

Fakultät für Mathematik – Campus **Essen**
Lehrstuhl für *Nichtlineare Analysis und Modellierung*
Universitätsstraße 2
45141 Essen

Diplomarbeit

Über die Konvexität der Funktion

$$C \longmapsto f(\det C)$$

auf $\mathbb{P}\text{Sym}(3)$

Cand.-Math. Stephan Lehmich

23. September 2011

Betreuer

PROF. DR. PATRIZIO NEFF

Inhaltsverzeichnis

1. Einleitung	5
2. Variationsrechnung	7
2.1. Einführung	7
2.2. Klassische Methoden	9
2.2.1. Euler-Lagrange-Gleichung	9
2.3. Direkte Methoden	11
2.3.1. Grundlagen	11
2.3.2. Ein Existenzsatz im skalaren Fall	17
2.3.3. Der vektorielle Fall	18
3. Konvexität von $f \circ \det$ auf $\mathbb{P}\text{Sym}(3)$	23
3.1. Eine Differentialungleichung an die Funktion f	23
3.1.1. Vorbereitende Lemmata und Definitionen	23
3.1.2. Differenzierbarkeit in normierten Räumen	27
3.1.3. Eine Konvexitätsbedingung	32
3.1.4. Die Differentialungleichung	38
3.2. Lösungen der Differentialungleichung(en)	45
3.2.1. Lineare Differentialgleichungen und der Grenzfall	45
3.2.2. Ober- und Unterfunktionen	47
4. Nichtlineare Elastizität	53
4.1. Das NEO-HOOKE-Modell	54
Literaturverzeichnis	57
A. Ein Gegenbeispiel	59

1. Einleitung

In modernen Entwicklungsprozessen kann kaum noch auf (numerische) Simulationsverfahren verzichtet werden. Die Anforderungen an Bauteile und Materialien werden stets größer.

Auch ist die Leistungsfähigkeit heutiger Computer hinreichend groß um viele Probleme simulieren zu können. Die Vorteile liegen dabei auf der Hand:

Im Entwicklungsprozess neuer Bauteile kann die Machbarkeit der Komponenten bereits in einem frühen Stadium überprüft werden. Dies bringt zum Einen den Vorteil mit sich, dass Entwicklungsprozesse beschleunigt und somit höchstwahrscheinlich Kosten gespart werden können. Zum Anderen kann die Produktqualität erheblich gesteigert werden. Auch können gewisse Größen mittels Experimenten überhaupt nicht bis lediglich unzureichend beobachtet werden, sodass Simulationen gar unabdingbar sind. Schlussendlich fehlt oftmals auch die Zeit, um das Langzeitverhalten von Bauteilen experimentell zu untersuchen.

Natürlich können Simulationen die Bauteilprüfungen nicht vollständig ersetzen, da Modelle eben immer nur ein Abbild der Realität darstellen und somit niemals exakt sind. Die Anzahl der erforderlichen Prüfungen lässt sich aber deutlich reduzieren. Neben der Simulationsmethode (in der Regel mittels der „Methode der finiten Elemente“ (FEM)), welche jedoch nicht Gegenstand dieser Arbeit sein soll, kommt der Modellbildung an sich eine entscheidende Rolle für die Güte der Vorhersage zu.

Für die Modellbildung ist eine korrekte Beschreibung des Werkstoffverhaltens vonnöten. Darunter versteht man den Zusammenhang zwischen einer äußeren aufgetragenen Kraft (bzw. mechanischen Spannung oder allgemein: Belastung) und der Reaktion des Werkstoffes hierauf, z. B. in Form einer Längenänderung (Dehnung). Die Bestimmung dieses Zusammenhangs geschieht in den Ingenieurwissenschaften beispielsweise durch Zugversuche an einem einfachen Probekörper. Solange die Belastung einen bestimmten Wert nicht übersteigt (Streckgrenze), liefern z. B. metallische Werkstoffe einen annähernd linearen Zusammenhang. In der Elastizität greifen lineare oder linearisierte Modelle jedoch zu kurz. Nicht nur, dass der Zusammenhang zwischen Spannung und Dehnung bei Elastomeren nichtlinear ist, so sind diese auch oftmals starken Verformungen mit Dehnungen über 100% ausgesetzt. Man arbeitet also besser mit nichtlinearen Werkstoffmodellen. Eine große Klasse solcher Modelle bilden die sogenannten „hyperelastischen Werkstoffmodelle“.

