) + I_3 \frac{d}{dI_3} \hat{\Psi}_{\text{iso}} \left( \frac{I_1}{I_3^{1/3}}, \frac{I_2}{I_3^{2/3}} \right) + I_3 \frac{d}{dI_3} f(I_3^{1/2}) \right) \\ &= \frac{2}{\sqrt{I_3}} \left( \frac{I_2}{I_3^{2/3}} \frac{\partial \hat{\Psi}_{\text{iso}}}{\partial \tilde{I}_2}(\tilde{I}_1, \tilde{I}_2) - \frac{1}{3} \frac{I_1}{I_3^{1/3}} \frac{\partial \hat{\Psi}_{\text{iso}}}{\partial \tilde{I}_1}(\tilde{I}_1, \tilde{I}_2) - \frac{2}{3} \frac{I_1}{I_3^{2/3}} \frac{\partial \hat{\Psi}_{\text{iso}}}{\partial \tilde{I}_2}(\tilde{I}_1, \tilde{I}_2) + \frac{\sqrt{I_3}}{2} f'(I_3^{1/2}) \right) \\ &= -\frac{2}{3} \frac{I_1}{I_3^{5/6}} \frac{\partial \hat{\Psi}_{\text{iso}}}{\partial \tilde{I}_1}(\tilde{I}_1, \tilde{I}_2) - \frac{2}{I_3^{7/6}} \left( \frac{2}{3} I_1 - I_2 \right) \frac{\partial \hat{\Psi}_{\text{iso}}}{\partial \tilde{I}_2}(\tilde{I}_1, \tilde{I}_2) + f'(I_3^{1/2}) \\ &= -\frac{2}{3} \frac{I_1}{I_3^{5/6}} \frac{\mu_1}{2} - \frac{2}{I_3^{7/6}} \left( \frac{2}{3} I_1 - I_2 \right) \frac{\mu_2}{2} + \gamma (I_3^{1/2} - 1)\end{aligned}$$

Dieser Ausdruck ist für große Volumenverzerrungen stets positiv. Zu vorgegebenen  $\mu_1, \mu_2, \gamma \in \mathbb{R}_+$  betrachten wir als Beispiel  $C = c \cdot \mathbb{1}$  mit  $c \in \mathbb{R}_+$ . Dann ist

$$\beta_0 = (3\mu_2 - \mu_1)c^{-3/2} - 2\mu_2 c^{-5/2} + \gamma c^{3/2} - \gamma \rightarrow \infty \quad \text{für} \quad c \rightarrow \infty.$$

Somit gilt für jede Wahl von  $\mu_1, \mu_2, \gamma \in \mathbb{R}_+$  **nicht**  $\beta_0 \leq 0$  und (E-TSS) ist nicht erfüllt. ■

<sup>8</sup>Eine Diskussion hierzu findet sich auch in Truesdell und Noll [79, S. 348 ff].

6. Semi-Invertierbarkeit und die empirischen Ungleichungen

**Beispiel 6.1.30.** Die kompressible *Neo-Hooke-Energie*  $W_{\text{NH}}: \text{GL}^+(3) \rightarrow \text{Sym}(3)$  mit

$$W_{\text{NH}}(F) = \frac{\mu}{2} \left( \left\| \frac{F}{\det F^{1/3}} \right\|^2 - 3 \right) + \frac{\gamma}{2} (\det F - 1)^2, \quad \mu, \gamma \in \mathbb{R}_+ \quad (6.1.39)$$

ist ein Spezialfall der kompressiblen Mooney-Rivlin-Energie mit  $\mu_1 = \mu$  und  $\mu_2 = 0$ .

**Beispiel 6.1.31.** Eine isochore Version der *Saint-Venant-Kirchhoff-Energie* (vgl. Beispiel 6.1.34) erhalten wir durch  $\widehat{W}: \text{GL}^+(3) \rightarrow \text{Sym}(3)$  mit

$$\widehat{W}(C) = \left\| \frac{C}{\det C^{1/3}} - \mathbf{1} \right\|^2. \quad (6.1.40)$$

Sie hat die Darstellung als isochore Funktion  $\widehat{\Psi}_{\text{iso}}$  in den Invarianten  $\tilde{I}_1 = \frac{I_1}{I_3^{1/3}}, \tilde{I}_2 = \frac{I_2}{I_3^{2/3}}$  von  $\tilde{C} = \frac{C}{\det C^{1/3}} \in \text{SL}(3)$

$$\widehat{W}(C) = \widehat{\Psi}_{\text{iso}}(\tilde{I}_1, \tilde{I}_2) = \tilde{I}_1^2 - 2\tilde{I}_1 - 2\tilde{I}_2 + 3 \quad (6.1.41)$$

und erfüllt wegen  $\frac{\partial \widehat{\Psi}_{\text{iso}}}{\partial \tilde{I}_2}(\tilde{I}_1, \tilde{I}_2) = -2 < 0$  laut Proposition 6.1.28 nicht die schwachen empirischen Ungleichungen.

Auch die folgende Energie weist einen additiv isochor-volumetrischen Split auf:

**Beispiel 6.1.32.** Die Energie  $\widehat{W}: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$\widehat{W}(C) = \frac{\mu}{4} \left\| \frac{C}{\det C^{1/3}} - \frac{C^{-1}}{\det C^{-1/3}} \right\|^2 + f(\det C), \quad \mu > 0. \quad (6.1.42)$$

weist einen additiv isochor-volumetrischen Split auf und hat die Darstellung in den Invarianten von  $C = F^T F$

$$\widehat{W}(C) = \widehat{\Psi}_{\text{iso}}(\tilde{I}_1, \tilde{I}_2) + f(I_3) \quad (6.1.43)$$

mit

$$\widehat{\Psi}_{\text{iso}}(\tilde{I}_1, \tilde{I}_2) = \frac{\mu}{4} (\tilde{I}_1^2 - 2\tilde{I}_2 + 2\tilde{I}_1 - \tilde{I}_2^2 - 6). \quad (6.1.44)$$

Diese Energie erfüllt laut Proposition 6.1.28 nicht die schwachen empirischen Ungleichungen, denn es gilt für die partiellen Ableitungen

$$\frac{\partial \widehat{\Psi}_{\text{iso}}}{\partial \tilde{I}_1}(\tilde{I}_1, \tilde{I}_2) = \frac{\mu}{2} (\tilde{I}_1 + 1) > 0 \quad \text{und} \quad \frac{\partial \widehat{\Psi}_{\text{iso}}}{\partial \tilde{I}_2}(\tilde{I}_1, \tilde{I}_2) = -\frac{\mu}{2} (\tilde{I}_2 + 1) < 0. \quad (6.1.45)$$

*Beweis zur Darstellung in den Invarianten.* Wir können die Energie auch schreiben als

$$\widehat{W}(C) = \frac{\mu}{4} \|\tilde{C} - \tilde{C}^{-1}\|^2 + f(\det C)$$

mit  $\tilde{C} = \frac{C}{\det C^{1/3}}$  und Invarianten  $\tilde{I}_1, \tilde{I}_2, \tilde{I}_3 = 1$  von  $\tilde{C}$ . Da nun nach Proposition 4.3.32 auf Seite 121

$$\|C\|^2 = I_1^2 - 2I_2 \quad \text{und} \quad \|C^{-1}\|^2 = 2\frac{I_1}{I_3} - \frac{I_2^2}{I_3^2}$$

gilt, können wir diese Ausdrücke auf  $\tilde{C}$  übertragen und erhalten

$$\|\tilde{C}\|^2 = \tilde{I}_1^2 - 2\tilde{I}_2 \quad \text{und} \quad \|\tilde{C}^{-1}\|^2 = 2\tilde{I}_1 - \tilde{I}_2^2,$$

da  $\tilde{I}_3 = 1$ . Mit  $\|\tilde{C} - \tilde{C}^{-1}\|^2 = \|\tilde{C}\|^2 - 6 + \|\tilde{C}^{-1}\|^2$  ist schließlich

$$\widehat{\Psi}_{\text{iso}}(\tilde{I}_1, \tilde{I}_2) = \frac{\mu}{4} ((\tilde{I}_1^2 - 2\tilde{I}_2) - 6 + (2\tilde{I}_1 - \tilde{I}_2^2)). \quad \blacksquare$$

### 6.1.4. Andere Energien

Mit ähnlicher Rechnung wie in Beispiel 6.1.32 erhalten wir:

**Beispiel 6.1.33.** Die Energie  $\widehat{W}: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$\widehat{W}(C) = \frac{\mu}{4} \|C - C^{-1}\|^2 + f(\det C), \quad \mu > 0 \quad (6.1.46)$$

hat die Darstellung in den Invarianten von  $C = F^T F$

$$\widehat{\Psi}(I_1, I_2, I_3) = \frac{\mu}{4} \left( I_1^2 - 2I_2 - 6 + 2 \frac{I_1}{I_3} - \frac{I_2^2}{I_3^2} \right) + f(I_3). \quad (6.1.47)$$

Diese Energie erfüllt laut Folgerung 6.1.5 nicht die schwachen empirischen Ungleichungen, denn es gilt für die partiellen Ableitungen

$$\frac{\partial \widehat{\Psi}}{\partial I_1}(I_1, I_2, I_3) = \frac{\mu}{2} \left( I_1 + \frac{1}{I_3} \right) > 0 \quad \text{und} \quad \frac{\partial \widehat{\Psi}}{\partial I_2}(I_1, I_2, I_3) = -\frac{\mu}{2} \left( 1 + \frac{I_2}{I_3} \right) < 0. \quad (6.1.48)$$

In Beispiel 6.1.31 haben wir bereits eine isochore Version (mit  $\mu = 4$  und  $\Lambda = 0$ ) der folgenden Energie kennengelernt. Sie erfüllt nicht die schwachen empirischen Ungleichungen. Die nicht-isochore Version, die wir in Proposition 5.4.2 bereits als *Saint-Venant-Kirchhoff-Energie* kennengelernt<sup>9</sup> haben, steht ihr dabei in Nichts nach.

**Beispiel 6.1.34.** Die *Saint-Venant-Kirchhoff-Energie*  $\widehat{W}_{\text{SVK}}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  mit

$$\widehat{W}_{\text{SVK}}(C) = \frac{\mu}{4} \|C - \mathbb{1}\|^2 + \frac{\Lambda}{8} (\text{tr}(C - \mathbb{1}))^2, \quad 3\Lambda + 2\mu > 0 \quad (6.1.49)$$

hat die Darstellung in den Invarianten von  $C = F^T F$

$$\widehat{\Psi}(I_1, I_2, I_3) = \frac{\mu}{4} (I_1^2 - 2I_2 - 2I_1 + 3) + \frac{\Lambda}{8} (I_1 - 3)^2. \quad (6.1.50)$$

Das aus dieser Energie resultierende Elastizitätsgesetz erfüllt laut Folgerung 6.1.5 nicht die schwachen empirischen Ungleichungen, denn es gilt für die partiellen Ableitungen

$$\frac{\partial \widehat{\Psi}}{\partial I_1}(I_1, I_2, I_3) = \frac{\mu}{4} (2I_1 - 2) + \frac{\Lambda}{4} (I_1 - 3) \quad \text{und} \quad \frac{\partial \widehat{\Psi}}{\partial I_2}(I_1, I_2, I_3) = -\frac{1}{2} \mu < 0. \quad (6.1.51)$$

Hierbei ist auch die Bedingungen  $\frac{\partial \widehat{\Psi}}{\partial I_1}(I_1, I_2, I_3) > 0$  für jedes  $\Lambda$  und  $\mu$  und hinreichend kleinem  $I_1$  nicht erfüllt.

*Zur Darstellung in den Invarianten.* Wir formen um

$$\begin{aligned} \widehat{W}_{\text{SVK}}(C) &= \frac{\mu}{4} \|C - \mathbb{1}\|^2 + \frac{\Lambda}{8} (\text{tr}(C - \mathbb{1}))^2 \\ &= \frac{\mu}{4} (\|C\|^2 - 2 \text{tr}(C) + 3) + \frac{\Lambda}{8} (\text{tr}(C) - 3)^2 \end{aligned}$$

Nun ist  $\|C\|^2 = I_1^2 - 2I_2$  und so erhalten wir

$$\widehat{\Psi}(I_1, I_2, I_3) = \frac{\mu}{4} (I_1^2 - 2I_2 - 2I_1 + 3) + \frac{\Lambda}{8} (I_1 - 3)^2. \quad \blacksquare$$

<sup>9</sup>Wir haben in Proposition 5.4.2 auf Seite 129 die Konvexität in  $C$  der Saint-Venant-Kirchhoff Energie gezeigt.

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

Nicht nur die schwachen empirischen Ungleichungen werden vom aus der Saint-Venant-Kirchhoff-Energie resultierenden Elastizitätsgesetz nicht erfüllt: Die bedeutend schwächeren Baker-Ericksen-Ungleichungen werden ebenso nicht erfüllt. Um dies zu zeigen, benötigen wir zunächst die Darstellung des Elastizitätsgesetzes als  $\hat{\sigma}$  in  $B$ :

**Proposition 6.1.35.** *Das durch die Saint-Venant-Kirchhoff-Energie induzierte Elastizitätsgesetz  $\hat{\sigma}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  lautet*

$$\hat{\sigma}(B) = \frac{1}{\sqrt{\det B}} \left( \mu B^2 + \frac{\Lambda}{2} \text{tr}(B) \cdot B - \frac{1}{2} (2\mu + 3\Lambda) B \right). \quad (6.1.52)$$

*Beweis.* Die Ableitung der Saint-Venant-Kirchhoff-Energie

$$\widehat{W}_{\text{SVK}}(B) = \frac{\mu}{4} (\text{tr}(B^2) - 2 \text{tr}(B) + 3) + \frac{\Lambda}{8} \left( (\text{tr}(B))^2 - 6 \text{tr}(B) + 9 \right)$$

lautet

$$\begin{aligned} D\widehat{W}_{\text{SVK}}(B) &= \frac{\mu}{4} (2C - 2 \cdot \mathbb{1}) + \frac{\Lambda}{8} (2 \text{tr}(B) \cdot \mathbb{1} - 6 \cdot \mathbb{1}) \\ &= \frac{\mu}{2} (B - \mathbb{1}) + \frac{\Lambda}{4} (\text{tr}(B) - 3) \cdot \mathbb{1} = \frac{\mu}{2} C + \frac{\Lambda}{4} \text{tr}(C) \cdot \mathbb{1} - \frac{1}{4} (2\mu + 3\Lambda) \cdot \mathbb{1}. \end{aligned}$$

Nach Lemma 4.3.27 auf Seite 113 gilt  $\hat{\sigma}(B) = \frac{2}{\sqrt{\det B}} D\widehat{W}(B) B$ . Einsetzen der Ableitung schließt den Beweis. ■

**Proposition 6.1.36.** *Das durch die Saint-Venant-Kirchhoff-Energie induzierte Elastizitätsgesetz  $\hat{\sigma}$  erfüllt nicht die strikten Baker-Ericksen Ungleichungen ( $\text{BE}^+$ ).*

*Beweis.* Sind  $\lambda_1^2, \lambda_2^2, \lambda_3^2 \in \mathbb{R}_+$  die Eigenwerte von  $B \in \text{Sym}^+(3)$ , dann gilt nach Proposition 6.1.35 für die Hauptspannungen

$$\sigma_i = \frac{1}{\lambda_1 \lambda_2 \lambda_3} \left( \mu \lambda_i^4 + \frac{\Lambda}{2} (\lambda_1^2 + \lambda_2^2 + \lambda_3^2) \lambda_i^2 - \frac{1}{2} (2\mu + 3\Lambda) \lambda_i^2 \right).$$

Nun angenommen,  $\hat{\sigma}$  würde die strikten Baker-Ericksen Ungleichungen erfüllen, dann würde nach Satz 6.1.1 auf Seite 135 für alle  $B \in \text{Sym}^+(3)$  mit Eigenwerten  $\lambda_1^2, \lambda_2^2, \lambda_3^2 \in \mathbb{R}_+$  und Hauptspannungen  $\sigma_1, \sigma_2, \sigma_3$  (also den Eigenwerten von  $\hat{\sigma}(B)$ ) gelten: Aus  $\lambda_i > \lambda_j$  folgt  $\sigma_i > \sigma_j$  für alle  $i, j \in \{1, 2, 3\}$ .

Zu vorgegebenen  $\mu \in \mathbb{R}_+$  und  $\Lambda \in \mathbb{R}$  mit  $3\Lambda + 2\mu > 0$  können wir  $x, y \in \mathbb{R}_+$  so klein wählen, dass

$$2(\mu + \Lambda)x + \Lambda y - (2\mu + 3\Lambda) < 0$$

gilt. Für hinreichend kleines  $\varepsilon > 0$  wählen wir nun  $\lambda_1 = \sqrt{x + \varepsilon}$ ,  $\lambda_2 = \sqrt{x - \varepsilon}$  und  $\lambda_3 = \sqrt{y}$ . Dann gilt  $\lambda_1 > \lambda_2$  und

$$\begin{aligned} \sigma_1 - \sigma_2 &= \frac{1}{(x^2 - \varepsilon^2)y} (\mu(x + \varepsilon)^2 - \mu(x - \varepsilon)^2 + \Lambda(2x + y)\varepsilon - (2\mu + 3\Lambda)\varepsilon) \\ &= \frac{1}{(x^2 - \varepsilon^2)y} (2(\mu + \Lambda)x + \Lambda y - (2\mu + 3\Lambda))\varepsilon < 0, \end{aligned}$$

also  $\sigma_1 < \sigma_2$ , ein Widerspruch. ■

In Ciarlet und Geymonat [17] findet sich die folgende Energiefunktion.

**Beispiel 6.1.37.** Die Ciarlet-Geymonat-Energie  $\widehat{W}: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$\widehat{W}_{\text{CG}}(C) = \frac{\mu}{2} (\text{tr}(C) - \log \det C - 3) + \frac{\Lambda}{4} (\det C - \log \det C - 1) \quad (6.1.53)$$

## 6. Semi-Invertierbarkeit und die empirischen Ungleichungen

hat die Darstellung in den Invarianten von  $C$

$$\widehat{\Psi}(I_1, I_2, I_3) = \frac{\mu}{2}(I_1 - \log I_3 - 3) + \frac{\Lambda}{4}(I_3 - \log I_3 - 1). \quad (6.1.54)$$

Das aus ihr resultierende Elastizitätsgesetz erfüllt die schwachen empirischen Ungleichungen (und somit auch die strikten Baker-Ericksen-Ungleichungen) und ist folglich semi-invertierbar, erfüllt aber nicht die „ $\beta_0 \leq 0$ “-Bedingung von Truesdells empirischer Ungleichungen.

*Beweis.* Die Gültigkeit der (WE-TSS) ist durch  $\frac{\partial \widehat{\Psi}}{\partial I_1} = \frac{\mu}{2} > 0$  und  $\frac{\partial \widehat{\Psi}}{\partial I_2} = 0$  mittels Folgerung 6.1.5 sichergestellt. Zudem erhalten wir mit Lemma 6.1.4 und  $\frac{\partial \widehat{\Psi}}{\partial I_3} = \frac{\Lambda}{4} - \left(\frac{\mu}{2} + \frac{\Lambda}{4}\right) I_3^{-1}$

$$\beta_0 = \frac{2}{\sqrt{I_3}} \left( I_2 \frac{\partial \widehat{\Psi}}{\partial I_2} + I_3 \frac{\partial \widehat{\Psi}}{\partial I_3} \right) = 2\sqrt{I_3} \left( \frac{\Lambda}{4} - \left( \frac{\mu}{2} + \frac{\Lambda}{4} \right) I_3^{-1} \right),$$

ein Ausdruck, der genau für  $I_3 > \frac{1}{2} \frac{\mu}{\Lambda} + 1$  positiv ist. ■

## 7. Shear

Was ist eine Scherung? Dazu betrachten wir das Szenario eines in  $e_1$ -Richtung und  $e_3$ -Richtung unendlich ausgedehnten Quaders eines homogenen elastischen Materials, der durch äußere Scherkräfte (auf der Oberseite in  $e_1$ -Richtung, auf der Unterseite in  $-e_1$ -Richtung) deformiert wird.

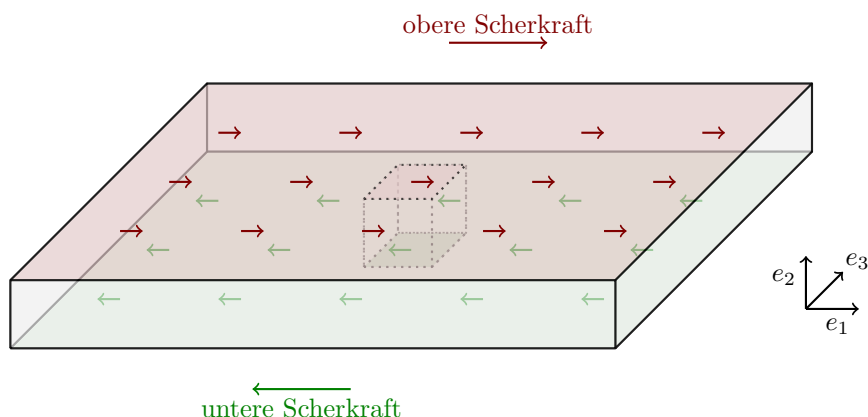


Abbildung 7.1.: Scherung

Hierbei wird eine Rotation des Quaders durch seine unendliche Ausdehnung verhindert. Diese verhindert auch, dass bis auf die Scherkräfte oben und unten sonstige Randbedingungen wirken. Durch die Homogenität (vgl. Abschnitt 4.3.2 auf Seite 103) können wir nun die gesamte Deformation untersuchen, indem wir die Einschränkung auf einen Teilwürfel betrachten. Es muss einen symmetrischen Spannungstensor  $T$  geben, der sämtliche Flächenkräfte, die auf den Teilwürfel wirken, charakterisiert: Auf einer Fläche mit Normale  $n$  wirkt die Kraft<sup>1</sup>  $T \cdot n$ . Die äußeren Scherkräfte erzwingen  $T \cdot e_2 = (s, 0, 0)^T$  bzw.  $T \cdot (-e_2) = (-s, 0, 0)$ , was die zweite Spalte von  $T$  festlegt.<sup>2</sup> Mit der Symmetrie von  $T$  ist auch seine zweite Zeile festgelegt. Weiterhin nehmen wir  $T \cdot e_3 = (0, 0, 0)^T$  an, was insgesamt auf den Spannungstensor

$$T = \begin{pmatrix} 0 & s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (7.0.1)$$

führt, den wir als *Pure-Shear-Spannung* bezeichnen.

So verstehen wir unter einer „Scherung“ klassischerweise eine *Simple-Shear-Deformation*, welche Ogden [62, S. 103] als volumenerhaltende und planare<sup>3</sup> Deformation, wie in Abbildung 7.2 zu sehen, beschreibt.

Wir werden in diesem Kapitel zeigen, dass unter Voraussetzung des linearen Modells jede Simple-Shear-Deformation von jedem linearen Elastizitätsgesetz auf eine Pure-Shear-Spannung abgebildet wird. Weiter werden wir zeigen, dass im Falle des finiten Modells jede (nichttriviale) Simple-Shear-Deformation mittels eines isotropen Elastizitätsgesetzes weder auf eine Pure-Shear-Cauchy-Spannung, noch auf eine Pure-Shear-Biot-Spannung (vergleiche dazu Abbildung 7.7) abgebildet wird.

<sup>1</sup>genauer Kraftdichte als Kraft pro Flächeneinheit

<sup>2</sup>Genau diese Symmetrie erwirkt die Drehmomentfreiheit, welche ein Rotieren des Würfels verhindert. In Flächenkräften ausgedrückt bedeutet dies, dass auf die Flächen mit Normalen  $e_1$  bzw.  $-e_1$  die vertikalen Kräfte  $(0, s, 0)^T$  bzw.  $(0, -s, 0)^T$  wirken, welche das Drehmoment der äußeren Scherkräfte kompensieren (vgl. Abbildung 7.7).

<sup>3</sup>d.h.  $\det F = 1$  und  $e_3$  ist Eigenvektor von  $F$  zum Eigenwert 1

## 7. Shear

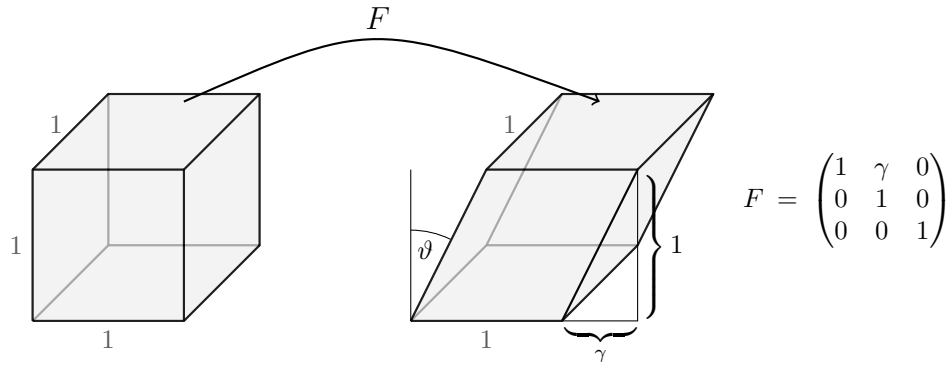


Abbildung 7.2.: Simple-Shear-Deformation mit Scherungswinkel  $\vartheta$  und Scherungsmaß  $\gamma = \tan(\vartheta)$

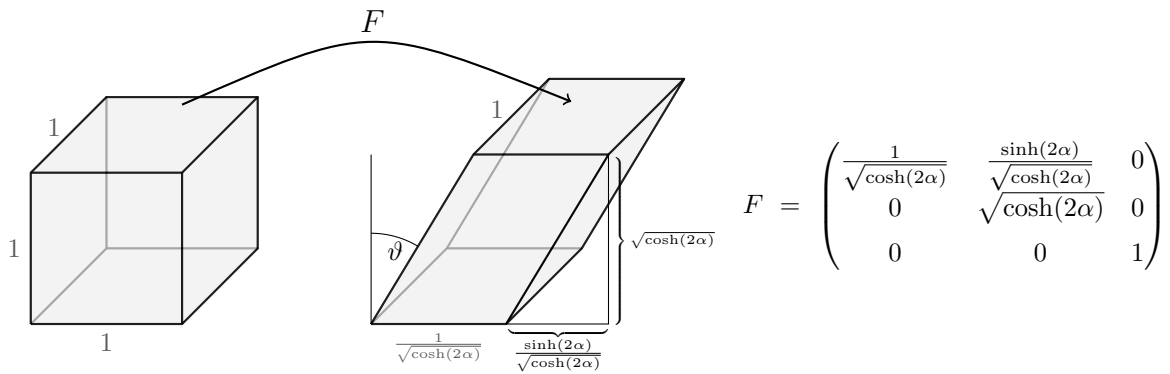


Abbildung 7.3.: linke Finite-Simple-Shear-Deformation mit  $\tan(\vartheta) = \tanh(2\alpha)$

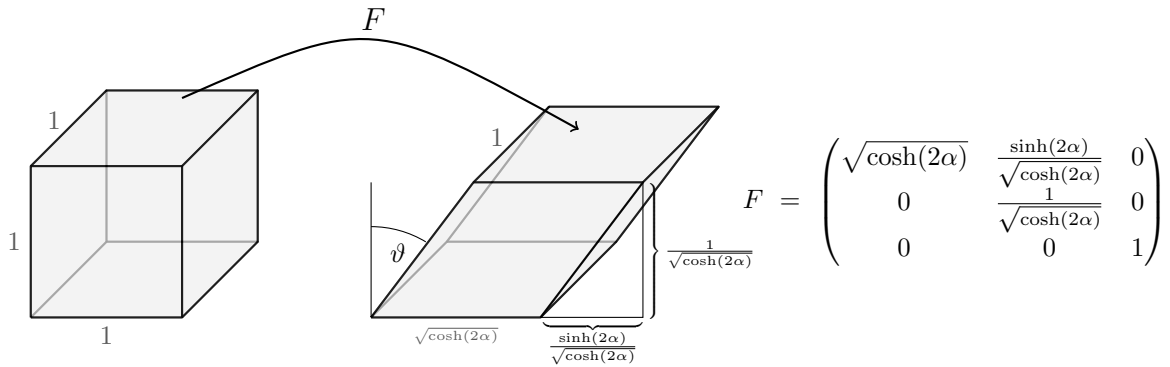


Abbildung 7.4.: rechte Finite-Simple-Biot-Shear-Deformation mit  $\tan(\vartheta) = \sinh(2\alpha)$

Als Ersatz werden wir dafür zwei dem Simple-Shear recht ähnliche Deformationen vorstellen: Wir werden zeigen, dass das auf eine Pure-Shear-Cauchy-Spannung führende Pendant zum Simple-Shear die *linke Finite-Simple-Shear-Deformation* (siehe Abbildung 7.3) und das auf eine Pure-Shear-Biot-Spannung führende Pendant die *rechte Finite-Simple-Shear-Deformation* (siehe Abbildung 7.4) ist.

Desweiteren werden wir neben der Simple-Shear-Deformation alle Deformationen, die im linearen Fall auf

## 7. Shear

Pure-Shear-Spannung führen, als

$$F = \mathbb{1} + \overbrace{\begin{pmatrix} 0 & \frac{\gamma}{2} & 0 \\ \frac{\gamma}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}^{\text{Simple-Shear}} + \underbrace{\begin{pmatrix} 0 & \frac{\gamma}{2} & 0 \\ -\frac{\gamma}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\in \mathfrak{so}(3)} + A \quad \text{mit } \gamma \in \mathbb{R} \text{ und } A \in \mathfrak{so}(3)$$

$\in \text{Sym}(3)$                        $\in \mathfrak{so}(3)$

identifizieren. Hierbei entsprechen die ersten beiden (symmetrischen) Summanden einer Dehnung (siehe Abbildung 7.6) und die letzten beiden (schiefsymmetrischen) Summanden einer linearisierten Drehung (vgl. Abschnitt 4.3.12).

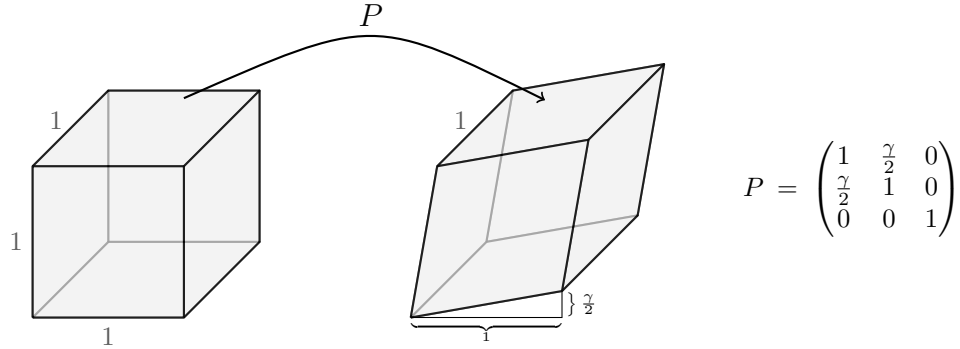


Abbildung 7.5.: lineare Pure-Shear-Dehnung

Auch für den finiten Fall werden wir neben der linken und rechten Finite-Simple-Shear-Deformation alle volumenerhaltenden und planaren Deformationen, die auf Pure-Shear-Spannungen führen, ausmachen. Wir werden beweisen, dass wenn die Pure-Shear-Spannung eine Cauchy-Spannung sein soll, sind dies alle  $F \in \text{GL}^+(3)$  der Form

$$F = \underbrace{\begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ \sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\in \text{Sym}(3)} \cdot \underbrace{\begin{pmatrix} \frac{\cosh(\alpha)}{\sqrt{\cosh(2\alpha)}} & \frac{\sinh(\alpha)}{\sqrt{\cosh(2\alpha)}} & 0 \\ -\frac{\sinh(\alpha)}{\sqrt{\cosh(2\alpha)}} & \frac{\cosh(\alpha)}{\sqrt{\cosh(2\alpha)}} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\in \text{SO}(3)} \cdot Q_1 \quad \text{mit } \alpha \in \mathbb{R},$$

$Q_1 \in \text{SO}(3),$

soll die Pure-Shear-Spannung eine Biot-Spannung sein, sind dies alle  $F \in \text{GL}^+(3)$  der Form

$$F = Q_2 \cdot \underbrace{\begin{pmatrix} \frac{\cosh(\alpha)}{\sqrt{\cosh(2\alpha)}} & \frac{\sinh(\alpha)}{\sqrt{\cosh(2\alpha)}} & 0 \\ -\frac{\sinh(\alpha)}{\sqrt{\cosh(2\alpha)}} & \frac{\cosh(\alpha)}{\sqrt{\cosh(2\alpha)}} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\in \text{SO}(3)} \cdot \underbrace{\begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ \sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\in \text{Sym}(3)} \quad \text{mit } \alpha \in \mathbb{R},$$

$Q_2 \in \text{SO}(3).$

Im ersten Fall folgt auf eine Rotation eine Dehnung, im zweiten Fall folgt auf eine Dehnung eine Rotation. Wir werden zeigen, dass die Dehnung in beiden Fällen identisch ist und sie als Pure-Shear-Dehnung (siehe Abbildung 7.6) bezeichnen.

Wir werden sogar sehen, dass die Linearisierungen der linken und der rechten Finite-Simple-Shear-Deformation beide genau der klassischen Simple-Shear-Deformation mit  $2\alpha = \gamma$  entsprechen.



## 7. Shear

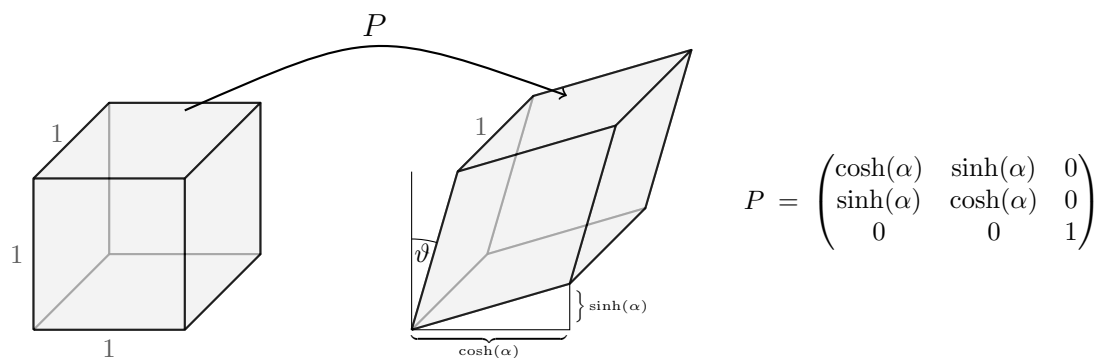


Abbildung 7.6.: Pure-Shear-Dehnung mit  $\tan(\vartheta) = \tanh(\alpha)$

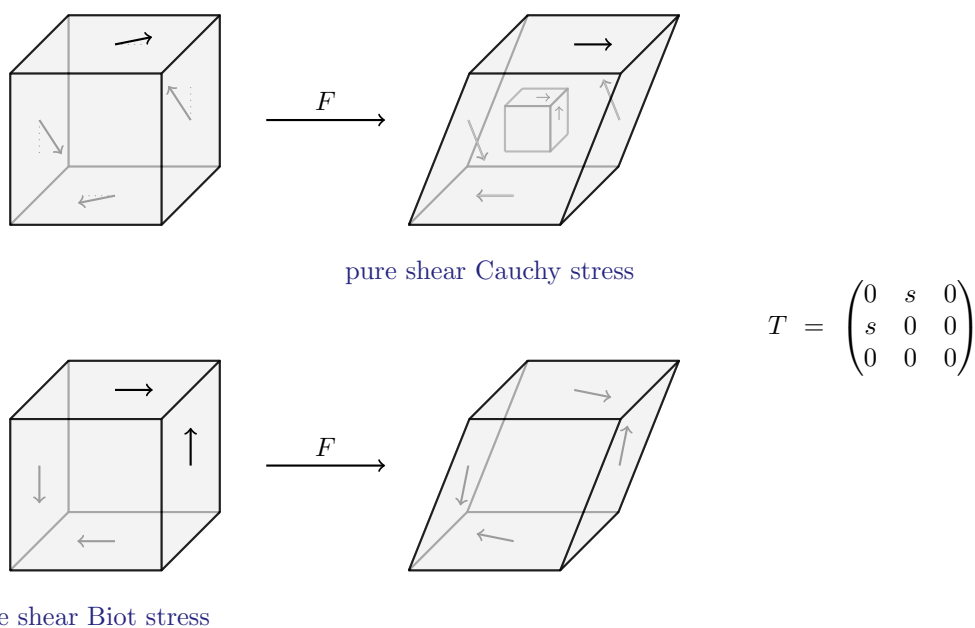


Abbildung 7.7.: Pure-Shear-Spannung dargestellt durch Kraftpfeile auf den Begrenzungsflächen eines undeformierten Referenzwürfels und des mittels einer Simple-Shear-Deformation verformten Würfels

Die Charakterisierungen der finiten Simple-Shear-Deformationen als Pure-Shear-Dehnung zuzüglich vorher bzw. nachher durchgeführter Rotation sind<sup>4</sup> lediglich notwendige Kriterien für Deformationen, um auf Pure-Shear-Cauchy-Spannungen bzw. auf Pure-Shear-Biot-Spannungen führen zu können. Wir werden aber in diesem Kapitel Energiefunktionen untersuchen, welche Elastizitätsgesetze induzieren, die eine dieser beiden Abbildungseigenschaften garantieren, also dass jede Pure-Shear-Dehnung auf eine Pure-Shear-Cauchy-Spannung bzw. jede Pure-Shear-Dehnung auf eine Pure-Shear-Biot-Spannung abgebildet wird.

Abschließend werden wir der Vollständigkeit halber die rechte und die linke Polarzerlegung des klassischen Simple-Shear-Deformationsgradienten untersuchen.

Wie bereits benutzt, verwenden wir in diesem Kapitel die folgenden Definitionen.

---

<sup>4</sup>modulo der Rotation

**Definitionen 7.0.38.** Wir nennen  $F \in \text{GL}^+(3)$  einen *Simple-Shear-Deformationsgradienten*, wenn  $F$  von der Form

$$F = \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{mit } \gamma \in \mathbb{R} \quad (7.0.2)$$

ist, einen *linken Simple-Finite-Shear-Deformationsgradienten*, wenn  $F$  von der Form

$$F = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} 1 & \sinh(2\alpha) & 0 \\ 0 & \cosh(2\alpha) & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix} \quad \text{mit } \alpha \in \mathbb{R} \quad (7.0.3)$$

ist und einen *rechten Simple-Finite-Shear-Deformationsgradienten*, wenn  $F$  von der Form

$$F = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} \cosh(2\alpha) & \sinh(2\alpha) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix} \quad \text{mit } \alpha \in \mathbb{R} \quad (7.0.4)$$

ist.

Wir nennen  $P \in \text{Sym}^+(3)$  eine *Pure-Shear-Dehnung*, wenn  $P$  von der Form

$$P = \begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ \sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{mit } \alpha \in \mathbb{R} \quad (7.0.5)$$

ist.

Wir nennen  $T \in \text{Sym}^+(3)$  eine *Pure-Shear-Spannung*, wenn  $T$  von der Form

$$T = \begin{pmatrix} 0 & s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{mit } s \in \mathbb{R} \quad (7.0.6)$$

ist.

## 7.1. Shear im linearen Modell – Simple-Shear-Deformationen und Pure-Shear-Spannungen

Der Zusammenhang zwischen einer Deformation und der aus ihr resultierenden Spannung wird im linearen Modell unter Annahme von Isotropie durch Lemma 4.3.35 auf Seite 123 beschrieben. Zu jedem isotropen linearen Elastizitätsgesetz gibt es Gewichtungskonstanten<sup>5</sup>  $\mu$  und  $\kappa$ , sodass sich die lineare Cauchy-Spannung zu gegebenem  $\varepsilon = \text{sym}(F - \mathbb{1})$  als gewichtete Summe des spurfreien Anteils  $\text{dev}_3 \varepsilon$  und des Spuranteils  $\text{tr}(\varepsilon) \cdot \mathbb{1}$  mittels

$$\sigma = 2\mu \text{dev}_3 \varepsilon + \kappa \text{tr}(\varepsilon) \cdot \mathbb{1} \quad (7.1.1)$$

ausdrücken lässt. Mit Hilfe dieser Darstellung können wir den folgenden Satz, der den Zusammenhang zwischen Deformationen und Spannung im isotropen linearen Modell beschreibt, leicht beweisen.

**Satz 7.1.1.** *Die linearisierte Cauchy-Spannung  $\sigma$  einer Deformation  $F \in \text{GL}^+(3)$  ist genau dann eine Pure-Shear-Spannung, wenn die Deformation von der Form*

$$F = \mathbb{1} + \begin{pmatrix} 0 & \frac{\gamma}{2} & 0 \\ \frac{\gamma}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + A \quad \text{mit } \gamma \in \mathbb{R} \text{ und } A \in \mathfrak{so}(3) \quad (7.1.2)$$

<sup>5</sup>Hierbei bezeichnet  $\mu$  den Lamé-Parameter und  $\kappa$  den Bulk-Modulus.

## 7. Shear

ist. Insbesondere ist jede Simple-Shear-Deformation  $F_\gamma$  von der Form (7.1.2).

*Beweis.* Sei  $F$  von der Form (7.1.2), dann ist

$$\varepsilon = \begin{pmatrix} 0 & \frac{\gamma}{2} & 0 \\ \frac{\gamma}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

und zusammen mit Lemma 4.3.35 erhalten wir

$$\sigma = 2\mu \operatorname{dev}_3 \varepsilon + \kappa \operatorname{tr}(\varepsilon) \cdot \mathbb{1} = \begin{pmatrix} 0 & \mu\gamma & 0 \\ \mu\gamma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

also eine Pure-Shear-Spannung mit  $s = \mu\gamma$ .

Sei andersherum  $\sigma$  eine Pure-Shear-Spannung. Da eine Pure-Shear-Spannung selbst spurfrei ist, muss aufgrund  $\sigma = 2\mu \operatorname{dev}_3 \varepsilon + \kappa \operatorname{tr}(\varepsilon) \cdot \mathbb{1}$  ebenfalls  $\operatorname{tr}(\varepsilon) = 0$ , also auch  $\operatorname{dev}_3 \varepsilon = \varepsilon$  gelten, was auf

$$\begin{pmatrix} 0 & s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 2\mu \varepsilon$$

führt. Damit können wir  $\varepsilon$  eindeutig auf die Form

$$\varepsilon = \begin{pmatrix} 0 & \frac{s}{2\mu} & 0 \\ \frac{s}{2\mu} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

festlegen. Diese Form führt schließlich auf (7.1.2) mit  $\gamma = \frac{s}{\mu}$  und  $A \in \mathfrak{so}(3)$ .

Siehe auch Vallée [80]. ■

Wir können die Form der Deformationen (7.1.2), die im isotropen linearen Modell auf eine Pure-Shear-Spannung führen noch weiter aufgesplitten und als

$$F = \mathbb{1} + \underbrace{\begin{pmatrix} 0 & \frac{\gamma}{2} & 0 \\ \frac{\gamma}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{=\varepsilon \in \operatorname{Sym}(3)} + \underbrace{\begin{pmatrix} 0 & \frac{\gamma}{2} & 0 \\ -\frac{\gamma}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{=B \in \mathfrak{so}(3)} + A \quad \text{mit } \gamma \in \mathbb{R} \text{ und } A \in \mathfrak{so}(3) \quad (7.1.3)$$

ausmachen. Für jedes solche  $F$  ist  $\mathbb{1} + \varepsilon + B$  genau eine Simple-Shear-Deformation mit Scherungsmaß  $\gamma$ , wie wir in Abbildung 7.8 sehen können und besteht aus der Dehnung  $\varepsilon \in \operatorname{Sym}(3)$  und der infinitesimalen Rotation  $B \in \mathfrak{so}(3)$ .

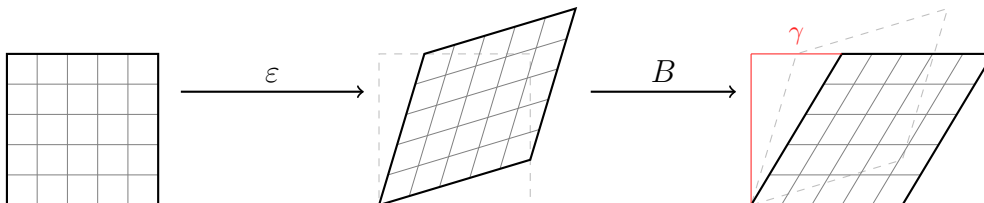


Abbildung 7.8.: Simple-Shear-Deformation

Wir wenden uns nun dem nichtlinearen Modell zu. Für den finiten Fall haben bereits Destrade et al. [23] gezeigt, dass sich der Zusammenhang zwischen Shear-Deformationen und Pure-Shear-Spannungen nicht so einfach darstellt.

## 7.2. Shear in nichtlinearen Modell – finite Simple-Shear-Deformationen und Pure-Shear-Spannungen

Für den finiten Fall haben Destrade et al. [23] gezeigt, dass keine Simple-Shear-Deformation, sondern nur eine Kombination aus einer Rotation, einer triaxialen Dehnung und einer Simple-Shear-Deformation auf eine Pure-Shear-Spannung führen kann. Wir werden in diesem Kapitel dieses Ergebnis bestätigen und zeigen, dass es zu jeder Simple-Shear-Deformation  $F_\gamma$  und Pure-Shear-Spannung  $T^s$  mit

$$F_\gamma = \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad T^s = \begin{pmatrix} 0 & s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (7.2.1)$$

eine triaxiale Dehnung  $\text{diag}(a, b, c)$  gibt, sodass  $F = F_\gamma \cdot \text{diag}(a, b, c) \cdot Q$  für alle  $Q \in \text{SO}(3)$  auf eine Pure-Shear-Cauchy-Spannung, also  $\widehat{\sigma}(FF^T) = T^s$ , führen kann, bzw. dass es eine triaxiale Dehnung  $\text{diag}(\tilde{a}, \tilde{b}, \tilde{c})$  gibt, sodass  $\tilde{F} = Q \cdot \text{diag}(\tilde{a}, \tilde{b}, \tilde{c}) \cdot F_\gamma$  für alle  $Q \in \text{SO}(3)$  auf eine Pure-Shear-Biot-Spannung, also  $\widehat{T}^{\text{Biot}}(F^T F) = T^s$  führen kann.<sup>6</sup>

Wir benötigen für unser Resultat, anders als Destrade et al., nicht die Semi-Invertierbarkeit des Elastizitätsgesetzes, also die Darstellbarkeit über die Repräsentationsformel (siehe. Satz 6.1.1)

$$B = \psi_0 \mathbb{1} + \psi_1 \sigma + \psi_2 \sigma^2 \quad \text{für alle } B \in \text{Sym}^+(3). \quad (7.2.2)$$

Destrade et al. verwenden sie in ihrem Beweis, fordern sie aber implizit; denn sie gehen einen Schritt weiter und fordern das Erfülltsein der schärferen empirischen Ungleichungen (E-TSS). Dies hat für sie den Vorteil, dass sie neben der Semi-Invertierbarkeit obendrein die Ordnung der Hauptdehnungen folgern können. Dennoch ist anzumerken, dass die deutlich schwächere Bedingung der strikten Baker-Ericksen-Ungleichungen (BE<sup>+</sup>) dafür auch gereicht hätten [25].<sup>7</sup>

Auf ähnliche Weise zeigt Batra [4, S. 110], dass für ein Elastizitätsgesetz, dass die empirischen Ungleichungen erfüllt, jede Dehnung, die auf eine uniaxiale Zugspannung führt, eine einfache Streckung sein muss. Und Marzano [48] zeigt für Elastizitätsgesetze, welche die strikte Baker-Ericksen Ungleichung (BE<sup>+</sup>) erfüllen: Für jede uniaxiale Spannung ( $T = \text{diag}(s, 0, 0)$ ) ist die auslösende Dehnung eine einfache Stauchung/Streckung ( $V - \mathbb{1} = \text{diag}(\alpha, 0, 0)$ ). Allerdings erweckt Marzano den Eindruck, dass die Umkehrung „aus der Implikation  $\sigma(V) = \text{diag}(s, 0, 0) \Rightarrow V - \mathbb{1} = \text{diag}(\alpha, 0, 0)$  folgt die Gültigkeit der Baker-Ericksen-Ungleichung (BE<sup>+</sup>)“ ebenso gilt. **Dies ist nicht korrekt!**<sup>8</sup>

<sup>6</sup>„führen kann“ bedeutet hierbei (wie am Ende dieses Abschnitts noch ausgeführt), dass diese Form von  $F$  bzw.  $\tilde{F}$  ein notwendiges Kriterium darstellt

<sup>7</sup>vgl. dazu die Implikationskette „(E-TSS)  $\Rightarrow$  (WE-TSS)  $\Rightarrow$  (BE<sup>+</sup>)  $\Rightarrow$  (semi)“ aus Satz 6.1.1 auf Seite 135.