Eines solcher hyperelastischen Werkstoffmodelle ist das physikalisch motivierte (d.h. es werden grundsätzlich physikalisch sinnvolle Ergebnisse vorhergesagt) Neo-HOOKE-Modell (da ähnlich dem HOOKEschen Gesetz). Es eignet sich zur Modellierung gummiartigen Materials, wie z. B. Airbags, Reifen oder gar Arterienwänden. Die freie Energiefunktion W eines solchen Materials (welches eine spezielle Klasse von Materialien des OGDEN-Typs

darstellt) kann z. B. durch

$$W(F) = a\|F\|^2 + f(\det F), \quad F = \nabla\varphi, \quad a \in \mathbb{R}_+ \quad (1.1)$$

beschrieben werden, siehe [12, S. 6]. φ beschreibt hier die Deformation einer „Referenzkonfiguration“. Dies führt auf ein Randwertproblem, welches man mit den direkten Methoden der Variationsrechnung behandelt. Man sucht also eine minimierende Deformation der freien Energiefunktion W . Eine der Bedingungen, die man an W für den Nachweis der Existenz von Lösungen des Randwertproblems stellen „muss“, ist die Polykonvexität. Da aber für Funktionen $W : \mathbb{R}^{N \times n} \rightarrow \mathbb{R}$, wobei $n, N \in \mathbb{N}$, stets

$$W \text{ konvex} \Rightarrow W \text{ polykonvex} \Rightarrow W \text{ quasikonvex} \Rightarrow W \text{ rank-1-konvex}$$

gilt [6, S. 59], $F \mapsto a\|F\|^2$ in Hilberträumen (also insbesondere auf $\mathbb{R}^{N \times n}$) sogar strikt konvex ist und sich Konvexität per Definition mit der Addition verträgt, motiviert dies die zugrunde liegende Fragestellung dieser Arbeit: Welcher Gestalt muss die Funktion f in (1.1) sein, damit auch $f \circ \det$ konvex ist?[†] Es genügt nicht, wie man zunächst vermuten könnte, dass die Funktion f selbst „nur“ konvex sein muss. Ein einfaches Gegenbeispiel für die sogar strikt konvexe Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^2$ liefert die Wahl

$$F := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} := \text{diag}(1, 1, 3) \quad , \quad G := \text{diag}(3, 1, 1) \quad , \quad \tau = \frac{3}{4} \quad ,$$

denn damit

$$\begin{aligned} f \circ \det((1 - \tau)F + \tau G) &\stackrel{!}{\leq} (1 - \tau) \cdot (f \circ \det(F)) + \tau \cdot (f \circ \det(G)) \\ \Leftrightarrow f \left(\det \left(\text{diag} \left(\frac{5}{2}, 1, \frac{3}{2} \right) \right) \right) &\leq \frac{1}{4} \cdot f(3) + \frac{3}{4} \cdot f(3) \\ \Leftrightarrow \left(\frac{15}{4} \right)^2 &\leq \frac{1}{4} \cdot 9 + \frac{3}{4} \cdot 9 = 9 \\ \Leftrightarrow 225 &\leq 144 \quad \text{!} \end{aligned}$$

wird die Konvexitätsungleichung verletzt.

Diese mathematisch recht vage formulierten Ausführungen wollen wir in der Arbeit fachlich unterfüttern. Zunächst folgt im nächsten Kapitel eine kurze Darstellung der Variationsrechnung mit Hauptaugenmerk auf den direkten Methoden. Die Notation ist natürlich bereits auf die Hyperelastizität ausgelegt. Mit der Materie vertraute Leser können direkt zum nächsten Kapitel übergehen, in welchem die eben angesprochene Problematik der Konvexität von $f \circ \det$ behandelt wird. Wir werden eine äquivalente Formulierung für zweimal stetig differenzierbare Funktionen f in Form einer Differentialungleichung herleiten. Das Kapitel ist in sich geschlossen und enthält insbesondere alle nötigen Definitionen, sodass der Leser nicht gezwungen ist, im vorherigen Kapitel nachzuschlagen. Im letzten Kapitel kommen wir auf die Hyperelastizität und das Neo-HOOKE-Modell zurück.

[†] Dies allerdings nicht als Funktion des Deformationsgradienten F , sondern nach einer Umformulierung als Funktion des sogenannten „rechten CAUCHY-GREEN-Deformationstensors“ $C := F^T F$.

2. Variationsrechnung

Die Ausführungen dieses Kapitels orientieren sich größtenteils an den Büchern von EVANS [8] und DACOROGNA [6], und/oder stellen klassische Resultate dar. Wenn auf andere Literatur zurückgegriffen wird, ist dies an den entsprechenden Stellen vermerkt.