<sup>8</sup>Wir sollten uns vor Augen führen: Offenkundig erfüllt jedes Elastizitätsgesetz, welches eine Dehnung  $V$  der Form  $V - \mathbb{1} = \text{diag}(\alpha, 0, 0)$  auf eine Spannung der Form  $T = \text{diag}(s, 0, 0)$  abbildet, für dieses  $V$  und das zugehörige  $\sigma(V)$  trivialerweise die Implikation  $\lambda_i > \lambda_j \Rightarrow \sigma_i > \sigma_j$  für alle  $i, j \in \{1, 2, 3\}$ . Die Baker-Ericksen-Ungleichung ist aber keine punktweise Eigenschaft und gilt für einzelne  $V$ ! Nur ein Elastizitätsgesetz kann die Baker-Ericksen-Ungleichung erfüllen, und dann gilt  $\lambda_i > \lambda_j \Rightarrow \sigma_i > \sigma_j$  für alle  $i, j \in \{1, 2, 3\}$  und **alle**  $V \in \text{Sym}^+(3)$ . Als Gegenbeispiel zu Marzanos Aussage betrachten wir das Elastizitätsgesetz

$$\sigma(V) = (1 - h(\lambda_1, \lambda_2, \lambda_3)) \cdot V - \mathbb{1}$$

mit der symmetrischen Funktion  $h: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  in den Eigenwerten von  $V$  mit

$$h(\lambda_1, \lambda_2, \lambda_3) = (\lambda_1 - \lambda_3)^2 (\lambda_2 - \lambda_3)^2 (\lambda_2 - \lambda_3)^2,$$

welche genau dann null ist, wenn zwei Eigenwerte gleich sind. Dann ist

$$\sigma(V) = \begin{pmatrix} s & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \Leftrightarrow \quad \begin{cases} (1 - h(\lambda_1, \lambda_2, \lambda_3)) \lambda_1 - 1 = s, \\ (1 - h(\lambda_1, \lambda_2, \lambda_3)) \lambda_2 - 1 = 0, \\ (1 - h(\lambda_1, \lambda_2, \lambda_3)) \lambda_3 - 1 = 0, \end{cases}$$

also  $\lambda_2 = \lambda_3$  und damit  $h(\lambda_1, \lambda_2, \lambda_3) = 0$ , was auf  $\lambda_1 = 1 + s$ ,  $\lambda_2 = \lambda_3 = 1$  und damit auf

$$V = \begin{pmatrix} 1+s & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Wir sollten uns aber bewusst machen, dass die Semi-Invertierbarkeit, ob wir sie nun direkt oder indirekt annehmen, immer eine Restriktion des Elastizitätsgesetzes mit sich bringt. Auf diese Einschränkung wollen wir verzichten. Wir fordern für den in diesem Kapitel von uns erbrachten Beweis nichts außer der Isotropie des Elastizitätsgesetzes. Unsere verwendeten Berechnungen sind denen von Destrade et al. zwar stellenweise ähnlich, allerdings liegt der Schlüssel unseres Beweises nicht in (7.2.2), sondern in Lemma 7.2.3: Aufgrund von Folgerung 4.3.22 auf Seite 112 wissen wir, dass für ein isotropes Elastizitätsgesetz der linke Dehnungstensor  $B = FF^T$  stets mit  $\hat{\sigma}(B)$  und der rechte Dehnungstensor  $C = F^T F$  stets mit  $\hat{T}^{\text{Biot}}(C)$  kommutiert. So können wir die allgemeine Form aller  $F \in \text{GL}^+(3)$  bestimmen, für die  $B = FF^T$  mit einer Pure-Shear-Cauchy-Spannung  $T^s$  kommutiert und die allgemeine Form aller  $F \in \text{GL}^+(3)$  bestimmen, für die  $C = F^T F$  mit einer Pure-Shear-Cauchy-Spannung  $T^s$  kommutiert.<sup>9</sup>

Allerdings sei an dieser Stelle explizit erwähnt, dass diese Kommutativität **nur ein notwendiges Kriterium** an die Deformation  $F \in \text{GL}^+(3)$  darstellt. Wir beweisen in diesem Abschnitt nicht, dass es ein Elastizitätsgesetz gibt, das solche  $F$  auf eine Cauchy oder Biot Pure-Shear-Spannung abbildet, sondern nur, dass **wenn** es eines gibt, dass dann die Deformationen zwingend die nachgewiesene allgemeine Form aufweisen.

Aber anders als Destrade et al. sind wir nicht direkt an Deformationen der Form

$$F = \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} Q \quad \text{bzw.} \quad F = Q \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (7.2.3)$$

also an Deformationen, die neben einer triaxialen Dehnung bloß eine Simple-Shear-Deformation enthalten, interessiert. Wir fordern, dass die Deformation als Ganzes die Eigenschaft hat, eine Scherung genannt werden zu können. Dafür muss die Deformation volumenerhaltend und planar sein. Wir werden zeigen, dass die Deformationen, die das erfüllen, die linke und die rechte Finite-Simple-Shear-Deformation sind: Jede Shear-Deformation, die auf eine Pure-Shear-Cauchy-Spannung führt, ist dabei eine linke Finite-Simple-Shear-Deformation und jede Shear-Deformation, die auf eine Pure-Shear-Biot-Spannung führt, ist dabei eine rechte Finite-Simple-Shear-Deformation.

### 7.2.1. Die allgemeine Form

Wie in (7.0.2) definiert, nennen wir  $T \in \text{Sym}^+(3)$  einen *Pure-Shear-Spannungstensor*, wenn  $T$  von der Form

$$T = \begin{pmatrix} 0 & s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{mit } s \in \mathbb{R} \quad (7.2.4)$$

ist.

**Proposition 7.2.1.** *Sei  $T \in \text{Sym}^+(3)$  ein Pure-Shear-Spannungstensor von der Form (7.2.4), dann hat  $T$  die Eigenwerte  $s, -s, 0$  zu den Eigenvektoren  $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0)^T, (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0)^T$  und  $(0, 0, 1)^T$ . Insbesondere lässt sich  $T$  mittels*

$$Q := \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \in \text{SO}(3), \quad (7.2.5)$$

zu  $Q \text{diag}(s, -s, 0)Q^T = T$  diagonalisieren.

**Bemerkung 7.2.2.** Wir sehen, entweder ist ein Pure-Shear-Spannungstensor entweder trivial (der spannungsfreie Fall des Nulltensors) oder er hat unterschiedliche Eigenwerte.

---

führt. Somit führt uniaxiale Spannung auf eine einfache Streckung. Aber  $\sigma$  erfüllt nicht die Baker-Ericksen-Ungleichungen, wie das Beispiel  $V = \text{diag}(3, 2, 1)$  zeigt. Denn es ist  $h(3, 2, 1) = 2$  und damit  $\sigma(\text{diag}(3, 2, 1)) = -\text{diag}(3, 2, 1) - \mathbb{1} = \text{diag}(-4, -3, -1)$ . Wir haben somit  $\lambda_1 = 3 > \lambda_2 = 2$  und  $\sigma_1 = -4 < \sigma_2 = -3$ , ein Widerspruch zu  $(\text{BE}^+)$ .

<sup>9</sup>Wir müssen allerdings anmerken, dass unsere eingesetzte Methode beim Problem der von Batra betrachteten uniaxialen Belastungen nicht funktioniert, da in diesem Fall der Spannungstensor nicht drei unterschiedliche Eigenwerte besitzt (vgl. Bemerkung 7.2.4).

## 7. Shear

*Beweis.* Einfaches Nachrechnen ergibt

$$\begin{aligned}
 Q \begin{pmatrix} s & 0 & 0 \\ 0 & -s & 0 \\ 0 & 0 & 0 \end{pmatrix} Q^T &= \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} s & 0 & 0 \\ 0 & -s & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \\
 &= \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} s & s & 0 \\ s & -s & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
 &= \frac{1}{2} \begin{pmatrix} 0 & 2s & 0 \\ 2s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = T. \quad \blacksquare
 \end{aligned}$$

Es soll nun  $\hat{\sigma}(B)$  für jedes  $B = FF^T$  mit  $F \in \text{GL}^+(3)$  ein Pure-Shear-Spannungstensor sein. Wir untersuchen, welche Einschränkungen wir unter Voraussetzung der Isotropie von  $\hat{\sigma}$  dafür am Deformationsgradienten  $F$  hinnehmen müssen. Die Isotropie stellt nach Folgerung 2.5.13 sicher, dass  $B$  mit  $\hat{\sigma}(B)$  kommutiert, daher untersuchen wir zunächst folgendes Lemma.

**Lemma 7.2.3.** *Seien  $T$  ein Pure-Shear-Spannungstensor mit  $s \neq 0$  und  $P \in \text{Sym}^+(3)$ . Dann kommutieren  $P$  und  $T$  genau dann, wenn  $P$  der Form*

$$P = \begin{pmatrix} p & q & 0 \\ q & p & 0 \\ 0 & 0 & r \end{pmatrix} \quad (7.2.6)$$

*genügt. Sind  $\mu_1, \mu_2, \mu_3 \in \mathbb{R}^+$  die Eigenwerte von  $P$ , dann gilt  $p = \frac{1}{2}(\mu_1 + \mu_2)$ ,  $q = \frac{1}{2}(\mu_1 - \mu_2)$  und  $r = \mu_3$ .*

*Beweis.* Angenommen,  $P$  und  $T$  lassen sich gemeinsam diagonalisieren. Lemma 2.5.20 besagt nun: Hat  $T$  zudem nur unterschiedliche Eigenwerte, so ist  $T$  koaxial zu  $P$ , d.h. jeder Eigenvektor von  $T$  ist auch Eigenvektor von  $P$ . In Proposition 7.2.1 haben wir gezeigt, dass wir  $T$  mittels  $Q$  wie in (7.2.5) diagonalisieren können und dass alle Eigenwerte von  $T$  genau dann unterschiedlich sind, wenn  $s \neq 0$  gilt, wobei wir letzteres in diesem Lemma voraussetzen. Somit ist  $T$  koaxial zu  $P$  und wir können  $P$  mittels  $Q$  in Diagonalgestalt bringen. Seien  $\mu_1, \mu_2, \mu_3 \in \mathbb{R}$  die Eigenwerte von  $P$ , dann erhalten wir

$$\begin{aligned}
 P &= Q \begin{pmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_2 & 0 \\ 0 & 0 & \mu_3 \end{pmatrix} Q^T = \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_2 & 0 \\ 0 & 0 & \mu_3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \\
 &= \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} \mu_1 & \mu_1 & 0 \\ -\mu_2 & \mu_2 & 0 \\ 0 & 0 & \sqrt{2}\mu_3 \end{pmatrix} \\
 &= \frac{1}{2} \begin{pmatrix} \mu_1 + \mu_2 & \mu_1 - \mu_2 & 0 \\ \mu_1 - \mu_2 & \mu_1 + \mu_2 & 0 \\ 0 & 0 & 2\mu_3 \end{pmatrix}.
 \end{aligned}$$

Also ist auch  $P$  von der Form (7.2.6).

Nehmen wir andersherum an, dass  $P$  von der Form (7.2.6) ist. Dann gilt

$$PT = \begin{pmatrix} p & q & 0 \\ q & p & 0 \\ 0 & 0 & r \end{pmatrix} \begin{pmatrix} 0 & s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} qs & ps & 0 \\ ps & qs & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p & q & 0 \\ q & p & 0 \\ 0 & 0 & r \end{pmatrix} = TP. \quad \blacksquare$$

**Bemerkung 7.2.4.** Es ist zu beachten, dass der Beweis von Lemma 7.2.3 den Umstand nutzt, dass eine (nichttriviale) Pure-Shear-Spannung ausschließlich einfache Eigenwerte besitzt; ansonsten wäre es ohne weitere Annahmen nicht möglich zu schließen, dass jeder Eigenvektor von  $T$  auch ein Eigenvektor von  $P$

## 7. Shear

ist. Im Hinblick auf Elastizitätstensoren muss dies nicht allgemein erfüllt sein, es sei denn die Hauptspannungen sind paarweise unterschiedlich oder wir nehmen zusätzliche Beschränkungen an das konstitutive Gesetz an. Unter den konstitutiven Voraussetzungen, welche diese „Bi-Koaxialität“ von Spannung und Dehnung garantieren, sind die empirischen Ungleichungen [78], obgleich die (schwächeren) Baker-Ericksen-Ungleichungen ebenso hinreichend sind.

Damit  $B = FF^T$  bzw.  $C = F^TF$  mit einem Pure-Shear-Spannungstensor  $T$  kommutiert, muss  $B$  bzw.  $C$  von der Form

$$\begin{pmatrix} p & q & 0 \\ q & p & 0 \\ 0 & 0 & r \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \lambda_1^2 + \lambda_2^2 & \lambda_1^2 - \lambda_2^2 & 0 \\ \lambda_1^2 - \lambda_2^2 & \lambda_1^2 + \lambda_2^2 & 0 \\ 0 & 0 & 2\lambda_3^2 \end{pmatrix} \quad (7.2.7)$$

sein. Hierbei sind  $\lambda_1, \lambda_2, \lambda_3$  die Singulärwerte von  $F$  und  $\lambda_1^2, \lambda_2^2, \lambda_3^2$  die Eigenwerte von  $B$  und von  $C$ . Zudem ist

$$\begin{pmatrix} p & q & 0 \\ q & p & 0 \\ 0 & 0 & r \end{pmatrix}^{1/2} = \frac{1}{2} \begin{pmatrix} \lambda_1 + \lambda_2 & \lambda_1 - \lambda_2 & 0 \\ \lambda_1 - \lambda_2 & \lambda_1 + \lambda_2 & 0 \\ 0 & 0 & 2\lambda_3 \end{pmatrix}. \quad (7.2.8)$$

Machen wir uns aber noch einmal explizit klar: Auch wenn über  $Q$  wie in (7.2.5) die Form von  $B$  bzw.  $C$  festgelegt sind, gilt das nicht für die Eigenwerte; für alle  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_+$  kommutieren  $B$  bzw.  $C$  mit  $T$ .

Lemma 7.2.3 ist von zentraler Bedeutung für die Charakterisierung der Deformationen, die einer Pure-Shear-Spannung entsprechen. Denn wir wissen durch Folgerung 4.3.22 auf Seite 112, dass  $B$  mit  $\widehat{\sigma}(B)$  und  $C$  mit  $\widehat{T}^{\text{Biot}}$  und mit  $\widehat{S}_2$  kommutiert. So können wir ohne weitere Vorarbeit direkt zeigen, dass eine (nichttriviale) Simple-Shear-Deformation niemals auf eine Pure-Shear-(Cauchy/Biot/Pioala-Kirchhoff-)Spannung, führen kann.

**Bemerkung 7.2.5.** Seien  $F$  ein Simple-Shear-Deformationsgradient mit  $\gamma \neq 0$  und  $T$  ein Pure-Shear-Spannungstensor mit  $s \neq 0$ . Da

$$FF^T = \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 + \gamma^2 & \gamma & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7.2.9)$$

und

$$F^TF = \begin{pmatrix} 1 & 0 & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \gamma & 0 \\ \gamma & 1 + \gamma^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7.2.10)$$

beide nicht von der Form (7.2.6) sind, kommutieren weder  $FF^T$  noch  $F^TF$  mit  $T$ .

Aber welche  $F \in \text{GL}^+(3)$  führen nun als der quadrierte positiv definite Faktor  $B = FF^T$  der linken Polarzerlegung von  $F$  oder als der quadrierte positiv definite Faktor  $C = F^TF$  der rechten Polarzerlegung von  $F$  auf die Form (7.2.6)? Wir wollen mit den nächsten zwei Lemmata dieser Frage nachgehen und die entsprechenden Deformationen charakterisieren.

**Lemma 7.2.6.** Sei  $P \in \text{Sym}^+(3)$  mit

$$P = \begin{pmatrix} p & q & 0 \\ q & p & 0 \\ 0 & 0 & r \end{pmatrix} \quad \text{und} \quad p > |q|, \quad r > 0 \quad (7.2.11)$$

gegeben. Dann ist  $F \in \text{GL}^+(3)$  mit  $FF^T = P$  bis auf beliebiges  $Q \in \text{SO}(3)$  eindeutig bestimmt durch

$$F = \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} Q \quad (7.2.12)$$

## 7. Shear

wobei

$$a = \sqrt{\frac{p^2 - q^2}{p}}, \quad b = \sqrt{p}, \quad c = \sqrt{r} \quad \text{und} \quad \gamma = \frac{q}{p}.$$

bzw. in den Singulärwerten  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_+$  von  $F$  ausgedrückt

$$a = \lambda_1 \lambda_2 \sqrt{\frac{2}{\lambda_1^2 + \lambda_2^2}}, \quad b = \sqrt{\frac{\lambda_1^2 + \lambda_2^2}{2}}, \quad c = \lambda_3, \quad \gamma = \frac{\lambda_1^2 - \lambda_2^2}{\lambda_1^2 + \lambda_2^2}.$$

**Lemma 7.2.7.** Sei  $P \in \text{Sym}^+(3)$  mit

$$P = \begin{pmatrix} p & q & 0 \\ q & p & 0 \\ 0 & 0 & r \end{pmatrix} \quad \text{und} \quad p > |q|, \quad r > 0 \quad (7.2.13)$$

gegeben. Dann ist  $F \in \text{GL}^+(3)$  mit  $F^T F = P$  bis auf beliebiges  $Q \in \text{SO}(3)$  eindeutig bestimmt durch

$$F = Q \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7.2.14)$$

wobei

$$a = \sqrt{p}, \quad b = \sqrt{\frac{p^2 - q^2}{p}}, \quad c = \sqrt{r} \quad \text{und} \quad \gamma = \frac{q}{p}. \quad (7.2.15)$$

bzw. in den Singulärwerten  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_+$  von  $F$  ausgedrückt

$$a = \sqrt{\frac{\lambda_1^2 + \lambda_2^2}{2}}, \quad b = \lambda_1 \lambda_2 \sqrt{\frac{2}{\lambda_1^2 + \lambda_2^2}}, \quad c = \lambda_3, \quad \gamma = \frac{\lambda_1^2 - \lambda_2^2}{\lambda_1^2 + \lambda_2^2}. \quad (7.2.16)$$

**Bemerkung 7.2.8.** Für die Wohldefiniertheit des Ausdrucks  $\sqrt{\frac{p^2 - q^2}{p}}$  für  $a$  bzw.  $b$  muss  $p > |q|$  gelten. Dieser Umstand führt auf die Schranke  $|\gamma| = \frac{|q|}{p} < 1$ .

**Bemerkung 7.2.9.** Mit den Lemmata 7.2.6 und 7.2.7 können wir auf das Hauptresultat von Destrade et al. schließen: Eine Pure-Shear-Spannung induziert keine Simple-Shear-Deformation, aber eine Kombination einer triaxialen Dehnung mit einer Simple-Shear-Deformation. Hierbei ist das Scherungsmaß  $\gamma$  der Simple-Shear-Deformation betraglich durch eins damit der Scherungswinkel durch  $45^\circ$  beschränkt. Neben der Isotropie ist unsere einzige Forderung an das konstitutive Gesetz die Eindeutigkeit bis auf Rotation der spannungsfreien Referenzkonfiguration, das heißt  $\hat{\sigma}(B) = 0$  genau dann, wenn  $B = \mathbb{1}$ . Wir benötigen nicht die Semi-Invertierbarkeit von  $\hat{\sigma}$ ! Zudem machen wir uns klar: Für ein vorgegebenes  $s \in \mathbb{R}$  fordern wir weder die Existenz, noch die Eindeutigkeit eines  $F \in \text{GL}^+(3)$  mit

$$\hat{\sigma}(FF^T) = \begin{pmatrix} 0 & s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{bzw.} \quad \hat{T}^{\text{Biot}}(F^T F) = \begin{pmatrix} 0 & s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (7.2.17)$$

Ob solch ein  $F$  existiert, hängt einzig und allein von der konstitutiven Beziehung ab. Aber was wir sagen können, ist: Falls solch ein  $F$  existiert, so ist es unweigerlich von der Form (7.2.12) bzw. (7.2.14).

*Beweis von Lemma 7.2.6.* Jedes  $F \in \text{GL}^+(3)$ , welches  $FF^T = P$  erfüllt, lässt sich eindeutig als  $F = VR$  mit  $V \in \text{Sym}^+(3)$  und  $R \in \text{SO}(3)$  und  $V^2 = P$  darstellen. Auf  $\text{Sym}^+(3)$  ist die Abbildung  $V^2 \mapsto V$  injektiv, sodass  $V$  eindeutig durch  $P$  bestimmt ist. Andererseits gilt dann

$$FQ(FQ)^T = FQQ^T F^T = P \quad \text{für jedes } Q \in \text{SO}(3).$$



## 7. Shear

Ohne Einschränkung setzen wir also

$$F := \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$$

und rechnen aus

$$\begin{aligned} FF^T &= \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a^2 & 0 & 0 \\ 0 & b^2 & 0 \\ 0 & 0 & c^2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a^2 & 0 & 0 \\ b^2\gamma & b^2 & 0 \\ 0 & 0 & c^2 \end{pmatrix} \\ &= \begin{pmatrix} a^2 + b^2\gamma^2 & b^2\gamma & 0 \\ b^2\gamma & b^2 & 0 \\ 0 & 0 & c^2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} p & q & 0 \\ q & p & 0 \\ 0 & 0 & r \end{pmatrix}. \end{aligned}$$

Damit nun  $a^2 + b^2\gamma^2 = b^2 = p$ ,  $b^2\gamma = q$  und  $c^2 = r$  erfüllt ist, muss  $b = \sqrt{p}$ ,  $c = \sqrt{r}$ ,  $\gamma = \frac{q}{b^2} = \frac{q}{p}$  und  $a = \sqrt{p - b^2\gamma^2} = \sqrt{p - \frac{q^2}{p}} = \sqrt{\frac{p^2 - q^2}{p}}$  gelten.

Nach Lemma 7.2.3 sind  $p = \frac{1}{2}(\lambda_1^2 + \lambda_2^2)$ ,  $q = \frac{1}{2}(\lambda_1^2 - \lambda_2^2)$  und  $r = \lambda_3^2$ , woraus sich sofort  $b = \sqrt{\frac{\lambda_1^2 + \lambda_2^2}{2}}$ ,  $c = \lambda_3$  und  $\gamma = \frac{\lambda_1^2 - \lambda_2^2}{\lambda_1^2 + \lambda_2^2}$  ergeben. Zudem ist  $p + q = \lambda_1^2$  und  $p - q = \lambda_2^2$ , woraus  $p^2 - q^2 = \lambda_1^2 \lambda_2^2$  und daraus dann  $a = \lambda_1 \lambda_2 \sqrt{\frac{2}{\lambda_1^2 + \lambda_2^2}}$  folgt. ■

*Beweis von Lemma 7.2.7.* Der Beweis verläuft analog zum Beweis von Lemma 7.2.6:

Jedes  $F \in \text{GL}^+(3)$ , welches  $F^T F = P$  erfüllt, lässt sich eindeutig als  $F = RU$  mit  $R \in \text{SO}(3)$  und  $U \in \text{Sym}^+(3)$  und  $U^2 = P$  darstellen. Auf  $\text{Sym}^+(3)$  ist die Abbildung  $U^2 \mapsto U$  injektiv, sodass  $U$  eindeutig durch  $P$  bestimmt ist. Andererseits gilt dann

$$(QF)^T QF = F^T Q^T QF = P \quad \text{für jedes } Q \in \text{SO}(3).$$

Ohne Einschränkung setzen wir also

$$F := \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und rechnen aus

$$\begin{aligned} F^T F &= \begin{pmatrix} 1 & 0 & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a^2 & 0 & 0 \\ 0 & b^2 & 0 \\ 0 & 0 & c^2 \end{pmatrix} \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a^2 & a^a\gamma & 0 \\ 0 & b^2 & 0 \\ 0 & 0 & c^2 \end{pmatrix} \\ &= \begin{pmatrix} a^2 & a^2\gamma & 0 \\ a^2\gamma & a^2\gamma^2 + b^2 & 0 \\ 0 & 0 & c^2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} p & q & 0 \\ q & p & 0 \\ 0 & 0 & r \end{pmatrix}. \end{aligned}$$

## 7. Shear

Damit nun  $a^2 = a^2\gamma^2 + b^2 = p$ ,  $a^2\gamma = q$  und  $c^2 = r$  erfüllt ist, muss  $a = \sqrt{p}$ ,  $c = \sqrt{r}$ ,  $\gamma = \frac{q}{a^2} = \frac{q}{p}$  und  $b = \sqrt{p - a^2\gamma^2} = \sqrt{p - \frac{q^2}{p}} = \sqrt{\frac{p^2 - q^2}{p}}$  gelten.

Nach Lemma 7.2.3 sind  $p = \frac{1}{2}(\lambda_1^2 + \lambda_2^2)$ ,  $q = \frac{1}{2}(\lambda_1^2 - \lambda_2^2)$  und  $r = \lambda_3^2$ , woraus sich sofort  $a = \sqrt{\frac{\lambda_1^2 + \lambda_2^2}{2}}$ ,  $c = \lambda_3$  und  $\gamma = \frac{\lambda_1^2 - \lambda_2^2}{\lambda_1^2 + \lambda_2^2}$  ergeben. Zudem ist  $p + q = \lambda_1^2$  und  $p - q = \lambda_2^2$ , woraus  $p^2 - q^2 = \lambda_1^2 \lambda_2^2$  und daraus dann  $b = \lambda_1 \lambda_2 \sqrt{\frac{2}{\lambda_1^2 + \lambda_2^2}}$  folgt. ■

### 7.2.2. Finite-Simple-Shear-Deformationen

Wir kennen nun die allgemeine Form der  $F \in GL^+(3)$ , für die  $B = FF^T$  bzw.  $C = F^T F$  mit einem Pure-Shear-Spannungstensor

$$T = \begin{pmatrix} 0 & s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (7.2.18)$$

kommutieren. Aber nicht jede Deformation dieser allgemeinen Form

$$F = \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} Q \quad \text{bzw.} \quad F = Q \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7.2.19)$$

ist geeignet, „shear“ genannt zu werden. Als charakteristische Eigenschaften einer Shear-Deformation  $F$  werden die drei folgenden Merkmale angenommen:

- i) Die Deformation  $F$  ist *volumenerhaltend*, d.h.  $\det F = 1$ .
- ii) Die Deformation  $F$  ist *planar*, d.h.  $F$  hat den Eigenwert 1 zum Eigenvektor  $e_3$ .
- iii) Die Deformation  $F$  ist *bodenparallel*, d.h.  $e_1$  ist ein Eigenvektor von  $F$ .

Hierbei sind  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_+$  die Singulärwerte von  $F$ .

Machen wir uns klar, dass die ersten beiden Eigenschaften zur Folge haben, dass zu den Singulärwerten  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_+$  von  $F$  ein  $\lambda \in \mathbb{R}_+$  gibt, sodass  $\lambda_1 = \lambda$ ,  $\lambda_2 = \frac{1}{\lambda}$  und  $\lambda_3 = 1$  gilt.

**Lemma 7.2.10.** *Seien  $T$  ein Pure-Shear-Spannungstensor mit  $s \neq 0$  und  $P \in \text{Sym}^+(3)$  mit  $\lambda_1 = \lambda$ ,  $\lambda_2 = \frac{1}{\lambda}$  und  $\lambda_3 = 1$  für ein  $\lambda \in \mathbb{R}_+$ , wobei  $\lambda_1^2, \lambda_2^2, \lambda_3^2$  die Eigenwerte von  $P$  sind. Dann kommutieren  $P$  und  $T$  genau dann, wenn  $P$  der Form*

$$\sqrt{P} = \begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ \sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{mit } \alpha \in \mathbb{R} \quad (7.2.20)$$

genügt, d.h. wenn  $\sqrt{P}$  eine finite Pure-Shear-Dehnung ist.

*Beweis.* Aufgrund von Lemma 7.2.3 kommutiert  $P$  genau dann mit  $T$ , wenn

$$P = \frac{1}{2} \begin{pmatrix} \lambda_1^2 + \lambda_2^2 & \lambda_1^2 - \lambda_2^2 & 0 \\ \lambda_1^2 - \lambda_2^2 & \lambda_1^2 + \lambda_2^2 & 0 \\ 0 & 0 & 2\lambda_3^2 \end{pmatrix}. \quad (7.2.21)$$

Verwenden wir nun die Vorgabe  $\lambda_1 = \lambda$ ,  $\lambda_2 = \frac{1}{\lambda}$ ,  $\lambda_3 = 1$  und setzen dabei  $\alpha := \log(\lambda) \in \mathbb{R}$ , dann ist

$$\begin{aligned} \sqrt{P} &= \frac{1}{2} \begin{pmatrix} \lambda_1 + \lambda_2 & \lambda_1 - \lambda_2 & 0 \\ \lambda_1 - \lambda_2 & \lambda_1 + \lambda_2 & 0 \\ 0 & 0 & 2\lambda_3 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} e^\alpha + e^{-\alpha} & e^\alpha - e^{-\alpha} & 0 \\ e^\alpha - e^{-\alpha} & e^\alpha + e^{-\alpha} & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ \sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad \blacksquare \end{aligned}$$

## 7. Shear

**Bemerkung 7.2.11.** Im Gegensatz zur linearen Pure-Shear-Dehnung aus Abschnitt 7.1 ist die finite Pure-Shear-Dehnung volumenerhaltend. Diesen Umstand kann man für ausgewählte  $\gamma$  bzw.  $\alpha$  in Abbildung 7.9 erkennen: Die Determinante der linearen Pure-Shear-Dehnung ist  $1 - \frac{\gamma^2}{4}$ , die der finiten Pure-Shear-Dehnung ist 1.

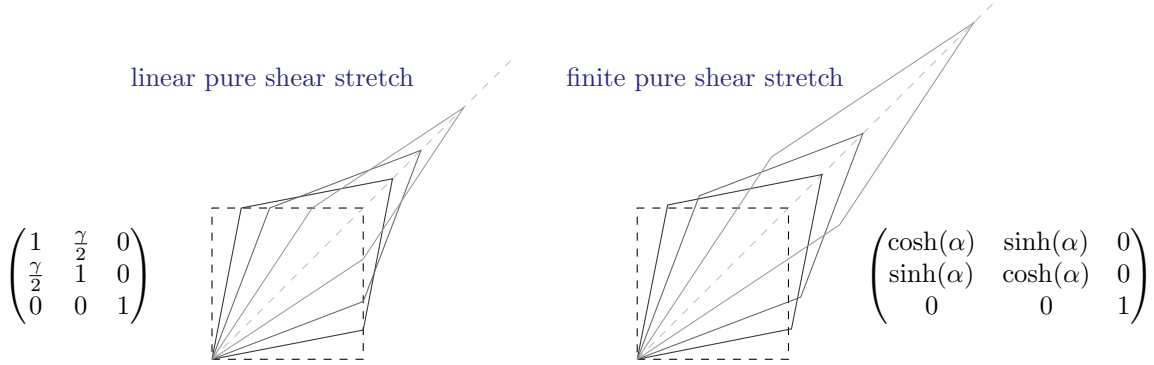


Abbildung 7.9.: lineare und finite Pure-Shear-Dehnung

**Lemma 7.2.12.** Sei  $P$  von der Form (7.2.20) gegeben. Dann ist  $F \in \text{GL}^+(3)$  mit  $FF^T = P$  bis auf beliebiges  $Q \in \text{SO}(3)$  eindeutig bestimmt durch

$$F = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} 1 & \sinh(2\alpha) & 0 \\ 0 & \cosh(2\alpha) & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix} \cdot Q. \quad (7.2.22)$$

**Lemma 7.2.13.** Sei  $P$  von der Form (7.2.20) gegeben. Dann ist  $F \in \text{GL}^+(3)$  mit  $F^T F = P$  bis auf beliebiges  $Q \in \text{SO}(3)$  eindeutig bestimmt durch

$$F = Q \cdot \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} \cosh(2\alpha) & \sinh(2\alpha) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix}. \quad (7.2.23)$$

*Beweis von Lemma 7.2.12.* Wir verwenden Lemma 7.2.6 und setzen dort direkt  $\lambda_1 = e^\alpha$ ,  $\lambda_2 = e^{-\alpha}$ ,  $\lambda_3 = 1$  ein. Dann erhalten wir

$$F = \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} Q,$$

mit

$$\begin{aligned} a &= \lambda_1 \lambda_2 \sqrt{\frac{2}{\lambda_1^2 + \lambda_2^2}} = e^\alpha e^{-\alpha} \cdot \sqrt{\frac{2}{e^{2\alpha} + e^{-2\alpha}}} = \frac{1}{\sqrt{\cosh(2\alpha)}} \\ b &= \sqrt{\frac{\lambda_1^2 + \lambda_2^2}{2}} = \sqrt{\frac{e^{2\alpha} + e^{-2\alpha}}{2}} = \sqrt{\cosh(2\alpha)} \\ c &= \lambda_3 = 1 \\ \gamma &= \frac{\lambda_1^2 - \lambda_2^2}{\lambda_1^2 + \lambda_2^2} = \frac{e^{2\alpha} - e^{-2\alpha}}{e^{2\alpha} + e^{-2\alpha}} = \tanh(2\alpha) \end{aligned}$$

## 7. Shear

und damit

$$\begin{aligned}
 F &= \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} Q \\
 &= \begin{pmatrix} 1 & \tanh(2\alpha) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{\cosh(2\alpha)}} & 0 & 0 \\ 0 & \sqrt{\cosh(2\alpha)} & 0 \\ 0 & 0 & 1 \end{pmatrix} Q \\
 &= \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} 1 & \sinh(2\alpha) & 0 \\ 0 & \cosh(2\alpha) & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix} Q. \quad \blacksquare
 \end{aligned}$$

*Beweis von Lemma 7.2.13.* Der zweite Beweis erfolgt analog zum ersten: Wir verwenden Lemma 7.2.7 und setzen dort direkt  $\lambda_1 = e^\alpha$ ,  $\lambda_2 = e^{-\alpha}$ ,  $\lambda_3 = 1$  ein. Dann erhalten wir

$$F = Q \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

mit

$$\begin{aligned}
 a &= \sqrt{\frac{\lambda_1^2 + \lambda_2^2}{2}} = \sqrt{\frac{e^{2\alpha} + e^{-2\alpha}}{2}} = \sqrt{\cosh(2\alpha)} \\
 b &= \lambda_1 \lambda_2 \sqrt{\frac{2}{\lambda_1^2 + \lambda_2^2}} = e^\alpha e^{-\alpha} \cdot \sqrt{\frac{2}{e^{2\alpha} + e^{-2\alpha}}} = \frac{1}{\sqrt{\cosh(2\alpha)}} \\
 c &= \lambda_3 = 1 \\
 \gamma &= \frac{\lambda_1^2 - \lambda_2^2}{\lambda_1^2 + \lambda_2^2} = \frac{e^{2\alpha} - e^{-2\alpha}}{e^{2\alpha} + e^{-2\alpha}} = \tanh(2\alpha)
 \end{aligned}$$

und damit

$$\begin{aligned}
 F &= Q \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= Q \begin{pmatrix} \sqrt{\cosh(2\alpha)} & 0 & 0 \\ 0 & \frac{1}{\sqrt{\cosh(2\alpha)}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \tanh(2\alpha) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= Q \cdot \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} \cosh(2\alpha) & \sinh(2\alpha) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix}. \quad \blacksquare
 \end{aligned}$$

Nun sind wir in der Lage, unser angekündigtes Hauptresultat schlusszufolgern.

**Satz 7.2.14.** *Jede Shear-Deformation<sup>10</sup>  $F \in \text{GL}^+(3)$ , die auf eine (nichttriviale) Pure-Shear-Cauchy-Spannung führt, ist eine linke Finite-Simple-Shear-Deformation*

$$F = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} 1 & \sinh(2\alpha) & 0 \\ 0 & \cosh(2\alpha) & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix} \quad \text{mit } \alpha \in \mathbb{R} \quad (7.2.24)$$

<sup>10</sup>d.h. es gilt  $\det F = 1$ ,  $e_3$  ist ein Eigenvektor von  $F$  zum Eigenwert 1 und  $e_1$  ist ein Eigenvektor von  $F$

## 7. Shear

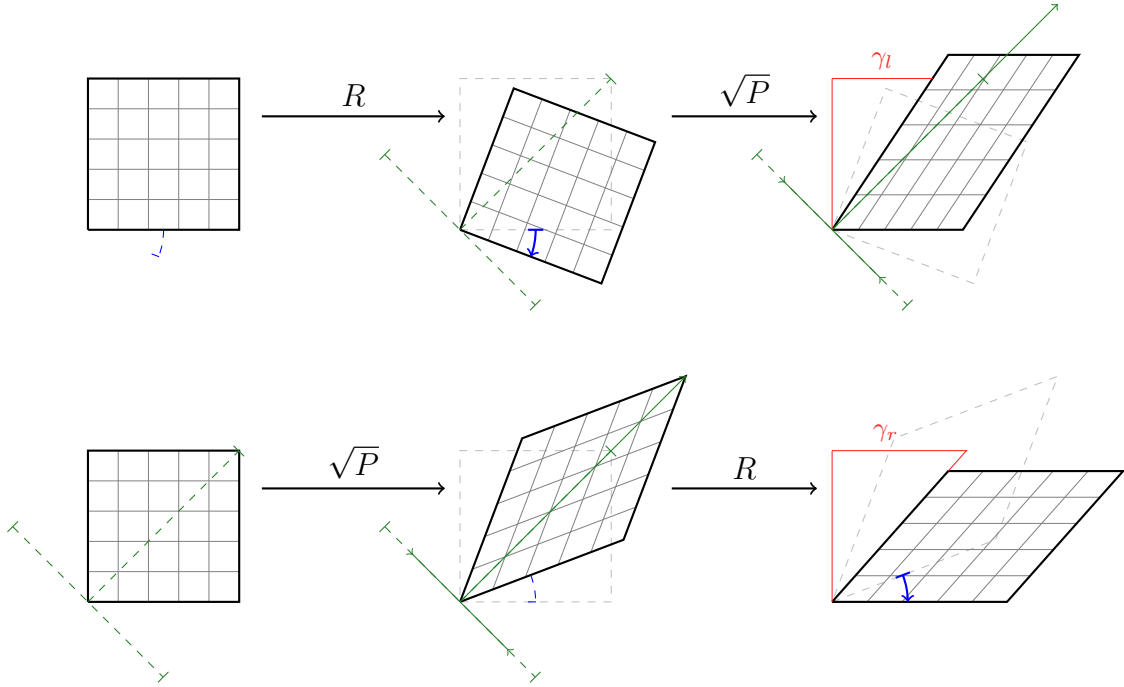


Abbildung 7.10.: linke Finite-Simple-Shear-Deformation mit Scherungsmaß  $\gamma_l = \tanh(2\alpha)$  und rechte Finite-Simple-Shear-Deformation mit Scherungsmaß  $\gamma_r = \sinh(2\alpha)$

und jede Shear-Deformation, die auf eine (nichttriviale) Pure-Shear-Biot-Spannung führt, ist dabei eine rechte Finite-Simple-Shear-Deformation

$$F = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} \cosh(2\alpha) & \sinh(2\alpha) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix} \quad \text{mit } \alpha \in \mathbb{R}. \quad (7.2.25)$$

*Beweis.* Seien  $T$  ein Pure-Shear-Spannungstensor mit  $s \neq 0$  und  $P \in \text{Sym}^+(3)$  mit  $\lambda_1 = \lambda$ ,  $\lambda_2 = \frac{1}{\lambda}$  und  $\lambda_3 = 1$  für ein  $\lambda \in \mathbb{R}_+$ , wobei  $\lambda_1^2, \lambda_2^2, \lambda_3^2$  die Eigenwerte von  $P$  sind. Nach Lemma 7.2.10 wissen wir, dass  $P$  und  $T$  genau dann kommutieren, wenn  $P$  von der Form (7.2.20) ist, also

$$\sqrt{P} = \begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ \sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{mit } \alpha \in \mathbb{R}$$

eine finite Pure-Shear-Dehnung ist.

Aufgrund von Folgerung 4.3.22 auf Seite 112 wissen wir, dass für ein isotropes Elastizitätsgesetz der linke Dehnungstensor  $B = FF^T$  stets mit  $\hat{\sigma}(B)$  und der rechte Dehnungstensor  $C = F^T F$  stets mit  $\hat{T}^{\text{Biot}}(C)$  kommutiert.

Für den Fall, dass  $\hat{\sigma}(B)$  eine Pure-Shear-Cauchy-Spannung ist, muss  $B = FF^T$  der Form (7.2.20) genügen (also ist  $V$  dann eine finite Pure-Shear-Dehnung). Mit Lemma 7.2.12 ist  $F$  bis auf eine Rotation  $Q \in \text{SO}(3)$  von rechts durch eine linke Finite-Simple-Shear-Deformation festgelegt.

Für den Fall, dass  $\hat{T}^{\text{Biot}}(C)$  eine Pure-Shear-Biot-Spannung ist, muss  $C = F^T F$  der Form (7.2.20) genügen (also ist  $U$  dann eine finite Pure-Shear-Dehnung). Mit Lemma 7.2.13 ist  $F$  bis auf eine Rotation  $Q \in \text{SO}(3)$  von links durch eine linke Finite-Simple-Shear-Deformation festgelegt.

Genau für  $Q = \mathbb{1}$  haben beide so entstandenen Deformationen die Eigenschaft der Bodenparallelität, d.h.  $e_1$  als Eigenvektor zu besitzen und wir erhalten zu einer Pure-Shear-Cauchy-Spannung die linke Finite-Simple-Shear-Deformation und zu einer Pure-Shear-Biot-Spannung die rechte Finite-Simple-Shear-Deformation. ■

Zum Abschluss beweisen wir noch einen auf den ersten Blick überraschenden Zusammenhang:

**Lemma 7.2.15.** Sei  $\alpha \in \mathbb{R}$  vorgegeben und

$$\sqrt{P} = \begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ \sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad R = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ -\sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix}. \quad (7.2.26)$$

Dann ist

$$\sqrt{P}R = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} 1 & \sinh(2\alpha) & 0 \\ 0 & \cosh(2\alpha) & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix} \quad (7.2.27)$$

die linke Finite-Simple-Shear-Deformation und

$$R\sqrt{P} = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} \cosh(2\alpha) & \sinh(2\alpha) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix} \quad (7.2.28)$$

die rechte Finite-Simple-Shear-Deformation zum vorgegebenen  $\alpha$ .

*Beweis.* Der Beweis erfolgt per Nachrechnen. Hierbei rufen wir uns die hyperbolischen Additionstheoreme

$$\cosh(\alpha)^2 + \sinh(\alpha)^2 = \cosh(2\alpha) \quad \text{und} \quad 2 \sinh(\alpha) \cosh(\alpha) = \sinh(2\alpha)$$

ins Gedächtnis. ■

**Bemerkung 7.2.16.** In anderen Worten ausgedrückt besagt das Lemma: Zu einem gegebenen  $\alpha \in \mathbb{R}$  bestehen die linke Finite-Simple-Shear-Deformation und die rechte Finite-Simple-Shear-Deformation aus genau der gleichen Dehnung und Rotation – nur wird im Fall der linken Finite-Simple-Shear-Deformation zunächst rotiert und dann gedehnt und im Fall der rechten Finite-Simple-Shear-Deformation zunächst gedehnt und dann rotiert.

### 7.3. Linearisierung von Finite-Simple-Shear

Die linke und rechte Finite-Simple-Shear-Deformationen tragen beide zu Recht den Namen Simple-Shear in sich, denn sie sind tatsächlich die finiten Entsprechungen der Simple-Shear-Deformation, wie der folgende Satz belegt.

**Satz 7.3.1.** Für jedes  $\alpha \in \mathbb{R}$  linearisiert sowohl die linke Finite-Simple-Shear-Deformation als auch die rechte Finite-Simple-Shear-Deformation zur Simple-Shear-Deformation mit  $\gamma = 2\alpha$ .

*Beweis.* Führen wir uns zunächst die Taylorentwicklungen der Matrixeinträge von Finite-Pure-Shear-Stretch, linker und rechter Finite-Simple-Shear-Deformation vor Augen:

$$\begin{aligned} \cosh(x) &= 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \mathcal{O}(x^6), & \frac{1}{\sqrt{\cosh(x)}} &= 1 - x^2 + \frac{7}{6}x^4 + \mathcal{O}(x^6), \\ \sinh(x) &= x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \mathcal{O}(x^6), & \frac{\cosh(x)}{\sqrt{\cosh(2x)}} &= 1 - \frac{1}{2}x^2 + \frac{17}{24}x^4 + \mathcal{O}(x^6), \\ \tanh(x) &= x - \frac{1}{3}x^3 + \frac{2}{15}x^5 + \mathcal{O}(x^6), & \frac{\sinh(x)}{\sqrt{\cosh(2x)}} &= x - \frac{5}{6}x^3 + \frac{121}{120}x^5 + \mathcal{O}(x^6), \\ \sqrt{\cosh(x)} &= 1 + x^2 - \frac{1}{6}x^4 + \mathcal{O}(x^6). \end{aligned}$$

## 7. Shear

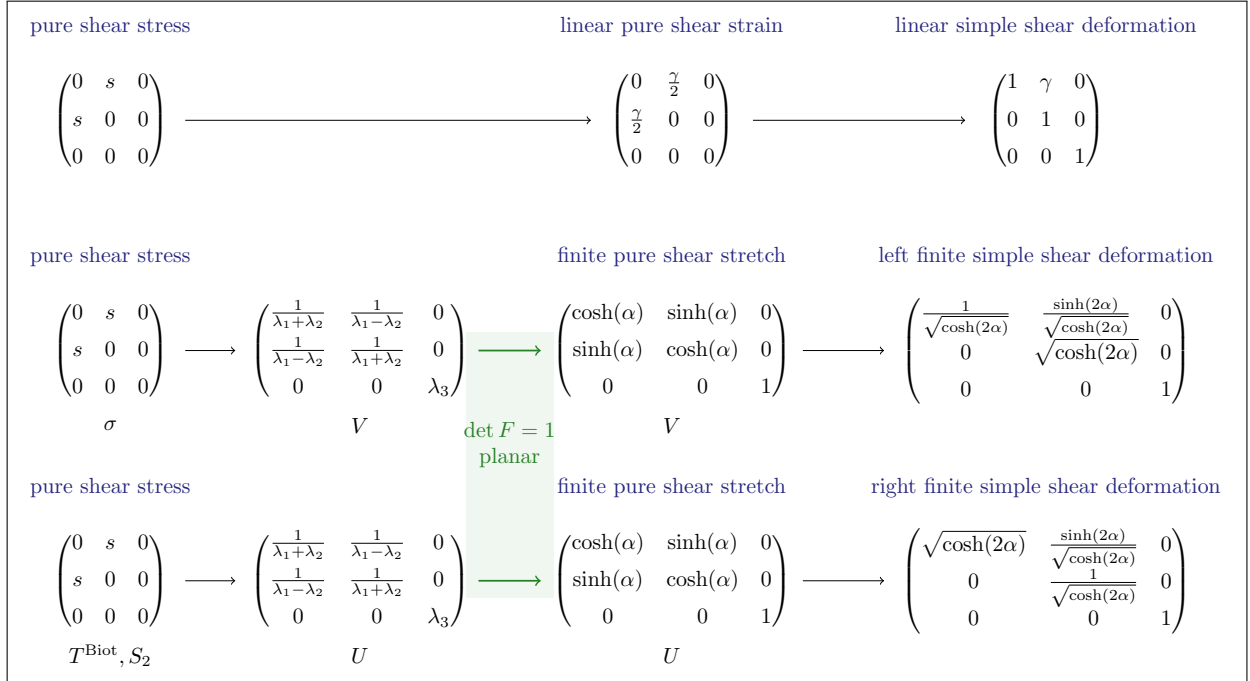


Abbildung 7.11.: Vergleich von Simple-Shear und Finite-Simple-Shear

So gilt für die Finite-Pure-Shear-Dehnung

$$P = \begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ \sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{1} + \begin{pmatrix} 0 & \alpha & 0 \\ \alpha & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mathcal{O}(\alpha^2),$$

sie linearisiert somit zur linearen Pure-Shear-Dehnung. Für die an beiden Finite-Simple-Shear-Deformationen beteiligte Rotation gilt

$$R = \begin{pmatrix} \frac{\cosh(\alpha)}{\sqrt{\cosh(2\alpha)}} & \frac{\sinh(\alpha)}{\sqrt{\cosh(2\alpha)}} & 0 \\ -\frac{\sinh(\alpha)}{\sqrt{\cosh(2\alpha)}} & \frac{\cosh(\alpha)}{\sqrt{\cosh(2\alpha)}} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{1} + \begin{pmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mathcal{O}(\alpha^2),$$

die Linearisierung entspricht dem schiefsymmetrischen Summanden der Zerlegung der Simple-Shear-Deformation.