2.1. Einführung

Die Variationsrechnung ist diejenige mathematische Theorie, welche sich mit der Lösung des folgenden Grundproblems befasst:

Betrachte das Funktional

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

wobei

- $n, N \in \mathbb{N}$, $0 \notin \mathbb{N}$ und
- $x = (x_1, \dots, x_n) \in \Omega$ mit $\Omega \subseteq \mathbb{R}^n$ eine offene Menge sei.
- Des Weiteren sei

$$\varphi = (\varphi^1, \dots, \varphi^N) : \Omega \longrightarrow \mathbb{R}^N$$

und somit der Gradient von φ an der Stelle x

$$\nabla\varphi(x) = (\partial_i\varphi^j(x))_{\substack{1 \leq j \leq N \\ 1 \leq i \leq n}} = \begin{pmatrix} \frac{\partial\varphi^1}{\partial x_1}(x) & \cdots & \frac{\partial\varphi^1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial\varphi^N}{\partial x_1}(x) & \cdots & \frac{\partial\varphi^N}{\partial x_n}(x) \end{pmatrix} \in \mathbb{R}^{N \times n} , \text{ sowie}$$

- $W : \Omega \times \mathbb{R}^N \times \mathbb{R}^{N \times n} \longrightarrow \mathbb{R}$, $W = W(x, \varphi, \xi)$ eine vorgegebene Funktion.¹
- Sprechweise: Falls $n = 1$ oder $N = 1$, so wollen wir das Problem als „skalar“, andernfalls als „vektoriell“ bezeichnen.

¹ Unter physikalischen Gesichtspunkten ist es manchmal sinnvoll, für W auch den Wert $+\infty$ zuzulassen. Falls die freie Energiefunktion also Werte in $\mathbb{R}^\infty := \mathbb{R} \cup \{+\infty\}$ annimmt, werden wir an entsprechender Stelle darauf hinweisen.

Unter gewissen Nebenbedingungen, welche wir später anführen werden, sucht man nun im Raum der Funktionen φ mit $\varphi : \Omega \rightarrow \mathbb{R}^N$ eine Funktion $\bar{\varphi}$ welche I minimiert. Bezeichnet man die Menge aller „zulässigen“ (engl.: admissible) Funktionen φ mit \mathcal{A} , so kann man auch kurz formulieren:

Gesucht ist $\bar{\varphi} \in \mathcal{A}$ mit

$$I(\bar{\varphi}) = \min_{\varphi \in \mathcal{A}} I(\varphi).$$

Man nennt $\bar{\varphi}$ Lösung des Variationsproblems.

Bevor wir erläutern wie man dieses Problem mathematisch handhabt, ein paar kurze Worte zu den Ursprüngen der Variationsrechnung und damit einhergehend der Motivation sich überhaupt mit Problemen dieser Art zu beschäftigen:

Schon um 200 vor Christus untersuchte ZENODOR das Problem der Isoperimetrischen Ungleichung, dessen mathematische Formulierung auf die eben von uns eingeführte variationelle Darstellung führt. 1662 entdeckte FERMAT, dass der Weg des Lichtes zwischen zwei Punkten derjenige ist, den es in der kürzesten Zeit zurücklegt. Auch dieses Problem der Brechung von Lichtstrahlen tritt in variationeller Form auf, wobei ($n = N = 1$)

$$W(x, \varphi, \xi) := g(x, \varphi) \sqrt{1 + \xi^2}.$$

Hier ist $g(\cdot, \cdot)$ der vom Ort abhängige Brechungsindex des Mediums, das der Lichtstrahl durchquert. Wir werden darauf später noch kurz zurück kommen.

Berühmt ist ebenfalls BERNOULLI's Brachistochronenproblem aus dem Jahre 1696, welches BALDER [4, S. 10] folgendermaßen zusammenfasst:

“Johann Bernoulli's famous 1696 brachistochrone problem asks for the optimal shape of a metal wire that connects two fixed points A and B in space. A bead of unit mass falls along this wire, without friction, under the sole influence of gravity. The shape of the wire is defined to be optimal if the bead falls from A to B in as short a time as possible.”

Mit der Gravitationskonstanten g (und erneut $n = N = 1$) ergibt sich für dieses Problem

$$W(x, \varphi, \xi) = W(\varphi, \xi) := \frac{\sqrt{1 + \xi^2}}{\sqrt{2g\varphi}}.$$

Auf LEONHARD EULER (1707-1783) geht schließlich der Name „Variationsrechnung“ zurück. Anhand dieser kurzen Auflistung lässt sich erahnen, dass viele Gesetze der Physik und anderer Wissenschaften direkt zu variationellen Formulierungen führen, bzw. so formuliert werden können. Und so ist die Variationsrechnung auch aus den Anwendungen heraus entstanden. Sie ist eng mit der Theorie der partiellen Differentialgleichungen verknüpft. Denn sei $\bar{\varphi}$ eine Lösung des Variationsproblems, dann gilt (womöglich) in symbolischer Schreibweise

$$DI(\bar{\varphi}) = 0, \tag{2.1}$$

welches i. Allg. eine nichtlineare partielle Differentialgleichung (kurz: PDGL) repräsentiert. Wohingegen es für gewöhnlich sehr schwer ist (2.1) direkt zu lösen, ist es unter Umständen wesentlich einfacher Minima (oder allgemein Extrema) von $I(\cdot)$ zu bestimmen.