Für die linke Finite-Simple-Shear-Deformation gilt

$$\begin{aligned} PR &= \begin{pmatrix} 1 & \alpha & 0 \\ \alpha & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \alpha & 0 \\ -\alpha & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \mathcal{O}(\alpha^2) \\ &= \begin{pmatrix} -\alpha^2 + 1 & 2\alpha & 0 \\ 0 & \alpha^2 + 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \mathcal{O}(\alpha^2) = \begin{pmatrix} 1 & 2\alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \mathcal{O}(\alpha^2). \end{aligned}$$

Für die rechte Finite-Simple-Shear-Deformation gilt

$$\begin{aligned} RP &= \begin{pmatrix} 1 & \alpha & 0 \\ -\alpha & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \alpha & 0 \\ \alpha & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \mathcal{O}(\alpha^2) \\ &= \begin{pmatrix} \alpha^2 + 1 & 2\alpha & 0 \\ 0 & -\alpha^2 + 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \mathcal{O}(\alpha^2) = \begin{pmatrix} 1 & 2\alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \mathcal{O}(\alpha^2). \end{aligned}$$

Beide Deformationen linearisieren somit zur linearen Shear-Deformation.

Betrachten wir noch die Scherungsmaße:

$$\begin{aligned}\gamma_l &= \tan \vartheta_l = \tanh(2\alpha) = 2\alpha + \mathcal{O}(\alpha^2), \\ \gamma_r &= \tan \vartheta_r = \sinh(2\alpha) = 2\alpha + \mathcal{O}(\alpha^2).\end{aligned}$$

Sowohl das linke als auch das rechte Scherungsmaß linearisieren zum Scherungsmaß der Simple-Shear-Deformation. ■

## 7.4. Konstitutive Bedingungen damit Pure-Shear-Dehnungen auf Pure-Shear-Spannungen führen

Wir haben im vorigen Abschnitt gezeigt, dass für eine linke Simple-Finite-Shear-Deformation

$$F = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} 1 & \sinh(2\alpha) & 0 \\ 0 & \cosh(2\alpha) & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix} \quad \text{mit } \alpha \in \mathbb{R} \quad (7.4.1)$$

alle Eigenvektoren von  $B = FF^T$  bzw. der zugehörigen Pure-Shear-Dehnung

$$V = \sqrt{B} = \begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ \sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7.4.2)$$

auch Eigenvektoren einer Pure-Shear-Spannung

$$T = \begin{pmatrix} 0 & s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{mit } s \in \mathbb{R} \quad (7.4.3)$$

sind, wir also gemeinsam mit

$$Q = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \text{SO}(3) \quad (7.4.4)$$

diagonalisieren können zu

$$V = Q \text{diag}\left(\lambda, \frac{1}{\lambda}, 1\right) Q^T \quad \text{bzw.} \quad B = Q \text{diag}\left(\lambda^2, \frac{1}{\lambda^2}, 1\right) Q^T \quad \text{mit} \quad \lambda = e^\alpha \quad (7.4.5)$$

und  $T = Q \text{diag}(s, -s, 1) Q^T$ . Ob ein konstitutives Gesetz  $\hat{\sigma}: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$  nun solch ein  $B$  auf solch ein  $T$  abbildet, hängt aufgrund der Isotropie von  $\hat{\sigma}$  ausschließlich davon ab, ob

$$\hat{\sigma}\left(\begin{pmatrix} \lambda^2 & 0 & 0 \\ 0 & \frac{1}{\lambda^2} & 0 \\ 0 & 0 & 1 \end{pmatrix}\right) = \begin{pmatrix} s & 0 & 0 \\ 0 & -s & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (7.4.6)$$

gilt. Oder anders ausgedrückt: Ob die Cauchy-Spannung  $\hat{\sigma}(FF^T)$  eines Simple-Finite-Shears eine Pure-Shear-Spannung ist, hängt einzig und allein von den Eigenwerten  $\sigma_1, \sigma_2, \sigma_3$  von  $\hat{\sigma}(FF^T)$  ab.

Insbesondere gilt: Für eine gegebene Spannungsantwort induziert eine Pure-Shear-Dehnung genau dann *immer* eine Pure-Shear-Spannung, wenn für alle  $\lambda \in \mathbb{R}_+$  ein  $s \in \mathbb{R}$  existiert, sodass aus  $(\lambda_1, \lambda_2, \lambda_3) = (\lambda, \frac{1}{\lambda}, 1)$  stets  $(\sigma_1, \sigma_2, \sigma_3) = (s, -s, 0)$  folgt, wobei  $\lambda_1, \lambda_2, \lambda_3$  die Singulärwerte von  $F$  sind und  $\sigma_i$  den  $i$ -ten Eigenwert von  $\hat{\sigma}(FF^T)$  bezeichnet.

Im Folgenden gehen wir von der *Hyperelastizität* des Elastizitätsgesetzes aus, d.h.  $\sigma$  ist induziert durch eine elastische Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$ , genauer  $\sigma(F) = DW(F)(\text{Cof } F)^{-1}$  für alle  $F \in \text{GL}^+(3)$ .



## 7. Shear

Zur Erinnerung: In Lemma 4.3.28 auf Seite 113 haben wir gezeigt, dass es zu einem gegebenen *isotropen* und *hyperelastischen* Elastizitätsgesetz  $\sigma: \text{Sym}^+(3) \rightarrow \text{Sym}(3)$ , stets eine nur von den Singulärwerten  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_+$  von  $F \in \text{GL}^+(3)$  abhängende Energiefunktion  $g: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  (und analog eine nur von den Eigenwerten  $\lambda_1^2, \lambda_2^2, \lambda_3^2 \in \mathbb{R}_+$  von  $FF^T$  abhängende Energiefunktion  $\widehat{g}: \mathbb{R}_+^3 \rightarrow \mathbb{R}$ ) existieren muss, sodass die  $\sigma$  induzierende Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  festgelegt ist durch

$$W(F) = g(\lambda_1, \lambda_2, \lambda_3) = \widehat{g}(\lambda_1^2, \lambda_2^2, \lambda_3^2). \quad (7.4.7)$$

Das Elastizitätsgesetz  $\sigma: \mathbb{R}_+^3 \rightarrow \mathbb{R}^3$  lässt sich dann ausdrücken durch

$$\sigma_i = \frac{\lambda_i}{\lambda_1 \lambda_2 \lambda_3} \cdot \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) \quad (7.4.8)$$

$$= \frac{2 \lambda_i^2}{\sqrt{\lambda_1^2 \lambda_2^2 \lambda_3^2}} \cdot \frac{\partial \widehat{g}}{\partial \lambda_i^2}(\lambda_1^2, \lambda_2^2, \lambda_3^2), \quad (7.4.9)$$

wobei die  $\sigma_i$  die Eigenwerte von  $\sigma$ , die sogenannten *Hauptspannungen*, sind.

### 7.4.1. Pure-Shear-Cauchy-Spannungen

Wir wollen nun isotrope elastische Energiefunktionen  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  untersuchen, die garantieren, dass das induzierte isotrope Elastizitätsgesetz  $\sigma$  Pure-Shear-Dehnungen  $V \in \text{Sym}^+(3)$  auf Pure-Shear-Spannungen  $T \in \text{Sym}(3)$  abbildet, also dass es zu jedem  $\lambda \in \mathbb{R}_+$  ein  $s \in \mathbb{R}$  mit  $\sigma(\text{diag}(\lambda, \frac{1}{\lambda}, 1)) = \text{diag}(s, -s, 0)$  gibt.

Zunächst definieren wir Energiefunktionen  $W$ , deren Argument  $F \in \text{GL}^+(3)$  wir in einen nur formändernden (und damit volumenerhaltenden) Teil  $\frac{F}{(\det F)^{1/3}} \in \text{SL}(3)$  und den Faktor der Volumenänderung  $\det F \in \mathbb{R}_+$  zerlegen und getrennt auswerten können.

**Definition 7.4.1.** Eine Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  weist einen *additiv isochor-volumetrischen Split* auf, wenn es stetig differenzierbare Funktionen  $W_{\text{iso}}: \text{SL}(3) \rightarrow \mathbb{R}$  und  $f: \mathbb{R}_+ \rightarrow \mathbb{R}$  gibt, sodass

$$W(F) = W_{\text{iso}}\left(\frac{F}{(\det F)^{1/3}}\right) + f(\det F) \quad \text{für alle } F \in \text{GL}^+(3). \quad (7.4.10)$$

**Proposition 7.4.2.** Gegeben sei eine isotrope elastische Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$ , die einen *additiven isochor-volumetrischen Split* aufweist. Dann ist die *Spur der Cauchy-Spannung* eine Funktion in der Determinante, d.h. es existiert eine Funktion  $\mathcal{F}: \mathbb{R}_+ \rightarrow \mathbb{R}$  mit

$$\text{tr}(\sigma(V)) = \mathcal{F}(\det V) \quad \text{für alle } V \in \text{Sym}^+(3). \quad (7.4.11)$$

*Beweis.* Die Funktion  $W$  lässt sich aufgrund ihrer Isotropie als Funktion in den Singulärwerten schreiben. Die Determinante lässt sich mit  $\det F = \lambda_1 \lambda_2 \lambda_3$  ebenso als Funktion in den Singulärwerten schreiben. Somit muss sich  $F \mapsto W_{\text{iso}}\left(\frac{F}{(\det F)^{1/3}}\right)$  als Differenz der beiden Funktionen auch als Funktion in den Singulärwerten schreiben lassen; also gibt es eine Funktion  $g_{\text{iso}}: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  mit  $g_{\text{iso}}(\lambda_1, \lambda_2, \lambda_3) = W_{\text{iso}}\left(\frac{F}{(\det F)^{1/3}}\right)$ . Offenkundig ist  $g_{\text{iso}}$  isochor, also  $t \mapsto g_{\text{iso}}(t \lambda_1, t \lambda_2, t \lambda_3)$  konstant und damit

$$0 = \frac{d}{dt} g_{\text{iso}}(t \lambda_1, t \lambda_2, t \lambda_3) = \sum_{i=1}^3 \lambda_i \frac{\partial g_{\text{iso}}}{\partial \lambda_i}(t \lambda_1, t \lambda_2, t \lambda_3) \quad \text{für alle } t \in \mathbb{R}_+.$$

Nun gilt mit  $t = 1$

$$\begin{aligned}\sigma_1 + \sigma_2 + \sigma_3 &= \frac{1}{\lambda_1 \lambda_2 \lambda_3} \sum_{i=1}^3 \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) \\ &= \frac{1}{\lambda_1 \lambda_2 \lambda_3} \underbrace{\sum_{i=1}^3 \lambda_i \frac{\partial g_{\text{iso}}}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3)}_{=0} + \frac{1}{\lambda_1 \lambda_2 \lambda_3} \sum_{i=1}^n \lambda_i \underbrace{\frac{\partial f}{\partial \lambda_i}(\lambda_1 \lambda_2 \lambda_3)}_{= \frac{\lambda_1 \lambda_2 \lambda_3}{\lambda_i} f'(\lambda_1 \lambda_2 \lambda_3)} \\ &= 3 f'(\lambda_1 \lambda_2 \lambda_3),\end{aligned}$$

also  $\text{tr}(\sigma(B)) = 3 f'(\det V)$ . Setzen wir  $\mathcal{F} = 3 f'$ , sind wir fertig.  $\blacksquare$

Wir können nun für Energien der Form (7.4.10) ein einfaches Kriterium angeben, das sicherstellt, dass aus Pure-Shear-Dehnungen immer Pure-Shear-Cauchy-Spannungen folgen.

**Lemma 7.4.3.** *Gegeben sei eine isotrope elastische Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$ , die einen additiven isochor-volumetrischen Split aufweist. Ist  $W$  streckungs-stauchungs-symmetrisch (tension-compression symmetric), d.h.  $W(F^{-1}) = W(F)$  für alle  $F \in \text{GL}^+(3)$ , dann gilt für alle Pure-Shear-Dehnungen  $V$ , dass  $\sigma(V)$  eine Pure-Shear-Spannung ist.*

*Beweis.* Sei  $\lambda_1 = \lambda$ ,  $\lambda_2 = \frac{1}{\lambda}$  und  $\lambda_3 = 1$ . Wie in Proposition 7.4.2 gezeigt ist die Spur von  $\sigma(B)$  eine Funktion  $\mathcal{F}$  in der Determinanten von  $V$ , also invariant gegenüber isochoren Veränderungen von  $B$ . Wegen  $\det B = \lambda_1 \lambda_2 \lambda_3 = 1$  und  $\sigma(\mathbb{1}) = 0$  gilt

$$\sigma_1 + \sigma_2 + \sigma_3 = \text{tr}(\sigma(V)) = \mathcal{F}(1) = \text{tr}(\sigma(\mathbb{1})) = 0.$$

Können wir nun  $\sigma_3 = 0$  zeigen, so folgt  $\sigma_1 = -\sigma_2$  und wir sind fertig. Hierzu definieren wir  $p: \mathbb{R}_+ \rightarrow \mathbb{R}$  mit  $p(t) := g(\lambda, \frac{1}{\lambda}, t)$  und der Ableitung  $p'(t) = \frac{\partial g}{\partial \lambda_3}(\lambda, \frac{1}{\lambda}, t)$ . Nun ist

$$W\left(\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \frac{1}{\lambda} & 0 \\ 0 & 0 & t \end{pmatrix}\right) \underset{\text{tension compr. sym.}}{=} W\left(\begin{pmatrix} \frac{1}{\lambda} & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \frac{1}{t} \end{pmatrix}\right) \underset{\text{sym. von } g}{=} W\left(\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \frac{1}{\lambda} & 0 \\ 0 & 0 & \frac{1}{t} \end{pmatrix}\right),$$

woraus  $p(t) = p(\frac{1}{t})$  für alle  $t \in \mathbb{R}_+$  folgt. Leiten wir beide Seiten ab, erhalten wir

$$p'(t) = \frac{d}{dt} p(t) = \frac{d}{dt} p\left(\frac{1}{t}\right) = -\frac{1}{t^2} p'\left(\frac{1}{t}\right) \quad \text{für alle } t \in \mathbb{R}_+.$$

Setzen wir  $t = 1$  ein, erhalten wir über  $p'(1) = -p'(1)$  den Wert der Ableitung  $0 = p'(1) = \frac{\partial g}{\partial \lambda_3}(\lambda, \frac{1}{\lambda}, 1)$ . Also ist

$$\sigma_3 = \frac{\lambda_3}{\lambda_1 \lambda_2 \lambda_3} \frac{\partial g}{\partial \lambda_3}\left(\lambda, \frac{1}{\lambda}, 1\right) = 0. \quad \blacksquare$$

**Definition 7.4.4.** Eine Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  hat die (verallgemeinerte) *Valanis-Landel Form*, wenn es differenzierbare Funktionen  $w, f: \mathbb{R}_+ \rightarrow \mathbb{R}$  mit  $f'(1) = w'(1) = 0$  gibt, sodass

$$W(F) = \sum_{i=1}^3 \underbrace{w(\lambda_i)}_{\text{Achsterm}} + \underbrace{f(\det F)}_{\text{Volumenterm}} \quad \text{für alle } F \in \text{GL}^+(3). \quad (7.4.12)$$

**Bemerkung 7.4.5.** Offenkundig ist  $W$  isotrop. Die Bedingung  $f'(1) = w'(1) = 0$  stellt sicher, dass  $F = \mathbb{1}$  eine spannungsfreie Referenzkonfiguration ist. Wir bemerken, dass eine Energie der Valanis-Landel Form im Allgemeinen keinen isochor-isometrischen Split aufweist.

**Lemma 7.4.6.** *Gegeben sei eine isotrope elastische Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  von der (verallgemeinerten) Valanis-Landel Form. Ist  $w$  streckungs-stauchungs-symmetrisch, d.h.  $w(\frac{1}{t}) = w(t)$  für alle  $t \in \mathbb{R}_+$ , dann gilt für alle Pure-Shear-Dehnungen  $V$ , dass  $\sigma(V)$  eine Pure-Shear-Spannung ist.*

## 7. Shear

*Beweis.* Sei  $\lambda_1 = \lambda$ ,  $\lambda_2 = \frac{1}{\lambda}$  und  $\lambda_3 = 1$ . Wir setzen  $g(\lambda_1, \lambda_2, \lambda_3) = W(F)$ , dann ist

$$\frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) = w'(\lambda_i) + \frac{\lambda_1 \lambda_2 \lambda_3}{\lambda_i} f'(\lambda_1 \lambda_2 \lambda_3)$$

und damit

$$\sigma_i = \frac{\lambda_i}{\lambda_1 \lambda_2 \lambda_3} \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) = \frac{\lambda_i}{\lambda_1 \lambda_2 \lambda_3} w'(\lambda_i) + f'(\lambda_1 \lambda_2 \lambda_3).$$

Einsetzen ergibt

$$\begin{aligned} \sigma_1 &= \lambda w'(\lambda) + f'(1) = \lambda w'(\lambda), \\ \sigma_2 &= \frac{1}{\lambda} w'\left(\frac{1}{\lambda}\right) + f'(1) = \frac{1}{\lambda} w'\left(\frac{1}{\lambda}\right), \\ \sigma_3 &= w'(1) + f'(1) = 0. \end{aligned}$$

Leiten wir beide Seiten der Bedingung  $w\left(\frac{1}{t}\right) = w(t)$  ab, erhalten wir

$$w'(t) = \frac{d}{dt} w(t) = \frac{d}{dt} w\left(\frac{1}{t}\right) = -\frac{1}{t^2} w'\left(\frac{1}{t}\right) \quad \text{für alle } t \in \mathbb{R}_+,$$

und damit  $\frac{1}{\lambda} w'\left(\frac{1}{\lambda}\right) = -\lambda w'(\lambda)$ , also  $\sigma_1 = -\sigma_2$ . ■

Im letzten Kapitel haben wir Energiefunktionen vom Hencky-Typ definiert. Wir erinnern uns:

**Definition 7.4.7** (Energiefunktion vom Hencky-Typ). Wir bezeichnen eine elastische Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  als vom *Hencky-Typ*, wenn sich  $W$  als Funktion in  $\|\text{dev}_3 \log U\|^2$  und  $|\text{tr} \log U|^2$  schreiben lässt, d.h. wenn  $W$  von der Form

$$W(F) = \mathcal{W}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) \tag{7.4.13}$$

mit einer differenzierbaren Funktion  $\mathcal{W}: \mathbb{R}_+^2 \rightarrow \mathbb{R}$  ist. Hierbei bezeichnet  $\log U = \log \sqrt{F^T F}$  den *Hencky-Dehnungstensor*.

**Definition 7.4.8** (quadratische Hencky-Energie). Wir bezeichnen die *quadratische Hencky-Energie*  $W_H: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$\begin{aligned} W_H(U) &= \mu \|\text{dev}_3 \log U\|^2 + \frac{\kappa}{2} (\text{tr}(\log U))^2 \\ &= \mu \left\| \log \frac{U}{\det U^{1/3}} \right\|^2 + \frac{\kappa}{2} (\log \det U)^2 \\ &= \mu \|\log U\|^2 + \frac{\Lambda}{2} (\text{tr}(\log U))^2. \end{aligned} \tag{7.4.14}$$

Hierbei sind  $\mu, \lambda \in \mathbb{R}$  mit  $\mu > 0$  und  $3\Lambda + 2\mu \geq 0$  die sogenannten *Lamé-Parameter* und  $\kappa \in \mathbb{R}_+$  der sogenannte *Bulk-Modulus*.

**Lemma 7.4.9.** *Gegeben sei eine isotrope elastische Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  vom Hencky-Typ. Dann gilt für alle Pure-Shear-Dehnungen  $V$ , dass  $\sigma(V)$  eine Pure-Shear-Spannung ist.*

*Beweis.* Wir setzen  $g(\lambda_1, \lambda_2, \lambda_3) = W(F)$ . Nach Proposition 2.5.38 auf Seite 37 können wir  $\|\text{dev}_3 \log U\|^2$  und  $|\text{tr} \log U|^2$  leicht in den Eigenwerten ausdrücken

$$g(\lambda_1, \lambda_2, \lambda_3) = \mathcal{W}\left(\underbrace{\sum_{k=1}^3 \left(\log \frac{\lambda_k}{\sqrt[3]{\lambda_1 \lambda_2 \lambda_3}}\right)^2}_{=\|\text{dev}_3 \log U\|^2}, \underbrace{(\log \lambda_1 \lambda_2 \lambda_3)^2}_{=|\text{tr}(\log U)|^2}\right).$$

## 7. Shear

Folgerung 2.5.37 liefert uns die Ableitungen von  $\|\text{dev}_3 \log U\|^2$  und  $|\text{tr}(\log u)|^2$  nach den Eigenwerten, sodass wir per Kettenregel die partiellen Ableitungen von  $g$  bestimmen können:

$$\begin{aligned} \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) &= \frac{\partial \mathcal{W}}{\partial x_1}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) \cdot \frac{2}{\lambda_i} \log \frac{\lambda_i}{(\det U)^{1/3}} \\ &\quad + \frac{\partial \mathcal{W}}{\partial x_2}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2) \cdot \frac{2}{\lambda_i} \text{tr} \log U. \end{aligned}$$

Einsetzen von  $\lambda_1 = \lambda$ ,  $\lambda_2 = \frac{1}{\lambda}$  und  $\lambda_3 = 1$  ergibt

$$\begin{aligned} \|\text{dev}_3 \log U\|^2 &= \frac{1}{3}((\log \lambda^2)^2 + (\log \lambda^{-2})^2 + (\log \lambda)^2) = 3(\log \lambda)^2, \\ |\text{tr} \log U|^2 &= \left| \log \lambda \frac{1}{\lambda} 1 \right|^2 = 0 \end{aligned}$$

also

$$\frac{\partial g}{\partial \lambda_i}(\lambda, \frac{1}{\lambda}, 1) = \frac{\partial \mathcal{W}}{\partial x_1}(3(\log \lambda)^2, 0) \cdot \frac{2}{\lambda_i} \log \lambda_i$$

und damit

$$\sigma_i = \lambda_i \frac{\partial g}{\partial \lambda_i}(\lambda, \frac{1}{\lambda}, 1) = 2 \log \lambda_i \frac{\partial \mathcal{W}}{\partial x_1}(3(\log \lambda)^2, 0).$$

Wir prüfen für  $i \in \{1, 2, 3\}$

$$\begin{aligned} \sigma_1 &= 2 \log \lambda \frac{\partial \mathcal{W}}{\partial x_1}(3(\log \lambda)^2, 0), \\ \sigma_2 &= 2 \log \frac{1}{\lambda} \frac{\partial \mathcal{W}}{\partial x_1}(3(\log \lambda)^2, 0) = -2 \log \lambda \frac{\partial \mathcal{W}}{\partial x_1}(3(\log \lambda)^2, 0), \\ \sigma_3 &= 2 \log 1 \frac{\partial \mathcal{W}}{\partial x_1}(3(\log \lambda)^2, 0) = 0 \end{aligned}$$

und sehen, dass  $\sigma_1 = -\sigma_2$  und  $\sigma_3 = 0$  sichergestellt ist. ■

**Folgerung 7.4.10.** Die Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  sei über Funktionen  $h, \tilde{f}: \mathbb{R}_+ \rightarrow \mathbb{R}$  bestimmt durch

$$W(F) = h(\|\text{dev}_3 \log U\|^2) + \tilde{f}(|\text{tr} \log U|^2). \quad (7.4.15)$$

Dann gilt für alle pure shear stretches  $V$ , dass  $\sigma(V)$  ein pure shear stress ist.

*Beweis.* Die Folgerung ist sowohl ein Spezialfall von Lemma 7.4.9 als auch von Lemma 7.4.3. Für das erste setzen wir einfach  $\overline{W}(x, y) = h(x) + \tilde{f}(y)$ . Im zweiten Fall können wir wegen  $\|\text{dev}_3 \log U\| = \left\| \log \frac{U}{(\det U)^{1/3}} \right\|$  die Funktion  $W_{\text{iso}} = h(\|\log \sqrt{X^T X}\|^2)$  und wegen  $\text{tr} \log U = \log \det U$  die Funktion  $f(x) = \tilde{f}(\log^2 x)$  setzen und erhalten damit den isochor-volumetrischen Split

$$W(F) = W_{\text{iso}}\left(\frac{F}{(\det F)^{1/3}}\right) + f(\det F).$$

Nun zeigen wir noch die Dehnungs-Stauchungs-Symmetrie von  $W$ , d.h.  $W(F^{-1}) = W(F)$ . Für den ersten Summanden gilt

$$W_{\text{iso}}\left(\frac{F}{(\det F)^{1/3}}\right) = h\left(\left\| \log \frac{\sqrt{F^T F}}{(\det F)^{1/3}} \right\|^2\right).$$

## 7. Shear

Mit  $F = RU$  ist

$$\begin{aligned}\sqrt{(F^{-1})^T F^{-1}} &= \sqrt{(RU)^{-T} (RU)^{-1}} \\ &= \sqrt{RU^{-T} U^{-1} R^{-1}} \\ &= \sqrt{RU^{-2} R^T} \\ &= R\sqrt{U^{-2} R^T} \\ &= RU^{-1} R^T\end{aligned}$$

und damit

$$\begin{aligned}W_{\text{iso}}\left(\frac{F^{-1}}{(\det F^{-1})^{1/3}}\right) &= h\left(\left\|\log \frac{RU^{-1} R^T}{(\det F^{-1})^{1/3}}\right\|^2\right) \\ &= h\left(\left\|\log \frac{U^{-1}}{(\det F)^{-1/3}}\right\|^2\right) \\ &= h\left(\left\|-\log \frac{U}{(\det F)^{1/3}}\right\|^2\right) \\ &= h\left(\left\|\log \frac{U}{(\det F)^{1/3}}\right\|^2\right) = W_{\text{iso}}\left(\frac{F}{(\det F)^{1/3}}\right).\end{aligned}$$

Für den zweiten Summanden ist

$$f(\det F^{-1}) = f((\det F)^{-1}) = \tilde{f}(\log^2(\det F)^{-1}) = \tilde{f}(\log^2 \det F) = f(\det F).$$

Wir erhalten damit

$$W(F^{-1}) = W_{\text{iso}}\left(\frac{F^{-1}}{(\det F^{-1})^{1/3}}\right) + f(\det F^{-1}) = W_{\text{iso}}\left(\frac{F}{(\det F)^{1/3}}\right) + f(\det F) = W(F).$$

Damit sind auch die Voraussetzungen für Lemma 7.4.3 erfüllt. ■

**Folgerung 7.4.11.** Die Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  sei über die Funktion  $\tilde{h}: \mathbb{R}_+ \rightarrow \mathbb{R}$  bestimmt durch

$$W(F) = \tilde{h}(\|\log U\|^2). \quad (7.4.16)$$

Dann gilt für alle Pure-Shear-Dehnungen  $V$ , dass  $\sigma(V)$  eine Pure-Shear-Spannung ist.

*Beweis.* Die Folgerung ist ein Spezialfall von Lemma 7.4.9. Wir setzen  $\mathcal{W}(s, t) := \tilde{h}(s + \frac{1}{3}t)$ . Da mit Proposition 2.2.18  $\|\log U\|^2 = \|\text{dev}_3 \log U\|^2 + \frac{1}{3}|\text{tr}(\log U)|^2$  gilt, erhalten wir

$$\mathcal{W}(\|\text{dev}_3 \log U\|^2, |\text{tr}(\log U)|^2) = \tilde{h}(\|\log U\|^2). \quad \blacksquare$$

**Lemma 7.4.12.** Gegeben sei eine isotrope elastische Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  von der Form

$$W(F) = \overline{W}(\|\text{dev}_3 \log U\|^2, \text{tr} \log U) \quad (7.4.17)$$

mit einer differenzierbaren Funktion  $\overline{W}: \mathbb{R}_+^2 \rightarrow \mathbb{R}$ , wobei  $\log U = \log \sqrt{F^T F}$  den Hencky-Dehnungstensor bezeichnet. Zudem sei

$$\frac{\partial \overline{W}}{\partial x_2}(x_1, 0) = 0 \quad \text{für alle } x_1 \in \mathbb{R}_+. \quad (7.4.18)$$

Dann gilt für alle pure shear stretches  $V$ , dass  $\sigma(V)$  ein pure shear stress ist.

## 7. Shear

*Beweis.* Wir gehen analog zum Beweis von Lemma 7.4.9 vor, setzen abermals  $g(\lambda_1, \lambda_2, \lambda_3) = W(F)$  und erhalten

$$g(\lambda_1, \lambda_2, \lambda_3) = \overline{W} \left( \underbrace{\sum_{k=1}^3 \left( \log \frac{\lambda_k}{\sqrt[3]{\lambda_1 \lambda_2 \lambda_3}} \right)^2}_{=\|\text{dev}_3 \log U\|^2}, \underbrace{\log \lambda_1 \lambda_2 \lambda_3}_{=\text{tr}(\log U)} \right).$$

Analog zu Folgerung 2.5.37 auf Seite 37 ist  $\frac{d}{d\lambda_i} \text{tr} \log U = \frac{2}{\lambda_i}$  und wir bestimmen die partiellen Ableitungen:

$$\begin{aligned} \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) &= \frac{\partial \overline{W}}{\partial x_1}(\|\text{dev}_3 \log U\|^2, \text{tr} \log U) \cdot \frac{2}{\lambda_i} \log \frac{\lambda_i}{(\det U)^{1/3}} \\ &\quad + \frac{\partial \overline{W}}{\partial x_2}(\|\text{dev}_3 \log U\|^2, \text{tr} \log U) \cdot \frac{2}{\lambda_i}. \end{aligned}$$

Einsetzen von  $\lambda_1 = \lambda$ ,  $\lambda_2 = \frac{1}{\lambda}$  und  $\lambda_3 = 1$  ergibt

$$\|\text{dev}_3 \log U\|^2 = 3(\log \lambda)^2 \quad \text{und} \quad \text{tr} \log U = 0,$$

also

$$\frac{\partial g}{\partial \lambda_i}(\lambda, \frac{1}{\lambda}, 1) = \frac{\partial \overline{W}}{\partial x_1}(3(\log \lambda)^2, 0) \cdot \frac{2}{\lambda_i} \log \lambda_i + \frac{\partial \overline{W}}{\partial x_2}(3(\log \lambda)^2, 0) \cdot \frac{2}{\lambda_i};$$

anders als im vorigen Lemma verschwindet der Summand mit der zweiten partiellen Ableitung von  $\overline{W}$  nicht! Wir erhalten

$$\sigma_i = \lambda_i \frac{\partial g}{\partial \lambda_i}(\lambda, \frac{1}{\lambda}, 1) = 2 \log \lambda_i \frac{\partial \overline{W}}{\partial x_1}(3(\log \lambda)^2, 0) + 2 \frac{\partial \overline{W}}{\partial x_2}(\|\text{dev}_3 \log U\|^2, \text{tr} \log U).$$

und damit

$$\begin{aligned} \sigma_1 &= 2 \log \lambda \frac{\partial \overline{W}}{\partial x_1}(3(\log \lambda)^2, 0) + 2 \frac{\partial \overline{W}}{\partial x_2}(3(\log \lambda)^2, 0), \\ \sigma_2 &= 2 \log \frac{1}{\lambda} \frac{\partial \overline{W}}{\partial x_1}(3(\log \lambda)^2, 0) = -2 \log \lambda \frac{\partial \overline{W}}{\partial x_1}(3(\log \lambda)^2, 0) + 2 \frac{\partial \overline{W}}{\partial x_2}(3(\log \lambda)^2, 0), \\ \sigma_3 &= 2 \log 1 \frac{\partial \overline{W}}{\partial x_1}(3(\log \lambda)^2, 0) = 2 \frac{\partial \overline{W}}{\partial x_2}(3(\log \lambda)^2, 0) \end{aligned}$$

und sehen, dass  $\sigma_1 = -\sigma_2$  und  $\sigma_3 = 0$  genau dann gilt, wenn  $\frac{\partial \overline{W}}{\partial x_2}(3(\log \lambda)^2, 0) = 0$ . ■

**Folgerung 7.4.13.** Die Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  sei über Funktionen  $\mu, g: \mathbb{R} \rightarrow \mathbb{R}$  bestimmt durch

$$W(F) = \mu(\text{tr} \log U) \cdot (\|\log U\|^2) + g(\text{tr} \log U), \quad (7.4.19)$$

mit

$$\mu(x) := \mu_r - (\mu_r - \mu_0) \exp\left(\frac{\text{sgn}(x) \cdot |x|^{1+\varepsilon}}{\varepsilon_r}\right), \quad (7.4.20)$$

$$g(x) := \kappa_0(\varepsilon_c)^2 \left(1 - \left(1 + \frac{x}{\varepsilon_c}\right) \exp\left(-\frac{x}{\varepsilon_c}\right)\right), \quad (7.4.21)$$

mit Konstanten  $\mu_0, \mu_r, \kappa_0, \varepsilon_r, \varepsilon_c, \varepsilon \in \mathbb{R}_+$ .

Dann gilt für alle pure shear stretches  $V$ , dass  $\sigma(V)$  ein pure shear stress ist.

**Bemerkung 7.4.14.** Es ist  $\text{sgn}(x) \cdot |x|^{1+\varepsilon} \rightarrow x$  für  $\varepsilon \rightarrow 0$ , somit entspricht  $\lim_{\varepsilon \searrow 0} W(F)$  genau der Energie aus Henann und Anand [36] mit  $\mu(x) = \mu_r - (\mu_r - \mu_0) \exp(\frac{x}{\varepsilon_r})$ .

## 7. Shear

*Beweis.* Wir verwenden Lemma 7.4.12 und setzen  $W(F) = \overline{W}(\|\text{dev}_3 \log U\|^2, \text{tr} \log U)$  mit

$$\overline{W}(x_1, x_2) := \mu(x_2) \cdot x_1 + g(x_2).$$

und

$$\frac{\partial \overline{W}}{\partial x_2}(x_1, x_2) = \mu'(x_2) \cdot x_1 + g'(x_2).$$

Es ist

$$\begin{aligned} \mu'(x) &= -\frac{1}{\varepsilon_r}(\mu_r - \mu_0)(1 + \varepsilon) \text{sgn}(x) \cdot |x|^\varepsilon \exp\left(\frac{\text{sgn}(x) \cdot |x|^{1+\varepsilon}}{\varepsilon_r}\right), \\ g'(x) &= \kappa_0(\varepsilon_c)^2 \left( \frac{1}{\varepsilon_c} \exp\left(-\frac{x}{\varepsilon_c}\right) + \left(1 + \frac{x}{\varepsilon_c}\right) \left(-\frac{1}{\varepsilon_c}\right) \exp\left(-\frac{x}{\varepsilon_c}\right) \right) \\ &= -\kappa_0 x \exp\left(-\frac{x}{\varepsilon_c}\right) \end{aligned}$$

und damit

$$\mu'(0) = 0 \quad \text{und} \quad g'(0) = 0.$$

Wir erhalten somit

$$\frac{\partial \overline{W}}{\partial x_2}(3(\log \lambda)^2, 0) = \mu'(0) \cdot 3(\log \lambda)^2 + g'(0) = 0. \quad \blacksquare$$

### 7.4.2. Pure-Shear-Biot-Spannungen

Gegebenenfalls ist es auch interessant, die Bedingungen an eine Energiefunktion  $W$  zu kennen, sodass aus jeder Pure-Shear-Dehnung eine Pure-Shear-Belastung folgt, d.h. es folgt  $T^{\text{Biot}}$  ist eine Pure-Shear-Spannung.

Zur Erinnerung: Für den Biot-Spannungstensor  $T^{\text{Biot}}: \text{GL}^+(3) \rightarrow \text{Sym}(3)$  gilt definitionsgemäß

$$T^{\text{Biot}}(F) = \det F \cdot R^T \sigma(F) F^{-T}, \quad (7.4.22)$$

wobei  $F = VR$  die Polarzerlegung von  $F \in \text{GL}^+(3)$  mit  $V \in \text{Sym}^+(3)$  und  $R \in \text{O}(3)$  ist. Eingeschränkt auf  $\text{Sym}^+(3)$  erhalten wir  $T^{\text{Biot}}(U) = \det U \cdot \sigma(U) U^{-1}$ , was für isotrope Energiefunktionen  $W(F) = g(\lambda_1, \lambda_2, \lambda_3)$  auf

$$T_i^{\text{Biot}}(U) = \frac{\det U}{\lambda_i} \sigma_i(U) = \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) \quad (7.4.23)$$

führt.

**Lemma 7.4.15.** *Gegeben sei eine isotrope elastische Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  von der (verallgemeinerten) Valanis-Landel Form. Ist nun  $w'(\frac{1}{t}) = -w'(t)$  für alle  $t \in \mathbb{R}_+$ , dann gilt für alle Pure-Shear-Dehnungen  $V$ , dass  $T^{\text{Biot}}(B)$  eine Pure-Shear-Spannung ist.*

*Beweis.* Sei  $\lambda_1 = \lambda$ ,  $\lambda_2 = \frac{1}{\lambda}$  und  $\lambda_3 = 1$ . Wir setzen  $g(\lambda_1, \lambda_2, \lambda_3) = W(F)$ , dann ist

$$T_i^{\text{Biot}} = \frac{\partial g}{\partial \lambda_i}(\lambda_1, \lambda_2, \lambda_3) = w'(\lambda_i) + \frac{\lambda_1 \lambda_2 \lambda_3}{\lambda_i} f'(\lambda_1 \lambda_2 \lambda_3).$$

Einsetzen ergibt

$$\begin{aligned} T_1^{\text{Biot}} &= w'(\lambda) + \frac{1}{\lambda} f'(1) = w'(\lambda), \\ T_2^{\text{Biot}} &= w'\left(\frac{1}{\lambda}\right) + \lambda f'(1) = -w'(\lambda), \\ T_3^{\text{Biot}} &= w'(1) + f'(1) = 0. \end{aligned}$$

und damit  $T_1^{\text{Biot}} = -T_2^{\text{Biot}}$  und  $T_3^{\text{Biot}} = 0$ . \blacksquare

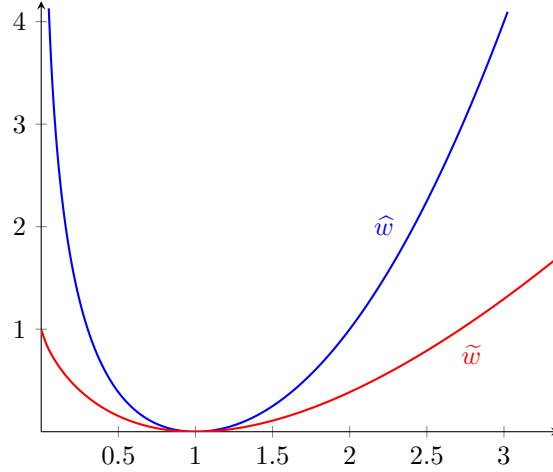


Abbildung 7.12.: Die Valanis-Landel Energieausdrücke  $\hat{w}$  und  $\tilde{w}$  wie in (7.4.25) bzw. (7.4.27) definiert.

**Bemerkung 7.4.16.** Mit der Bedingung  $w'(\frac{1}{t}) = -w'(t)$  für alle  $t \in \mathbb{R}_+$  ist  $w: \mathbb{R}_+ \rightarrow \mathbb{R}$  eindeutig durch die Abbildungsvorschrift auf  $[1, \infty)$  festgelegt, denn für alle  $t > 0$  gilt

$$w\left(\frac{1}{t}\right) = w(1) + \int_1^t \frac{d}{ds} w\left(\frac{1}{s}\right) ds = w(1) + \int_1^t \frac{1}{s^2} w'\left(\frac{1}{s}\right) ds = w(1) - \int_1^t \frac{w'(s)}{s^2} ds. \quad (7.4.24)$$

**Beispiel 7.4.17.** Die Funktion  $\hat{w}: \mathbb{R}_+ \rightarrow \mathbb{R}$  mit

$$\hat{w}(t) = \begin{cases} (t-1)^2 & \text{für } t \geq 1 \\ 2(t - \log t - 1) & \text{für } t < 1 \end{cases} \quad (7.4.25)$$

erfüllt die Bedingung  $\hat{w}'(\frac{1}{t}) = -\hat{w}'(t)$  für alle  $t \in \mathbb{R}_+$ .

**Beispiel 7.4.18.** Die isotrope Energiefunktion  $W: \text{GL}^+(3) \rightarrow \mathbb{R}$  mit

$$W(F) = 2\mu \langle U, \log U - \mathbb{1} \rangle + 3 \quad (7.4.26)$$

der (verallgemeinerten) Valanis-Landel Form  $W(F) = \sum_{i=1}^n \tilde{w}(\lambda_i) + f(\det U)$  mit  $f \equiv 0$  und

$$\tilde{w}(t) = 2\mu (t(\log t - 1) + 1) \quad (7.4.27)$$

erfüllt die Bedingung  $\tilde{w}'(\frac{1}{t}) = -\tilde{w}'(t)$  für alle  $t \in \mathbb{R}_+$ .

**Bemerkung 7.4.19.** Die Energiefunktion aus Beispiel 7.4.18 induziert das Beckersche Elastizitätsgesetz im querkontraktionsfreien Spezialfall [53]. Dieses konstitutive Gesetz wurde als erstes vom Geologen G.F. Becker [5] beschrieben, der es aus einer Reihe von Axiomen einschließlich der Äquivalenz von Pure-Shear-Dehnung und Pure-Shear-Biot-Spannung hergeleitet hat.

## 7.5. Polarzerlegung des Simple-Shear-Deformationsgradienten

**Lemma 7.5.1.** Gegeben sei ein Simple-Shear-Deformationsgradient  $F \in \text{GL}^+(3)$ . Dann lautet seine rechte Polarzerlegung:

$$\underbrace{\begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{F \in \text{GL}^+(3)} = \underbrace{\frac{1}{\sqrt{4+\gamma^2}} \cdot \begin{pmatrix} 2+\gamma^2 & \gamma & 0 \\ \gamma & 2 & 0 \\ 0 & 0 & \sqrt{4+\gamma^2} \end{pmatrix}}_{V \in \text{Sym}^+(3)} \cdot \underbrace{\frac{1}{\sqrt{4+\gamma^2}} \cdot \begin{pmatrix} 2 & \gamma & 0 \\ -\gamma & 2 & 0 \\ 0 & 0 & \sqrt{4+\gamma^2} \end{pmatrix}}_{R \in \text{SO}(3)}. \quad (7.5.1)$$



## 7. Shear

Und für die linke Polarzerlegung gilt:

$$\underbrace{\begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{F \in \text{GL}^+(3)} = \underbrace{\frac{1}{\sqrt{4+\gamma^2}} \cdot \begin{pmatrix} 2 & \gamma & 0 \\ -\gamma & 2 & 0 \\ 0 & 0 & \sqrt{4+\gamma^2} \end{pmatrix}}_{R \in \text{SO}(3)} \cdot \underbrace{\frac{1}{\sqrt{4+\gamma^2}} \cdot \begin{pmatrix} 2 & \gamma & 0 \\ \gamma & 2+\gamma^2 & 0 \\ 0 & 0 & \sqrt{4+\gamma^2} \end{pmatrix}}_{U \in \text{Sym}^+(3)}. \quad (7.5.2)$$

*Beweis.* Wir machen uns daran, den Verzerrungstensor (stretch)  $V$  zu bestimmen. Dies erreichen wir über das Wurzelziehen des Fingertensors

$$B = FF^T = \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+\gamma^2 & \gamma & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (7.5.3)$$

Das charakteristische Polynom von  $B$  lautet

$$\begin{aligned} \det(B - \mu \cdot \mathbb{1}) &= \begin{vmatrix} 1+\gamma^2 - \mu & \gamma & 0 \\ \gamma & 1 - \mu & 0 \\ 0 & 0 & 1 - \mu \end{vmatrix} = (1 - \mu)^2(1 + \gamma^2 - \mu) - (1 - \mu)\gamma^2 \\ &= (1 - \mu)((1 - \mu)(1 - \mu + \gamma^2) - \gamma^2) = (1 - \mu)((1 - \mu)^2 + \gamma^2 - \mu\gamma^2 - \gamma^2) \\ &= (1 - \mu)(1 - 2\mu + \mu^2 - \mu\gamma^2) = (1 - \mu)(\mu^2 + (-2 - \gamma^2)\mu + 1), \end{aligned}$$

seine Nullstellen sind somit  $\mu = 1$  und

$$\begin{aligned} \mu &= 1 + \frac{1}{2}\gamma^2 \pm \sqrt{\left(1 + \frac{1}{2}\gamma^2\right)^2 - 1} = 1 + \frac{1}{2}\gamma^2 \pm \sqrt{\gamma^2 + \frac{1}{4}\gamma^4} \\ &= \frac{1}{4}(\gamma^2 + 4 \pm 2\gamma\sqrt{\gamma^2 + 4} + \gamma^2) = \frac{1}{4}(\sqrt{\gamma^2 + 4} \pm \gamma)^2. \end{aligned}$$

Setzen wir  $\alpha := \frac{\sqrt{\gamma^2 + 4} + \gamma}{2} > 1$ , dann ist

$$\frac{1}{\alpha} = \frac{2}{\sqrt{\gamma^2 + 4} + \gamma} = \frac{\gamma^2 + 4 - \gamma^2}{2(\sqrt{\gamma^2 + 4} + \gamma)} = \frac{(\sqrt{\gamma^2 + 4} - \gamma)(\sqrt{\gamma^2 + 4} + \gamma)}{2(\sqrt{\gamma^2 + 4} + \gamma)} = \frac{\sqrt{\gamma^2 + 4} - \gamma}{2} < 1.$$