2.2. Klassische Methoden

Die klassischen Methoden greifen den letzten Gedanken auf. Es wird versucht bekannte Methoden zur Bestimmung von Extrema von Funktionen, die auf Teilmengen des \mathbb{R}^N erklärt sind, auf variationelle Formulierungen anzuwenden. Schwierigkeiten bereitet auf den ersten Blick die Unendlichdimensionalität des Definitionsbereiches des Funktionals I . Durch geeignete Annahmen können diese aber beherrscht werden. Dies wollen wir im nächsten Unterabschnitt ausführen, um die sogenannten EULER-LAGRANGE-Gleichungen herzuleiten.

2.2.1. Euler-Lagrange-Gleichung

Mit obigen Bezeichnungen seien fortan Ω beschränkt mit glattem Rand $\partial\Omega$ und

$$W : \bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

ebenfalls glatt (C^2 genügt). Es gilt also $N = 1$, wobei dies nicht obligatorisch ist. Es erspart lediglich notationelle Klümmzüge. Mit W_{x_i} (respektive W_φ und W_{ξ_i}) sei die partielle Ableitung von $W = W(x, \varphi, \xi)$ nach x_i (respektive φ und ξ_i) bezeichnet. Außerdem wollen wir nun eine der anfangs erwähnten Nebenbedingungen an die Menge \mathcal{A} unserer zulässigen Funktionen stellen: die Randbedingung $\varphi = g$ auf $\partial\Omega$, wobei $g \in C^2(\partial\Omega)$. Nun können wir folgenden Satz formulieren:

Satz 2.1 (EULER-LAGRANGE-GLEICHUNG)

Unter obigen Voraussetzungen existiere ein glatter Minimierer

$$\bar{\varphi} \in \mathcal{A} := \{\varphi \in C^2(\bar{\Omega}) : \varphi = g \text{ auf } \partial\Omega\}$$

von $I(\cdot)$, also eine Lösung des Variationsproblems.

Dann löst $\bar{\varphi}$ die EULER-LAGRANGE-Gleichung (ELG). Dies ist das Randwertproblem

$$\sum_{i=1}^n (W_{\xi_i}(x, \bar{\varphi}(x), \nabla \bar{\varphi}(x)))_{x_i} = W_\varphi(x, \bar{\varphi}(x), \nabla \bar{\varphi}(x)) \quad x \in \Omega, \tag{ELG}$$

$$\bar{\varphi}(x) = g(x) \quad x \in \partial\Omega,$$

also eine nichtlineare PDGL zweiten Grades.

Beweis Zum Beweis wählen wir eine beliebige Testfunktion $v \in C_0^\infty(\Omega)$. Da v kompakten Träger hat, gilt $\bar{\varphi} + \tau v = g$ auf $\partial\Omega$ für alle $\tau \in \mathbb{R}$. Die Funktion

$$i : \mathbb{R} \longrightarrow \mathbb{R}$$

$$\tau \longmapsto i(\tau) := I(\bar{\varphi} + \tau v) = \int_\Omega W(x, \bar{\varphi} + \tau v, \nabla \bar{\varphi} + \tau \nabla v) dx$$

ist folglich wohldefiniert und hat nach Voraussetzung ein Minimum in $\tau = 0$. Als Komposition von C^2 -Funktionen ist auch $i \in C^2(\mathbb{R})$ und somit gilt $i'(0) = 0$. Wegen

$$i'(\tau) = \int_{\Omega} W_{\varphi}(x, \bar{\varphi} + \tau v, \nabla \bar{\varphi} + \tau \nabla v) v + \sum_{i=1}^n W_{\xi_i}(x, \bar{\varphi} + \tau v, \nabla \bar{\varphi} + \tau \nabla v) v_{x_i} dx$$

folgt also

$$\int_{\Omega} W_{\varphi}(x, \bar{\varphi}, \nabla \bar{\varphi}) v + \sum_{i=1}^n W_{\xi_i}(x, \bar{\varphi}, \nabla \bar{\varphi}) v_{x_i} dx = 0. \quad (2.2)$$

Gleichung (2.2) nennt man auch „erste Variation“. Beachte: Zur Herleitung dieser Beziehung wäre $W, \bar{\varphi} \in C^1$ hinreichend gewesen. Wir dürfen aber in (2.2) noch einmal partiell integrieren, und erhalten aufgrund des kompakten Trägers von v und deshalb verschwindenden Randtermen

$$\forall v \in C_0^{\infty}(\Omega) : \int_{\Omega} \left[W_{\varphi}(x, \bar{\varphi}, \nabla \bar{\varphi}) - \sum_{i=1}^n (W_{\xi_i}(x, \bar{\varphi}, \nabla \bar{\varphi}))_{x_i} \right] v dx = 0.$$