Die Eigenwerte von  $B$  lauten also  $(\mu_1, \mu_2, \mu_3) = (\alpha^2, 1, \frac{1}{\alpha^2})$  (die Singulärwerte von  $F$  sind  $(\lambda_1, \lambda_2, \lambda_3) = (\alpha, 1, \frac{1}{\alpha})$ ). Zu diesen Eigenwerten bestimmen wir nun die Eigenräume:

Für  $\mu_2$  ist die Rechnung recht kurz

$$\begin{aligned} Bx = \mu_2 \cdot x &\Leftrightarrow (B - 1 \cdot \mathbb{1})x = 0 \Leftrightarrow \begin{pmatrix} \gamma^2 & \gamma & 0 \\ \gamma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &\Leftrightarrow x \in \text{span} \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}. \end{aligned} \quad (7.5.4)$$

## 7. Shear

Im Falle von  $\mu_1$  und  $\mu_3$  brauchen wir etwas Vorarbeit:

$$\begin{aligned}
 Bx = \mu_{1,3} \cdot x &\Leftrightarrow (B - \lambda_{1,3} \cdot \mathbb{1})x = 0 \\
 \Leftrightarrow &\begin{pmatrix} 1 + \gamma^2 - (1 + \frac{1}{2}\gamma^2 \pm \sqrt{\gamma^2 + \frac{1}{4}\gamma^4}) & \gamma & 0 \\ \gamma & 1 - (1 + \frac{1}{2}\gamma^2 \pm \sqrt{\gamma^2 + \frac{1}{4}\gamma^4}) & 0 \\ 0 & 0 & 1 - (1 + \frac{1}{2}\gamma^2 + \sqrt{\gamma^2 + \frac{1}{4}\gamma^4}) \end{pmatrix} \\
 \Leftrightarrow &\begin{pmatrix} \frac{1}{2}\gamma(\gamma \mp \sqrt{4 + \gamma^2}) & \gamma & 0 \\ \gamma & -\frac{1}{2}\gamma(\gamma \pm \sqrt{4 + \gamma^2}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \Leftrightarrow &\begin{pmatrix} \frac{1}{2}(\gamma \mp \sqrt{4 + \gamma^2}) & 1 & 0 \\ 1 & -\frac{1}{2}(\gamma \pm \sqrt{4 + \gamma^2}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \Leftrightarrow &\begin{pmatrix} -\frac{1}{4}(\gamma \mp \sqrt{4 + \gamma^2})(\gamma \pm \sqrt{4 + \gamma^2}) & -\frac{1}{2}(\gamma \pm \sqrt{4 + \gamma^2}) & 0 \\ 1 & -\frac{1}{2}(\gamma \pm \sqrt{4 + \gamma^2}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & -\frac{1}{2}(\gamma \pm \sqrt{4 + \gamma^2}) & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
 \end{aligned}$$

Nun ergeben sich die Eigenräume zu

$$Bx = \alpha^2 \cdot x \Leftrightarrow x \in \text{span} \left\{ \begin{pmatrix} -\frac{1}{2}(\gamma + \sqrt{\gamma^2 + 4}) \\ -1 \\ 0 \end{pmatrix} \right\} = \text{span} \left\{ \begin{pmatrix} \alpha \\ 1 \\ 0 \end{pmatrix} \right\} \quad (7.5.5)$$

und

$$Bx = \frac{1}{\alpha^2} \cdot x \Leftrightarrow x \in \text{span} \left\{ \begin{pmatrix} -\frac{1}{2}(\gamma - \sqrt{\gamma^2 + 4}) \\ -1 \\ 0 \end{pmatrix} \right\} = \text{span} \left\{ \begin{pmatrix} \frac{1}{\alpha} \\ -1 \\ 0 \end{pmatrix} \right\}. \quad (7.5.6)$$

Es ist leicht zu sehen, dass

$$\left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \alpha \\ 1 \\ 0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{\alpha} \\ -1 \\ 0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} \alpha \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\alpha} \\ -1 \\ 0 \end{pmatrix} \right\rangle = 0$$

sowie

$$\left\| \begin{pmatrix} \alpha \\ 1 \\ 0 \end{pmatrix} \right\| = \sqrt{1 + \alpha^2}, \quad \left\| \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\| = 1 \quad \text{und} \quad \left\| \begin{pmatrix} \frac{1}{\alpha} \\ -1 \\ 0 \end{pmatrix} \right\| = \sqrt{\frac{1}{\alpha^2} + 1} = \frac{\sqrt{1 + \alpha^2}}{\alpha}$$

gilt. Somit spannen

$$\begin{pmatrix} \alpha \\ 1 \\ 0 \end{pmatrix} / \sqrt{1 + \alpha^2} = \begin{pmatrix} \frac{\alpha}{\sqrt{1 + \alpha^2}} \\ \frac{1}{\sqrt{1 + \alpha^2}} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \frac{1}{\alpha} \\ -1 \\ 0 \end{pmatrix} / \frac{\sqrt{1 + \alpha^2}}{\alpha} = \begin{pmatrix} \frac{1}{\sqrt{1 + \alpha^2}} \\ -\frac{\alpha}{\sqrt{1 + \alpha^2}} \\ 0 \end{pmatrix}$$

ein Orthonormalensystem aus Eigenvektoren zu  $B$  auf und wir erhalten

$$Q = \begin{pmatrix} \frac{\alpha}{\sqrt{1 + \alpha^2}} & 0 & \frac{1}{\sqrt{1 + \alpha^2}} \\ \frac{1}{\sqrt{1 + \alpha^2}} & 0 & -\frac{\alpha}{\sqrt{1 + \alpha^2}} \\ 0 & 1 & 0 \end{pmatrix}. \quad (7.5.7)$$

## 7. Shear

Wir können nun auch  $V$  bestimmen. Dazu multiplizieren wir:

$$\begin{aligned}
 V &= Q \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3) Q^T = \begin{pmatrix} \frac{\alpha}{\sqrt{1+\alpha^2}} & 0 & \frac{1}{\sqrt{1+\alpha^2}} \\ \frac{1}{\sqrt{1+\alpha^2}} & 0 & -\frac{\alpha}{\sqrt{1+\alpha^2}} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\alpha} \end{pmatrix} \begin{pmatrix} \frac{\alpha}{\sqrt{1+\alpha^2}} & \frac{1}{\sqrt{1+\alpha^2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{1+\alpha^2}} & -\frac{\alpha}{\sqrt{1+\alpha^2}} & 0 \end{pmatrix} \\
 &= \begin{pmatrix} \frac{\alpha^2}{\sqrt{1+\alpha^2}} & 0 & \frac{\alpha^{-1}}{\sqrt{1+\alpha^2}} \\ \frac{\alpha}{\sqrt{1+\alpha^2}} & 0 & -\frac{1}{\sqrt{1+\alpha^2}} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\alpha}{\sqrt{1+\alpha^2}} & \frac{1}{\sqrt{1+\alpha^2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{1+\alpha^2}} & -\frac{\alpha}{\sqrt{1+\alpha^2}} & 0 \end{pmatrix} \\
 &= \begin{pmatrix} \frac{\alpha^3}{1+\alpha^2} + \frac{\alpha^{-1}}{1+\alpha^2} & \frac{\alpha^2}{1+\alpha^2} - \frac{1}{1+\alpha^2} & 0 \\ \frac{\alpha^2}{1+\alpha^2} - \frac{1}{1+\alpha^2} & \frac{\alpha}{1+\alpha^2} + \frac{\alpha}{1+\alpha^2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} \frac{\alpha(\alpha^2+\alpha^{-2})}{1+\alpha^2} & \frac{\alpha^2-1}{1+\alpha^2} & 0 \\ \frac{\alpha^2-1}{1+\alpha^2} & \frac{2\alpha}{1+\alpha^2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7.5.8}
 \end{aligned}$$

Wir können noch etwas vereinfachen: Es ist

$$\alpha = \frac{1}{2}(\sqrt{\gamma^2+4} + \gamma) \quad \text{und} \quad \frac{1}{\alpha} = \frac{1}{2}(\sqrt{\gamma^2+4} - \gamma),$$

somit

$$\frac{1+\alpha^2}{\alpha} = \frac{1}{\alpha} + \alpha = \sqrt{\gamma^2+4} \quad \text{und} \quad \frac{\alpha^2-1}{\alpha} = \alpha - \frac{1}{\alpha} = \gamma. \tag{7.5.9}$$

Damit ist

$$\frac{2\alpha}{1+\alpha^2} = 2 \cdot \frac{1}{\sqrt{\gamma^2+4}} \quad \text{und} \quad \frac{\alpha^2-1}{1+\alpha^2} = \frac{\alpha^2-1}{\alpha} \cdot \frac{\alpha}{1+\alpha^2} = \gamma \cdot \frac{1}{\sqrt{\gamma^2+4}}. \tag{7.5.10}$$

Weiterhin ist

$$\alpha^2 + \frac{1}{\alpha^2} = \left( \alpha^2 - 2 \cdot \alpha \cdot \frac{1}{\alpha} + \frac{1}{\alpha^2} \right) + 2 = \left( \alpha - \frac{1}{\alpha} \right)^2 + 2 = \gamma^2 + 2 \tag{7.5.11}$$

und damit

$$\frac{\alpha(\alpha^2 + \alpha^{-2})}{1+\alpha^2} = \frac{\alpha^2 + \frac{1}{\alpha^2}}{\alpha + \frac{1}{\alpha}} = \frac{\gamma^2 + 2}{\sqrt{\gamma^2+4}}. \tag{7.5.12}$$

Insgesamt erhalten wir also

$$V = \begin{pmatrix} \frac{2+\gamma^2}{\sqrt{4+\gamma^2}} & \frac{\gamma}{\sqrt{4+\gamma^2}} & 0 \\ \frac{\gamma}{\sqrt{4+\gamma^2}} & \frac{2}{\sqrt{4+\gamma^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \frac{1}{\sqrt{4+\gamma^2}} \cdot \begin{pmatrix} 2+\gamma^2 & \gamma & 0 \\ \gamma & 2 & 0 \\ 0 & 0 & \sqrt{4+\gamma^2} \end{pmatrix}. \tag{7.5.13}$$

Mit der Kenntnis von  $V$  einschließlich seiner Diagonalisierung können wir leicht die Polarzerlegung  $F =$

## 7. Shear

$VR = RU$  bestimmen. Da benötigen wir zunächst die Inverse

$$\begin{aligned}
 V^{-1} &= Q \operatorname{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \lambda_3^{-1}) Q^T = \begin{pmatrix} \frac{\alpha}{\sqrt{1+\alpha^2}} & 0 & \frac{1}{\sqrt{1+\alpha^2}} \\ \frac{1}{\sqrt{1+\alpha^2}} & 0 & -\frac{\alpha}{\sqrt{1+\alpha^2}} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\alpha} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{pmatrix} \begin{pmatrix} \frac{\alpha}{\sqrt{1+\alpha^2}} & \frac{1}{\sqrt{1+\alpha^2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{1+\alpha^2}} & -\frac{\alpha}{\sqrt{1+\alpha^2}} & 0 \end{pmatrix} \\
 &= \begin{pmatrix} \frac{\alpha}{\sqrt{1+\alpha^2}} & 0 & \frac{1}{\sqrt{1+\alpha^2}} \\ \frac{1}{\sqrt{1+\alpha^2}} & 0 & -\frac{\alpha}{\sqrt{1+\alpha^2}} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{1+\alpha^2}} & \frac{\alpha^{-1}}{\sqrt{1+\alpha^2}} & 0 \\ 0 & 0 & 1 \\ \frac{\alpha}{\sqrt{1+\alpha^2}} & -\frac{\alpha^2}{\sqrt{1+\alpha^2}} & 0 \end{pmatrix} \\
 &= \begin{pmatrix} \frac{\alpha}{1+\alpha^2} + \frac{\alpha}{1+\alpha^2} & \frac{1}{1+\alpha^2} - \frac{\alpha^2}{1+\alpha^2} & 0 \\ \frac{1}{1+\alpha^2} - \frac{\alpha^2}{1+\alpha^2} & \frac{\alpha^{-1}}{1+\alpha^2} + \frac{\alpha^3}{1+\alpha^2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} \frac{2\alpha}{1+\alpha^2} & \frac{1-\alpha^2}{1+\alpha^2} & 0 \\ \frac{1-\alpha^2}{1+\alpha^2} & \frac{\alpha(\alpha^2+\alpha^{-2})}{1+\alpha^2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} \frac{2}{\sqrt{4+\gamma^2}} & -\frac{\gamma}{\sqrt{4+\gamma^2}} & 0 \\ -\frac{\gamma}{\sqrt{4+\gamma^2}} & \frac{2+\gamma^2}{\sqrt{4+\gamma^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \frac{1}{\sqrt{4+\gamma^2}} \cdot \begin{pmatrix} 2 & -\gamma & 0 \\ -\gamma & 2+\gamma^2 & 0 \\ 0 & 0 & \sqrt{4+\gamma^2} \end{pmatrix}. \tag{7.5.14}
 \end{aligned}$$

Wegen  $F = VR$  ist

$$\begin{aligned}
 R &= V^{-1}F = \frac{1}{\sqrt{4+\gamma^2}} \cdot \begin{pmatrix} 2 & -\gamma & 0 \\ -\gamma & 2+\gamma^2 & 0 \\ 0 & 0 & \sqrt{4+\gamma^2} \end{pmatrix} \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \frac{1}{\sqrt{4+\gamma^2}} \cdot \begin{pmatrix} 2 & \gamma & 0 \\ -\gamma & 2 & 0 \\ 0 & 0 & \sqrt{4+\gamma^2} \end{pmatrix}. \tag{7.5.15}
 \end{aligned}$$

Zudem ist  $F = RU$ , also

$$\begin{aligned}
 U &= R^T F = \frac{1}{\sqrt{4+\gamma^2}} \cdot \begin{pmatrix} 2 & -\gamma & 0 \\ \gamma & 2 & 0 \\ 0 & 0 & \sqrt{4+\gamma^2} \end{pmatrix} \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \frac{1}{\sqrt{4+\gamma^2}} \cdot \begin{pmatrix} 2 & \gamma & 0 \\ \gamma & 2+\gamma^2 & 0 \\ 0 & 0 & \sqrt{4+\gamma^2} \end{pmatrix}. \quad \blacksquare
 \end{aligned}$$

# 8. Zusammenfassung und Ausblick

## 8.1. Zusammenfassung der Haupt- und Nebenresultate

Im Rahmen dieser Arbeit wurden die folgenden neuen Ergebnisse erzielt:

### Beweis der Sum-of-Squared-Logarithms-Ungleichung im beliebigdimensionalen Fall

Im Rahmen dieser Arbeit wurde ein Beweis der Sum-of-Squared-Logarithms-Ungleichung (SSLI), der bis dato nur für  $n \leq 4$  vorlag, entwickelt. Dieser Beweis funktioniert für beliebige Dimensionen und verwendet recht elementare Konzepte der Mathematik und kann so leicht nachvollzogen werden. Wir beweisen: Seien  $n \in \mathbb{N}$  und  $x, y \in \mathbb{R}_+^n$ , sodass

$$e_k(x) \leq e_k(y) \quad \text{für alle } k \in \{1, \dots, n-1\} \quad \text{und} \quad e_n(x) = e_n(y).$$

Dann gilt

$$\sum_{i=1}^n (\log x_i)^2 \leq \sum_{i=1}^n (\log y_i)^2.$$

Für die Anwendung auf die nichtlineare Elastizitätstheorie wurde die folgende Matrix-Variante vorgestellt: Seien  $U, \tilde{U} \in \text{Sym}^+(n)$ , sodass

$$I_k(U) \leq I_k(\tilde{U}) \quad \text{für alle } k \in \{1, \dots, n-1\} \quad \text{und} \quad \det U = \det \tilde{U}.$$

Dann gilt

$$\|\log U\|^2 \leq \|\log \tilde{U}\|^2.$$

### Neue Kriterien, die ebenfalls die Gültigkeit der SSLI sicherstellen

Wir erforschen weitere Bedingungen, welche die Ungleichung  $\sum_{i=1}^n (\log x_i)^2 \leq \sum_{i=1}^n (\log y_i)^2$  sicherstellen. Diese Bedingungen sind keine Verschärfung der klassischen Bedingung an die elementarsymmetrischen Polynomen zweier Vektoren und erschließen so die Ungleichung auch für weitere Szenarien.

### Beweis der Sum-of-Powered-Logarithms-Ungleichung im beliebigdimensionalen Fall

Es wird eine mit der SSLI sehr verwandte Ungleichung präsentiert, dessen Bedingung ein weiteres Beispiel für die Relevanz des Konzeptes der auch in dieser Arbeit behandelten Majorization darstellt. Wir zeigen: Seien  $a, b \in \mathbb{R}^n$  mit  $a_i, b_i > 1$ . Angenommen  $a$  wird schwach von  $b$  majorisiert ( $a \prec^w b$ ), dann gilt

$$\sum_{i=1}^n (\log a_i)^p \leq \sum_{i=1}^n (\log b_i)^p, \quad \text{für alle } p < 0.$$

### Neue Interpretation der SSLI in Bezug auf die Hencky-Energie

Wir wenden die SSLI auf die Hencky-Energie  $W_H: \text{Sym}^+(3) \rightarrow \mathbb{R}$  mit

$$W_H(U) = \mu \|\log U\|^2 + \frac{\Lambda}{2} (\text{tr}(\log U))^2 \quad \text{für alle } U \in \text{Sym}^+(3)$$

an und zeigen die Monotonie in den Invarianten von  $U$ , d.h. von  $\Psi(I_1, I_2, I_3) = W(U)$ . Wir zeigen, dass  $\Psi$  unendlich oft differenzierbar ist und für alle  $U$  gilt:<sup>1</sup>

$$\frac{\partial \Psi}{\partial I_1}(I_1, I_2, I_3) \geq 0 \quad \text{und} \quad \frac{\partial \Psi}{\partial I_2}(I_1, I_2, I_3) \geq 0.$$

Wir führen die Klasse der Energien vom Hencky-Typ mit

$$W(F) = \mathcal{W}(\|\text{dev}_3 \log U\|^2, |\text{tr} \log U|^2)$$

ein und zeigen, dass unter Voraussetzung des monotonen Wachstums von  $\mathcal{W}$  bezüglich des ersten Argument die SSLI die Gültigkeit der Baker-Ericksen-Ungleichung impliziert.

### Strukturierte Herangehensweise an Shear und Finden der finiten Entsprechung des linearen Simple-Shear

Die auf eine Pure-Shear-Cauchy-Spannung bzw. Pure-Shear-Biot-Spannung führende Deformation (bestehend aus Rotation, triaxialer Dehnung und Shear bzw. Shear, triaxiale Dehnung und Rotation) wird von uns unter der Voraussetzung „Shear“ zu sein, also Volumenerhaltung, Planarität und Bodenparallelität<sup>2</sup>, als Produkt  $PR$  bzw.  $RP$  aus Pure-Shear-Dehnung  $P$  und Rotation  $R$  der Form

$$P = \begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ \sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad R = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) & 0 \\ -\sinh(\alpha) & \cosh(\alpha) & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix}.$$

ausgemacht. Wir führen den linken Simple-Finite-Shear-Deformationsgradienten

$$F = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} 1 & \sinh(2\alpha) & 0 \\ 0 & \cosh(2\alpha) & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix} \quad \text{mit } \alpha \in \mathbb{R}$$

und den rechten Simple-Finite-Shear-Deformationsgradienten

$$F = \frac{1}{\sqrt{\cosh(2\alpha)}} \begin{pmatrix} \cosh(2\alpha) & \sinh(2\alpha) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{\cosh(2\alpha)} \end{pmatrix} \quad \text{mit } \alpha \in \mathbb{R}$$

als finite Cauchy- bzw. Biot-Entsprechung des linearen Simple-Shear ein und belegen die Linearisierung zu Letzterem.

Zudem bietet die Arbeit die nachstehenden Nebenresultate:

- **Zusammenstellung und Erweiterung von Ergebnissen aus dem Bereich der konstitutiven Bedingungen mit Fokus auf empirische Ungleichungen und Semi-Invertierbarkeit von Spannungsantwortfunktionen**
- **Systematische Aufarbeitung des Modellierungsansatzes der nichtlinearen Elastizitätstheorie**
- **Eine hinreichende Bedingung, damit eine Lösung der Euler-Lagrange-Gleichung zum Energiefunktional globaler Minimierer ist**
- **Untersuchung von Energiefunktionen der Form  $\widehat{W}(C) = \alpha \text{tr}(C) + \beta \text{tr}(C)^2 + \gamma \text{tr}(C^2) - \delta \log \det C + \zeta$  mit  $\alpha, \beta, \gamma, \delta, \zeta \geq 0$  sodass  $D\widehat{W}(\mathbb{1}) = 0$  bezüglich der Konvexität in  $C$  und Anwendung obengenannter Resultate**

<sup>1</sup>Hier nehmen wir explizit die Vielfachen der Einheitsmatrix aus der Betrachtung heraus; ein Umstand, der uns keinerlei Probleme beim Schließen auf die verwendeten Aussagen bereitet.

<sup>2</sup>Es gilt  $\det F = 1$ ,  $F$  hat den Eigenwert 1 zum Eigenvektor  $e_3$  und  $e_1$  ist ein Eigenvektor von  $F$ .

## 8.2. Ausblick auf weiterführende Forschungsmöglichkeiten

In vielen der betrachteten Bereiche verbleiben offene Fragestellungen, die Gegenstand zukünftiger Forschungsarbeiten werden könnten.

### Verallgemeinerungen der Sum-of-Squared-Logarithms-Ungleichung

Es ist bislang nicht bekannt, ob die SSLI unter schwächeren Bedingungen an die Vektoren  $x, y \in \mathbb{R}_+^n$  gilt. Eine mögliche Erweiterung betrifft den Fall  $e_n(x) < e_n(y)$ , obgleich gezeigt wurde, dass  $e_k(x) \leq e_k(y)$  nicht ausreichend für die Gültigkeit der SSLI ist. Wäre es möglich, dass andere Restriktionen als die Gleichheit  $e_n(x) = e_n(y)$  hinreichend für das Gelten der SSLI sind?

### Konsequenzen aus der Semi-Invertierbarkeit, die nicht notwendig die empirischen Ungleichungen voraussetzen

In Abschnitt 7.2 wurde gezeigt, dass viele positive Eigenschaften von elastischen Energiefunktionen, die in der Literatur häufig aus den empirischen Ungleichungen gefolgert werden, auch von schwächeren konstitutiven Bedingungen impliziert werden, beispielsweise die Semi-Invertierbarkeit bzw. die Bi-Koaxialität der Cauchy-Spannungsantwort. Eine weitere mögliche Fragestellung betrifft die Erweiterbarkeit dieser Ergebnisse: Gibt es leicht zu überprüfende Äquivalenzkriterien für die Bi-Koaxialität? Welche weiteren Konsequenzen der empirischen Ungleichungen folgen aus schwächeren Bedingungen, wie z.B. den schwachen empirischen Ungleichungen?

### Alternative Eigenschaften und Bedingungen für die Konvexität von Energiefunktionen in $C$

Das in Abschnitt 5.1 bewiesene Kriterium für die Minimierungseigenschaft von Lösungen der elastischen Gleichgewichtsgleichung erfordert zusätzlich zur Konvexität von  $C = F^T F$  die positive Semidefinitheit der zur Lösung gehörenden zweiten Piola-Kirchhoff-Spannung. Diese Bedingung ist in der Praxis nur schwer zu überprüfen. Es stellt sich daher die Frage, ob ein praxistaugliches Kriterium für diese Bedingung existiert oder ob die Aussage auch unter zugänglicheren Bedingungen bewiesen werden kann.

### Kriterien für den Zusammenhang zwischen Pure-Shear-Dehnungen und Pure-Shear-Spannungen

Die in Abschnitt 7.4 aufgezeigte hinreichende Bedingung an das konstitutive Gesetz dafür, dass Pure-Shear-Deformationen stets Pure-Shear-Spannungen entsprechen, könnte um notwendige Kriterien ergänzt werden. Desweiteren könnte es von Interesse sein, Kriterien aufzufinden, unter denen Pure-Shear-Spannungen stets Pure-Shear-Deformationen erzeugen.