Das Fundamentallema der Variationsrechnung liefert nun die Behauptung. ■

Bemerkung 2.2

- (i) *Im Falle $N > 1$ erhält man ein gekoppeltes System von nichtlinearen partiellen Differentialgleichungen.*
- (ii) *Mit Hilfe der Bedingung $i''(0) \geq 0$ kann man eine sogenannte „zweite Variation“ herleiten, wie EVANS in [8] zeigt.*
- (iii) *Die Gleichung (2.2) bezeichnet man zudem auch als „schwache Version der EULER-LAGRANGE-Gleichung“. Dies ist darin begründet, dass man die erste Variation unter Bedingungen an das Wachstumsverhalten der partiellen Ableitungen von W auf den Sobolevraum $W_0^{1,p}(\Omega) \ni v$ (wir erinnern daran, dass $W, \bar{\varphi} \in C^1$ hinreichend für (2.2) war) verallgemeinern kann. Die Bedeutung erschließt sich mit der Abhandlung der direkten Methoden im nächsten Abschnitt, in denen Sobolevräume geeignete Funktionenräume für das Funktional I darstellen.*
- (iv) *Wir wissen nun, dass hinreichend glatte Minimierer auch das Randwertproblem (ELG) lösen. Interessant ist aber auch die Frage, ob Lösungen von (ELG) $I(\cdot)$ minimieren. Im eindimensionalen Fall $n = N = 1$ kann man zeigen, dass dies zutrifft, falls zusätzlich*

$$\forall x \in \bar{\Omega} \quad (\varphi, \xi) \mapsto W(x, \varphi, \xi) : \mathbb{R}^2 \longrightarrow \mathbb{R} \quad (2.3)$$

konvex ist, siehe [6, S. 125f.].

2.3. Direkte Methoden

Die klassischen Methoden der Variationsrechnung wurden in der Entstehung der Variationsrechnung zuerst entwickelt. Aus mehreren Gründen sind diese aber beginnend mit HILBERT seit der 20. Jahrhundertwende durch die „direkten Methoden der Variationsrechnung“ abgelöst worden.

Viele Kriterien wie z. B. in (2.3) sind nur für den Fall $n = 1$ aufgestellt worden, welcher interessante Anwendungen ausklammert. Oftmals stellt auch die Lösung des Randwertproblems (ELG) ein weit schwierigeres Unterfangen als die Minimierung des Variationsfunktionals dar. Man führt ein leichter zu lösendes Problem auf ein schwieriger zu lösendes „zurück“. Letztlich ist auch die Forderung an $\bar{\varphi}$ zweimal stetig differenzierbar zu sein nicht nur sehr stark, sondern oftmals auch inadäquat. Wir erinnern an FERMATs Forschungen über Lichtstrahlen. Geht ein Lichtstrahl von einem Medium in ein anderes über, z. B. von Luft nach Wasser, so hat der Lichtweg an der Wasseroberfläche einen Knick. Die Lösung des Variationsproblems wird also höchstens stückweise stetig differenzierbar sein. Wie bereits in der vorigen Bemerkung angedeutet, muss der Raum der Funktionen in denen nach Lösungen gesucht wird, folglich geschickter gewählt werden.² Damit einhergehend muss aber auch die Methodik zur Bestimmung von Minimierern neu aufgestellt werden. Wir können nicht mehr davon ausgehen, dass eine Lösung des Variationsproblems im klassischen Sinne differenzierbar ist. Außerdem geht man den im Vergleich zu den klassischen Methoden umgekehrten Weg: Zu einer gegebenen elliptischen partiellen Differentialgleichung sucht man eine variationelle Formulierung, deren EULER-LAGRANGE-Gleichung gerade die elliptische PDGL ergibt. Unter gewissen Voraussetzungen und in einem gewissen Sinne löst dann ein Minimum des Variationsfunktionals diese elliptische PDGL.

Im nächsten Unterabschnitt werden wir die für die Anpassung nötigen Definitionen und Grundlagen sowohl bereitstellen als auch motivieren.

2.3.1. Grundlagen

Es sei nach wie vor $\Omega \subseteq \mathbb{R}^n$ offen, aber für die nachfolgenden Definitionen nicht notwendigerweise beschränkt. Im Folgenden betrachten wir wie auch bei den klassischen Methoden den skalaren Fall reellwertiger Funktionen ($N = 1$). Dies hat mehrere Gründe: Die Aussagen und Definitionen lassen sich durch komponentenweise Betrachtung unmittelbar auf den vektoriellen Fall übertragen. Außerdem wird unser erster Existenzsatz für den skalaren Fall formuliert. Zwar gelten die Existenzsätze zumeist auch fast wortwörtlich im vektoriellen Fall, können dann aber erheblich verbessert werden. Wie wir sehen werden, gibt es physikalisch und mathematisch interessante Probleme, welche die Voraussetzungen der skalaren Existenzsätze nicht erfüllen: Sie verletzen die Konvexitätsbedingung. Dies macht eine gesonderte Behandlung des vektoriellen Falls ohnehin notwendig.