### Zusammenstellung der Eigenschaften von verschiedenen in der Literatur erwähnten Energiefunktionen

Ziel einer zukünftigen Forschungsarbeit könnte auch die systematische Zusammenstellung eines „Kompendiums der Elastizitätsgesetze“ sein, in dem eine möglichst große Vielfalt konstitutiver Gesetze der nichtlinearen Elastizitätstheorie gesammelt und hinsichtlich ihrer konstitutiven Eigenschaften untersucht wird. Dies würde unter anderem das Auffinden von Gegenbeispielen zu möglichen Implikationen konstitutiver Eigenschaften ermöglichen.

## 9. Literaturverzeichnis

- [1] M. Baker und J.L. Ericksen. „Inequalities restricting the form of the stress-deformation relation for isotropic elastic solids and Reiner-Rivlin fluids“. *J. Washington Acad. Sci.* 44 (1954), S. 33–35.
- [2] J. Ball. „Constitutive inequalities and existence theorems in nonlinear elastostatics“. *Nonlinear Analysis and Mechanics: Heriot-Watt Symposium*. Bd. 1. Pitman Publishing Ltd. Boston. 1977, S. 187–241.
- [3] J. Ball. „Differentiability properties of symmetric and isotropic functions“. *Duke Math. J* 51.3 (1984), S. 699–728.
- [4] R. C. Batra. „Deformation produced by a simple tensile load in an isotropic elastic body“. *Journal of Elasticity* 6.1 (1976), S. 109–111.
- [5] G. F. Becker. „The Finite Elastic Stress-Strain Function“. *American Journal of Science* 46 (1893). newly typeset version available at [https://www.uni-due.de/imperia/md/content/mathematik/ag\\_neff/becker\\_latex\\_new1893.pdf](https://www.uni-due.de/imperia/md/content/mathematik/ag_neff/becker_latex_new1893.pdf), S. 337–356.
- [6] D. S. Bernstein. „Inequalities for the trace of matrix exponentials“. *SIAM Journal on Matrix Analysis and Applications* 9.2 (1988), S. 156–158.
- [7] D. S. Bernstein. *Matrix mathematics: theory, facts, and formulas*. Princeton University Press, 2009.
- [8] P. Biler und A. Witkowski. *Problems in mathematical analysis*. M. Dekker, 1990.
- [9] M. Birsan, P. Neff und J. Lankeit. „Sum of squared logarithms – an inequality relating positive definite matrices and their matrix logarithm“. *Journal of Inequalities and Applications* 2013.168 (2013). open access, S. 1–16.
- [10] J. A. Blume. „Elastic materials with coincident principal stress and strain axes“. *Journal of Elasticity* 35.1-3 (1994), S. 275–280.
- [11] C. Boor. „A naive proof of the representation theorem for isotropic, linear asymmetric stress-strain relations“. *Journal of elasticity* 15.2 (1985), S. 225–227.
- [12] L. Borisov, P. Neff, S. Sra und C. Thiel. „The sum of squared logarithms inequality in arbitrary dimensions“. *to appear in Linear Algebra and its Applications* (2016). preprint available at arXiv:1508.04039. DOI: [10.1016/j.laa.2016.06.026](https://doi.org/10.1016/j.laa.2016.06.026).
- [13] S. Bosch. *Lineare Algebra*. Springer-Verlag, 2014.
- [14] G. Cardano. *Ars Magna*. Petreius, 1545.
- [15] P. G. Ciarlet. *Three-Dimensional Elasticity*. Studies in mathematics and its applications 1. Elsevier Science, 1988. ISBN: 9780080875415.
- [16] P. G. Ciarlet. „Mathematical Elasticity, Volume I“. *Elsevier* (1988).
- [17] P. G. Ciarlet und G. Geymonat. „Sur les lois de comportement en élasticité non linéaire compressible“. *CR Acad. Sci. Paris Sér. II* 295 (1982), S. 423–426.
- [18] P. G. Ciarlet und C. Mardare. „The pure displacement problem in nonlinear three-dimensional elasticity: intrinsic formulation and existence theorems“. *Comptes Rendus Mathématique* 347.11 (2009), S. 677–683.
- [19] T. Cover und J. Thomas. *Elements of information theory*. Wiley, 2012. ISBN: 9781118585771.
- [20] F. Cucker und A. G. Corbalan. „An alternate proof of the continuity of the roots of a polynomial“. *American Mathematical Monthly* (1989), S. 342–345.



## 9. Literaturverzeichnis

- [21] F. Dannan, P. Neff und C. Thiel. „On the sum of squared logarithms inequality and related inequalities“. *Journal of Mathematical Inequalities* (2015). open access, available at arXiv:1411.1290.
- [22] C. Davis. „All convex invariant functions of hermitian matrices“. *Archiv der Mathematik* 8.4 (1957), S. 276–278. DOI: [10.1007/BF01898787](https://doi.org/10.1007/BF01898787).
- [23] M. Destrade, J.G. Murphy und G. Saccomandi. „Simple shear is not so simple“. *International Journal of Non-Linear Mechanics* 47.2 (2012), S. 210–214.
- [24] L. Dorfmann und R.W. Ogden. *Nonlinear Mechanics of Soft Fibrous Materials*. Springer, 2014.
- [25] J.E. Dunn. „Elastic materials of coaxial type and inequalities sufficient to ensure strong ellipticity“. *International Journal of Solids and Structures* 20.5 (1984), S. 417–427.
- [26] L.C. Evans. *Partial Differential Equations: Second Edition*. Bd. 19. Graduate Studies in Mathematics. American Mathematical Society, 2010. ISBN: 9780821849743.
- [27] J. Finger. „Das Potential der inneren Kräfte und die Beziehungen zwischen den Deformationen und den Spannungen in elastisch isotropen Körpern bei Berücksichtigung von Gliedern, die bezüglich der Deformationselemente von dritter, beziehungsweise zweiter Ordnung sind“. *Sitzungsberichte der Akademie der Wissenschaften in Wien* 44 (1894). ISSN: 0378-8644.
- [28] D.Y. Gao und R.W. Ogden. „Closed-form solutions, extremality and nonsmoothness criteria in a large deformation elasticity problem“. *Zeitschrift für angewandte Mathematik und Physik* 59.3 (2008), S. 498–517.
- [29] D.Y. Gao. „Global extremum criteria for finite elasticity“. *Zeitschrift für angewandte Mathematik und Physik* 43.5 (1992), S. 924–937.
- [30] D.Y. Gao und E. Hajilarov. „Analytic solutions to three-dimensional finite deformation problems governed by St Venant–Kirchhoff material“. *Mathematics and Mechanics of Solids* (2015). DOI: 10.1177/1081286515591084.
- [31] D.Y. Gao, P. Neff, I. Roventa und C. Thiel. „On the convexity of nonlinear elastic energies in the right Cauchy–Green tensor“. *to appear in Journal of Elasticity* (2016). preprint available at arXiv:1508.05721.
- [32] G. Green. „On the propagation of light in crystallized media“. *Transactions of the Cambridge Philosophical Society* 7 (1841), S. 121.
- [33] A. Gupta, D.J. Steigmann und J.S. Stölken. „On the evolution of plasticity and incompatibility“. *Mathematics and Mechanics of Solids* 12.6 (2007), S. 583–610.
- [34] G.H. Hardy, J.E. Littlewood und G. Pólya. *Inequalities*. Cambridge University Press, 1952.
- [35] S. Hartmann und P. Neff. „Polyconvexity of generalized polynomial-type hyperelastic strain energy functions for near-incompressibility“. *International journal of solids and structures* 40.11 (2003), S. 2767–2791.
- [36] D.L. Henann und L. Anand. „A large strain isotropic elasticity model based on molecular dynamics simulations of a metallic glass“. *Journal of Elasticity* 104.1-2 (2011), S. 281–302.
- [37] H. Hencky. „Welche Umstände bedingen die Verfestigung bei der bildsamen Verformung von festen isotropen Körpern?“ *Zeitschrift für Physik* 55 (1929). available at [www.uni-due.de/imperia/md/content/mathematik/ag\\_neff/hencky1929.pdf](http://www.uni-due.de/imperia/md/content/mathematik/ag_neff/hencky1929.pdf), S. 145–155.
- [38] H. Heuser. „Lehrbuch der Analysis, Teil 2“. *BG Teubner, Stuttgart* (1991).
- [39] B.E. Johnson und A. Hoger. „The dependence of the elasticity tensor on residual stress“. *Journal of Elasticity* 33.2 (1993), S. 145–165. DOI: [10.1007/BF00705803](https://doi.org/10.1007/BF00705803).
- [40] D.F. Jones und L.R.G. Treloar. „The properties of rubber in pure homogeneous strain“. *Journal of Physics D: Applied Physics* 8.11 (1975), S. 1285.
- [41] R. Jozsa und G. Mitchison. „Symmetric polynomials in information theory: Entropy and subentropy“. *Journal of Mathematical Physics* 56.6 (2015), S. 062201.

## 9. Literaturverzeichnis

- [42] H. Le Dret und A. Raoult. „Quasiconvex Envelopes of Stored Energy Densities that are Convex with Respect to the Strain Tensor“. *Calculus of Variations, Applications and Computations*. Hrsg. von C. Bandle, M. Chipot, J. S. J. Paulin, J. Bemelmans und I. Shafrir. Bd. 326. CRC Press, 1995, S. 138–146.
- [43] H. Le Dret und A. Raoult. „The quasiconvex envelope of the Saint Venant–Kirchhoff stored energy function“. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* 125.06 (1995), S. 1179–1192.
- [44] S. Lehmich, P. Neff und J. Lankeit. „On the convexity of the function  $C \mapsto f(\det C)$  on positive-definite matrices“. *Mathematics and Mechanics of Solids* 19.4 (2014), S. 369–375.
- [45] A. Marshall, I. Olkin und B. Arnold. *Inequalities: Theory of majorization and its applications: Theory of majorization and its applications*. Springer Series in Statistics. Springer, 2010. ISBN: 9780387682761.
- [46] R. J. Martin. „Logarithmic strain and its role in nonlinear elasticity“. Dissertation. Universität Duisburg-Essen, 2016.
- [47] R. J. Martin und P. Neff. „Some remarks on the monotonicity of primary matrix functions on the set of symmetric matrices“. *Archive of Applied Mechanics* 85.12 (2015), S. 1761–1778.
- [48] S. Marzano. „An interpretation of Baker-Ericksen inequalities in uniaxial deformation and stress“. *Meccanica* 18.4 (1983), S. 233–235.
- [49] G. Mitchison und R. Jozsa. „Towards a geometrical interpretation of quantum-information compression“. *Physical Review A* 69.3 (2004), S. 032304.
- [50] D. S. Mitrinović, J. E. Pečarić und A. M. Fink. *Classical and new inequalities in analysis*. Springer, 1993.
- [51] P. Neff, B. Eidel, F. Osterbrink und R. J. Martin. „A Riemannian approach to strain measures in nonlinear elasticity“. *Comptes Rendus Mécanique* 342.4 (2014), S. 254–257.
- [52] P. Neff, J. Lankeit und A. Madeo. „On Grioli’s minimum property and its relation to Cauchy’s polar decomposition“. *International Journal of Engineering Science* 80 (2014), S. 209–217. ISSN: 0020-7225.
- [53] P. Neff, I. Münch und R. Martin. „Rediscovering G.F. Becker’s early axiomatic deduction of a multiaxial nonlinear stress-strain relation based on logarithmic strain“. *to appear in Mathematics and Mechanics of Solids*, doi: 10.1177/1081286514542296 (2014). available at arXiv:1403.4675. DOI: [10.1177/1081286514542296](https://doi.org/10.1177/1081286514542296).
- [54] P. Neff. „On Korn’s first inequality with non-constant coefficients“. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* 132.01 (2002), S. 221–243.
- [55] P. Neff. *The sum of squared logarithms conjecture*. [www.mathoverflow.net/questions/207845/the-sum-of-squared-logarithms-conjecture](http://www.mathoverflow.net/questions/207845/the-sum-of-squared-logarithms-conjecture). Mai 2015.
- [56] P. Neff, B. Eidel und R. J. Martin. „Geometry of logarithmic strain measures in solid mechanics“. *Archive for Rational Mechanics and Analysis* 222.2 (2016). available at arXiv:1505.02203, S. 507–572. DOI: [10.1007/s00205-016-1007-x](https://doi.org/10.1007/s00205-016-1007-x).
- [57] P. Neff, I.-D. Ghiba und J. Lankeit. „The exponentiated Hencky-logarithmic strain energy. Part I: constitutive issues and rank-one convexity“. *Journal of Elasticity* 121.2 (2015). available at arXiv:1403.3843, S. 143–234.
- [58] P. Neff, J. Lankeit, I. D. Ghiba, R. J. Martin und D. J. Steigmann. „The exponentiated Hencky-logarithmic strain energy. Part II: Coercivity, planar polyconvexity and existence of minimizers“. *Zeitschrift für angewandte Mathematik und Physik* 66.4 (2015). available at arXiv:1408.4430, S. 1671–1693.
- [59] P. Neff, Y. Nakatsukasa und A. Fischle. „A logarithmic minimization property of the unitary polar factor in the spectral and Frobenius norms“. *SIAM Journal on Matrix Analysis and Applications* 35.3 (2014), S. 1132–1154.

## 9. Literaturverzeichnis

- [60] P. Neff, Y. Nakatsukasa und A. Fischle. „A logarithmic minimization property of the unitary polar factor in the spectral norm and the Frobenius matrix norm“. *SIAM Journal on Matrix Analysis and Applications* 35.3 (2014), S. 1132–1154.
- [61] R. W. Ogden, G. Saccomandi und I. Sgura. „Fitting hyperelastic models to experimental data“. *Computational Mechanics* 34.6 (2004), S. 484–502.
- [62] R. W. Ogden. *Non-linear elastic deformations*. Courier Corporation, 1997.
- [63] A. C. Pipkin. „Relaxed energy densities for large deformations of membranes“. *IMA Journal of Applied Mathematics* 52.3 (1994), S. 297–308.
- [64] W. Pompe und P. Neff. „On the generalized sum of squared logarithms inequality“. *Journal of Inequalities and Applications* (2015). open access, available at arXiv:1410.2706.
- [65] A. Raoult. „Non-polyconvexity of the stored energy function of a Saint Venant–Kirchhoff material“. *Aplikace matematiky* 31.6 (1986), S. 417–419.
- [66] H. Richter. „Das isotrope Elastizitätsgesetz“. *Zeitschrift für Angewandte Mathematik und Mechanik* 28.7/8 (1948). available at [https://www.uni-due.de/imperia/md/content/mathematik/ag\\_neff/richter\\_isotrop\\_log.pdf](https://www.uni-due.de/imperia/md/content/mathematik/ag_neff/richter_isotrop_log.pdf), S. 205–209.
- [67] H. Richter. „Zum Logarithmus einer Matrix“. *Archiv der Mathematik* 2.5 (1949). available at [https://www.uni-due.de/imperia/md/content/mathematik/ag\\_neff/richter\\_log.pdf](https://www.uni-due.de/imperia/md/content/mathematik/ag_neff/richter_log.pdf), S. 360–363. ISSN: 0003-889X. DOI: [10.1007/BF02036865](https://doi.org/10.1007/BF02036865).
- [68] H. Richter. „Über Matrixfunktionen“. *Mathematische Annalen* 122.1 (1950), S. 16–34. ISSN: 0025-5831. DOI: [10.1007/BF01342947](https://doi.org/10.1007/BF01342947).
- [69] R. S. Rivlin und J. L. Ericksen. „Stress-Deformation Relations for Isotropic Materials“. *Collected Papers of R. S. Rivlin*. Hrsg. von G. I. Barenblatt und D. D. Joseph. Springer New York, 1997, S. 911–1013. DOI: [10.1007/978-1-4612-2416-7\\_61](https://doi.org/10.1007/978-1-4612-2416-7_61). URL: [http://dx.doi.org/10.1007/978-1-4612-2416-7\\_61](http://dx.doi.org/10.1007/978-1-4612-2416-7_61).
- [70] R. Rivlin und D. Saunders. „Large elastic deformations of isotropic materials“. *Collected Papers of RS Rivlin*. Springer, 1997, S. 157–194.
- [71] J. Schröder und P. Neff. *Poly-, quasi- and rank-one convexity in applied mechanics*. Bd. 516. Springer Science & Business Media, 2010.
- [72] B. R. Seth. *Generalized strain measure with applications to physical problems*. Techn. Ber. Defense Technical Information Center, 1961.
- [73] M. Šilhavý. „The Convexity of  $C \mapsto h(\det C)$ “. *Technische Mechanik* 35.1 (2015), S. 60–61.
- [74] M. Šilhavý. „A functional inequality related to analytic continuation“. *preprint Institute of Mathematics AS CR IM-2015-37* (2015). available at [www.math.cas.cz/fichier/preprints/IM\\_20150623102729\\_44.pdf](http://www.math.cas.cz/fichier/preprints/IM_20150623102729_44.pdf).
- [75] S. J. Spector. „A Note on the Convexity of  $C \mapsto h(\det C)$ “. *Journal of Elasticity* 118.2 (2015), S. 251–256. DOI: [10.1007/s10659-014-9486-1](https://doi.org/10.1007/s10659-014-9486-1).
- [76] M. Tomic. „Théoreme de Gauss relatif au centre de gravité et son application“. *Bull. Soc. Math. Phys. Serbie* 1 (1949), S. 31–40.
- [77] L. R. G. Treloar. „The elasticity and related properties of rubbers“. *Reports on progress in physics* 36.7 (1973), S. 755.
- [78] C. Truesdell und H. Moon. „Inequalities sufficient to ensure semi-invertibility of isotropic functions“. *Journal of Elasticity* 5.34 (1975).
- [79] C. Truesdell und W. Noll. *The Non-Linear Field Theories of Mechanics*. originally published as Volume III/3 of the Encyclopedia of Physics in 1965. Springer, 2004.
- [80] C. Vallée. „Lois de comportement élastique isotropes en grandes déformations“. *International Journal of Engineering Science* 16.7 (1978), S. 451–457.
- [81] H. Weyl. „Inequalities between the two kinds of eigenvalues of a linear transformation“. *Proceedings of the National Academy of Sciences of the United States of America* 35.7 (1949), S. 408.

# 10. Stichwortverzeichnis

- additiv isochor-volumetrischer Split, 146
- Baker-Ericksen-Ungleichungen, 40, 159
  - strikte, 40, 43
  - strikte, von isotropen Elastizitätsgesetzen, 135
- Beschleunigungsfeld, 101
- Bi-Koaxialität, 43
  - von isotropen Elastizitätsgesetzen, 135
- Ciarlet-Geymonat-Energie, 151
- Deformation
  - equibiaxial, 115
  - finite, Simple-Shear, 156
  - Pure-Shear, 116
  - Simple-Shear, 153, 159
  - uniaxial, 114
- Deformationsabbildung, 95
- Dehnung, 104, 122, 155, 169
  - triaxiale, 159
- Dehnungs-Spannungs-Zusammenhang, 116
- Determinante, 4
- Differenzierbarkeit, 8
  - in den Invarianten, 66
- Diskriminante eines Polynoms, 65
- Divergenzfreiheit von  $S_1$ , 103
- Elastizitätsgesetz, *siehe* konstitutives Gesetz
  - lineares, 123
- elementarsymmetrische Polynome, 28, 33, 62, 67–69, 74, 80, 82, 120, 139
  - Eigenschaften, 69–70
- empirische Ungleichungen, 40, 42
  - schwache, 41, 42
  - schwache, von isotropen Elastizitätsgesetzen, 135
  - von isotropen Elastizitätsgesetzen, 135
- Energiefunktional, 125
- Energiefunktionen vom Hencky-Typ, 141
- Entropie, 82
- Euler-Lagrange-Gleichung, 54, 102
  - schwache Form, 54
  - schwache Lösung, 125
- Flächenkraftabbildung, 95, 101
- Frobenius-Skalarprodukt, 18
- Frobeniusnorm, 22
- Fundamentallemma der Variationsrechnung, 52
- Gaußscher Integralsatz, 11
- gewichtete Frobeniusnorm, 23
- gewichtetes isotropes Skalarprodukt, 20
- Gibbs-Ungleichung, 92
- Gleichgewicht der Kräfte, 96
- Hauptachsen, 24
- Hauptspannungen, 111, 113, 142, 151, 162, 172
- Hencky-Energie
  - exponentiated, 145
  - inkompressible, quadratische, 117
  - quadratische, 39, 144, 174
- Heron-Verfahren, 48
- Homogenität einer Deformation, 103
- Hyperelastizität, 106
- Invarianten einer Matrix, 28, 33, 34, 38–42, 66–69, 72, 120, 131, 134–152
- Invertierbarkeit, 43
  - von isotropen Elastizitätsgesetzen, 136
- Isotrope Mengen, 27
- Isotropie, 27–38
  - von Elastizitätsgesetzen, 110
  - von Energiefunktionen, 112
- Jacobi-Matrix, 8
- Jensen-Ungleichung, 91
- Koaxialität, 27–35
- Kofaktor, 14, 99
- konstitutives Gesetz, 104
- konvexe Funktion, 57
- konvexe Menge, 57
- Konvexität
  - von Energiefunktionen, 124–133
- Kreuzprodukt, 15
- Lebesgue-Raum, 52
- Lokalität einer Deformation, 103
- MacLaurin-Ungleichung, 68
- Majorization, 59, 75
  - schwache, 59
  - schwache logarithmische, 61

## STICHWORTVERZEICHNIS

- Matrixexponential, 44
- Matrixlogarithmus, 46
- Mooney-Rivlin-Energie
  - kompressible, 148
- Neo-Hooke-Energie
  - kompressible, 149
- Newton-Ungleichung, 68
- Norm, 7
- Normalenfeld, äußeres, 10
- Normalenvektor
  - äußerer, 11, 95–97, 102, 105, 107
- Objektivität
  - von Elastizitätsgesetzen, 107
  - von Energiefunktionen, 108
- Piola-Identität, 99
- Polarzerlegung, 26
  - eines Deformationsgradienten, 103
- Polynom
  - Koeffizienten, 63
  - Nullstellen, 63
- primäre Matrixfunktion, 44
- Rotation, 56, 96, 104, 122, 153, 155, 156, 159, 163, 168, 169
  - lineare, 122
- Saint-Venant-Kirchhoff-Energie, 116, 129, 150
  - isochor, 149
- Satz von Cauchy, 98
- Scherung, *siehe* shear
- Schur-Konvexität, 62, 69, 85, 86
- schwache Differenzierbarkeit, 52
- Semi-Invertierbarkeit, 39, 43, 134, 152, 159, 160, 163
  - von isotropen Elastizitätsgesetzen, 136
- Shear
  - Szenario, 153
- shear, 153–183
- Simple-Shear
  - finit, 158–169
  - linear, 157–158
- Skalarprodukt, 6
- Sobolev-Raum, 53
- Spannungsantwort, 134
- spannungsfreie Referenzkonfiguration, 104, 128, 134, 163, 173
- Spannungsprinzip von Euler und Cauchy, 97
- Spannungstensor
  - Biot-Spannungstensor, 106
  - Cauchy-Spannungstensor, 105
  - Erster Piola-Kirchhoff-Spannungstensor, 106
  - Pure Shear, 159
  - zweiter Piola-Kirchhoff-Spannungstensor, 106
- SSLI, *siehe* Sum-of-Squared-Logarithms-Ungleichung
- Sum-of-Powered-Logarithms-Ungleichung, 85
- Sum-of-Squared-Logarithms-Ungleichung, 71–86, 135, 140
  - Matrixform, 71
- symmetrische Funktion, 63
- symmetrische Menge, 63
- Tangentialraum, 25
- Tschebyscheff-Summenungleichung, 49
- Umordnungsungleichung, 48
- Valanis-Landel-Energie, 117, 127
  - verallgemeinerte Form, 173, 178
- Verzerrung, 121
- Verzerrungstensor
  - linearer, 121
- Volumenkraftabbildung, 95, 101
- Wärmeleitungsgleichung, 11