Die geeigneten Funktionenräume für das Variationsproblem sind die (allgemein für $k \in \mathbb{N}$, $p \in [1, \infty]$ definierten) „Sobolevräume“ $W^{k,p}(\Omega)$, und von diesen speziell $W^{1,p}(\Omega)$ für

² Dies ist so zu verstehen, dass wir die zulässigen Funktionen nicht nur zweckmäßiger wählen, sondern damit einhergehend insbesondere auch die Anzahl der Kandidaten vergrößern. Dies erhöht im Allgemeinen die Chance, tatsächlich Minimierer zu finden.

$1 < p < \infty$.

$W^{k,p}(\Omega)$ sind diejenigen Funktionen, deren partiellen Ableitungen bis zur Ordnung k im schwachen Ableitungssinne alle existieren und zu $L^p(\Omega)$ gehören. Hierbei nennt man bekanntermaßen $f \in L^p(\Omega)$ „schwach differenzierbar nach x_i mit schwacher Ableitung $g \in L^p(\Omega)$ “, falls

$$\forall v \in C_0^\infty(\Omega) \quad \int_{\Omega} g(x)v(x) dx = - \int_{\Omega} f(x)\partial_i v(x) dx .$$

Höhere schwache Ableitungen der Ordnung $|\alpha| \leq k$ (in Multiindexschreibweise) werden nach obigem Muster entsprechend via

$$\forall v \in C_0^\infty(\Omega) \quad \int_{\Omega} g(x)v(x) dx = (-1)^{|\alpha|} \int_{\Omega} f(x)D^\alpha v(x) dx$$

erklärt.

$$\|u\|_{W^{k,p}(\Omega)} := \|u\|_{k,p,\Omega} := \begin{cases} \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^p(\Omega)}^p \right)^{1/p} & \text{für } p \in [1, \infty) , \\ \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty(\Omega)} & \text{für } p = \infty \end{cases}$$

für $u \in W^{k,p}(\Omega)$ definiert eine Norm auf $W^{k,p}(\Omega)$. Alle $W^{k,p}(\Omega)$ sind vollständig, also BANACHräume. Speziell folgt, dass $W^{k,2}(\Omega)$ ein HILBERTraum ist, da dies für $L^2(\Omega)$ gilt. Man schreibt dann $H^k(\Omega) := W^{k,2}(\Omega)$. Für $k = 0$ ist $W^{0,p}(\Omega)$ als $L^p(\Omega)$ und speziell $H^0(\Omega)$ als $L^2(\Omega)$ aufzufassen.

Warum wir den Fall $k = 1$ erörtern ist aus den bisherigen Überlegungen sofort ersichtlich. Dass wir uns auf den Fall $p \in (1, \infty)$ beschränken bedarf hingegen einer Erklärung. Diese werden wir später nachreichen.

Zunächst möchten wir darauf aufmerksam machen, dass wir, wie es bei den klassischen Methoden bereits der Fall war, im Allgemeinen Randwertprobleme behandeln. Außer für den Anwendungsfall uninteressante Spezialfälle ist jedoch der Rand $\partial\Omega$ einer offenen und beschränkten Menge Ω eine Nullmenge. Von Randwerten von Funktionen des Raumes $W^{1,p}(\Omega)$ zu reden, macht also zunächst keinen Sinn, da L^p -Funktionen nur fast überall erklärt sind. Um dieses Problem zu umgehen, fassen wir Randwerte immer im Sinne des Spuroperators auf ($p < \infty$):

Satz 2.3 (SPUROPERATOR)

Falls Ω einen C^1 -Rand besitzt, so existiert ein stetiger linearer Operator

$$T : W^{1,p}(\Omega) \longrightarrow L^p(\partial\Omega)$$

derart, dass

$$Tu = u|_{\partial\Omega} \quad \text{für } u \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$$

und

$$\forall u \in W^{1,p}(\Omega) \quad \|Tu\|_{L^p(\partial\Omega)} \leq C \|u\|_{1,p,\Omega} ,$$

wobei C eine Konstante ist, die nur von Ω und p abhängt.

Wir nennen Tu die „Spur von u auf $\partial\Omega$ “. Dank des Spuroperators und eines weiteren Resultates können wir nun diejenigen Funktionen, deren Spur auf $\partial\Omega$ verschwindet, auf folgende Art und Weise charakterisieren:

Satz 2.4

Unter obigen Voraussetzungen und Bezeichnungen gilt für $u \in W^{1,p}(\Omega)$

$$u \in W_0^{1,p}(\Omega) \quad \Leftrightarrow \quad Tu = 0 \text{ auf } \partial\Omega .$$

Hierbei ist allgemein

$$W_0^{k,p}(\Omega) := \overline{C_0^\infty(\Omega)}^{W^{k,p}(\Omega)}$$

erklärt. Eine Funktion in $W^{1,p}(\Omega)$ deren Spur auf $\partial\Omega$ verschwindet, ist also Grenzwert von $C_0^\infty(\Omega)$ -Funktionen bezüglich der $W^{1,p}(\Omega)$ -Norm. Da T ein linearer Operator ist, ist nun auch klar was es bedeutet, wenn zwei Funktionen im Spursinne übereinstimmende Randwerte besitzen: Die Differenz der beiden Funktionen liegt in $W_0^{1,p}(\Omega)$. Entsprechend ist bei Randwertproblemen die Schreibweise

$$u \in u_0 + W_0^{1,p}(\Omega)$$

zu verstehen: $u = u_0$ auf $\partial\Omega$ und $u \in W^{1,p}(\Omega)$.

Nachdem wir die Menge der zulässigen Funktionen angepasst haben, ist die Frage, welche weiteren Forderungen wir an unser Funktional $I(\cdot)$ stellen müssen um die Existenz von Minimierern zu garantieren. Diese Forderungen werden wir natürlich letztlich an W stellen.³ Klar ist:

Zunächst einmal brauchen wir nicht nach Minimierern zu suchen, falls es entweder keine Kandidaten gibt, oder $\{I(\varphi) : \varphi \in \mathcal{A}\} \subseteq \mathbb{R}$ nicht nach unten beschränkt ist. Die natürlichsten aller Forderungen stellen folglich $\mathcal{A} \neq \emptyset$ und

$$m := \inf_{\varphi \in \mathcal{A}} \{I(\varphi)\} > -\infty$$

dar. Insbesondere wenn wir Werte in \mathbb{R}^∞ zulassen, ist $m = \infty$ denkbar. Auch diesen uninteressanten Sonderfall wollen wir von nun an ausklammern, denn dann wäre jedes $\varphi \in \mathcal{A}$ Lösung des Variationsproblems.

Um die weitere Vorgehensweise zu motivieren, stelle man sich den endlich-dimensionalen Fall $I : \mathbb{R}^N \rightarrow \mathbb{R}$ vor: Wählen wir eine sogenannte „Minimalfolge“ $(x_k)_{k \in \mathbb{N}}$, das bedeutet $I(x_k) \searrow m$, und ist

- sowohl I stetig,
- als auch $(x_k) \subseteq \mathbb{R}^N$ (relativ) kompakt,

so gilt $\lim_{k \rightarrow \infty} x_{\pi(k)} = \bar{x}$ für eine Teilfolge $(x_{\pi(k)})_{k \in \mathbb{N}}$, und somit

$$\lim_{k \rightarrow \infty} I(x_{\pi(k)}) = I(\lim_{k \rightarrow \infty} x_{\pi(k)}) = I(\bar{x}) .$$

³ Vereinfacht gesprochen übertragen sich die Eigenschaften ja aufgrund der Linearität des Integrals.

Das Infimum wird also tatsächlich angenommen und \bar{x} ist ein Minimum des Variationsfunktional.

Wir wollen jedoch die Voraussetzungen so gering wie möglich halten. Als erstes sieht man, dass die Stetigkeitsforderung an das Funktional I abgeschwächt werden kann. Es genügt offenbar,

$$x_k \xrightarrow{k \rightarrow \infty} \bar{x} \quad \Rightarrow \quad I(\bar{x}) \leq \liminf_{k \rightarrow \infty} I(x_k)$$

für alle gegen \bar{x} konvergente Folgen $(x_k)_{k \in \mathbb{N}}$ zu verlangen.⁴ Solche Funktionen nennt man im Punkt \bar{x} „von unten folgenhalbstetig“, bzw. „unterhalbstetig“ oder auch kurz „UHS“. Insbesondere ist also jede in \bar{x} stetige Funktion dort auch UHS.

Im Folgenden sei X stets ein normierter Raum.

Wesentlich für den Nachweis der Existenz eines Minimums war jedoch unsere zweite Forderung, das Kompaktheitskriterium. Um ein eindimensionales Beispiel ($I : \mathbb{R} \rightarrow \mathbb{R}$) zu geben: Funktionen I der Bauart $I(x) := 1/(1+x^2)$ möchten wir ausschließen, denn dann würde $I(x_k) \searrow m$ nicht mehr die Beschränktheit der Folge $(x_k)_{k \in \mathbb{N}}$ implizieren. Dass dieser Fall nicht eintreten kann, garantiert die folgende Bedingung:

Definition 2.5 (KOERZIVITÄT)

Sei $I : X \rightarrow \mathbb{R}$ ein Funktional. Dann nennen wir I „koerziv“, falls

$$\|x_k\| \rightarrow \infty \quad \Rightarrow \quad I(x_k) \rightarrow \infty. \quad (2.4)$$

Diese Definition gemäß Werner [14] ist bewusst allgemein gehalten. Je nach Problemstellung und entsprechendem Existenzsatz werden wir eine passende Koerzitivitätsbedingung an W stellen. Wir wissen nun:

$$\forall m < C \in \mathbb{R} \quad \exists \tilde{C} > 0 \quad \text{mit} \quad |I(x)| \leq C \quad \Rightarrow \quad \|x\| \leq \tilde{C}.$$

In unserer endlich-dimensionalen Anschauung stünden nun alle Zutaten bereit, um sofort einen Existenzsatz mit trivialem Beweis anzugeben. Lassen wir aber von unserer Anschauung ab, und erinnern uns daran, dass unsere zulässigen Funktionen in dem unendlichdimensionalen Sobolevraum $W^{1,p}(\Omega)$ liegen, so stellt uns eine direkte Konsequenz aus dem Lemma von RIESZ vor ein Problem:

Satz 2.6 (KOMPAKTHEIT DER EINHEITSKUGEL)

Folgende Aussagen sind äquivalent:

- (i) $\dim X < \infty$.
- (ii) $\overline{B_1(0)} := \{x \in X : \|x\| \leq 1\}$ ist kompakt.
- (iii) Jede in X beschränkte Folge hat eine konvergente Teilfolge.

Beweis [14, S. 27f.]

⁴ Denn es folgt dann $I(\bar{x}) \leq \liminf_{k \rightarrow \infty} I(x_k) = m = \inf_{x \in \mathbb{R}^N} \{I(x)\}$.

Bemerkung 2.7

Die Aussage (ii) gilt dann natürlich auch für jede andere Kugel $\overline{B_R(x)}$ mit $x \in X$ und $R > 0$.

Die Existenz einer bzgl. der $W^{1,p}(\Omega)$ -Norm konvergenten Teilfolge ist somit nicht mehr garantiert. Abhilfe schafft das folgende Resultat, welches auch als Spezialfall des Satzes von BANACH-ALAOĞLU-BOURBAKI bewiesen werden kann:

Satz 2.8 (SCHWACHE FOLGENKOMPAKTHEIT)

In einem reflexiven normierten Raum⁵ X besitzt jede beschränkte Folge eine schwach konvergente Teilfolge.

Beweis [14, S. 107]

Um die Aussagen nachvollziehen zu können, müssen wir zunächst ein paar neue Begrifflichkeiten klären:

Für einen normierten Raum X ist dessen „Dualraum“ X' vermöge $X' := \mathcal{L}(X, \mathbb{K})$ erklärt. X' ist also die Menge aller stetigen und linearen Operatoren des Raumes X in seinen Grundkörper \mathbb{K} . Bei uns gilt immer $\mathbb{K} = \mathbb{R}$. X' ist dann wieder ein normierter Raum und immer vollständig, da \mathbb{K} vollständig ist. Nun kann man natürlich auch den Dualraum $(X)'$ zu X' bilden (usw.). Diesen wollen wir mit X'' bezeichnen und den „Bidualraum“ zu X nennen. Es gilt also $X'' = \mathcal{L}(\mathcal{L}(X, \mathbb{K}), \mathbb{K})$. Durch

$$\begin{aligned} J : X &\longrightarrow X'' \\ x &\longmapsto Jx \end{aligned}$$

mit $Jx(x') := x'(x)$ wird dann eine wohldefinierte, lineare und isometrische (und damit injektive) Abbildung definiert, wie man leicht nachprüft. Ist J darüber hinaus sogar surjektiv, so sind also X und X'' isomorph. In diesem Fall nennt man X „reflexiv“. In Räumen mit dieser angenehmen Eigenschaft besteht dann die Kette X, X', X'', \dots im Wesentlichen nur aus den beiden Räumen X und X' .

Nun müssen wir noch erklären, was wir unter schwacher Konvergenz verstehen:

Definition 2.9 (SCHWACHE KONVERGENZ)

Eine Folge $(x_k)_{k \in \mathbb{N}} \subseteq X$ heißt „*schwach konvergent mit Grenzwert $x \in X$* “, falls

$$\forall x' \in X' \quad x'(x_k) \xrightarrow{k \rightarrow \infty} x'(x).$$

Schreibweise: $x_k \rightharpoonup x$.

⁵ Per Definition ist dann X sogar vollständig, da isomorph zum vollständigen Bidualraum X'' . Genauer könnte man auch formulieren, dass die abgeschlossene Einheitskugel (und damit auch jede andere abgeschlossene Kugel, s.o.) in reflexiven BANACH-Räumen „*schwach kompakt*“, also kompakt bzgl. der schwachen Topologie ist.

Bemerkung 2.10

- (i) *In endlich-dimensionalen Räumen sind schwache Konvergenz und (die übliche bzw. fortan manchmal als „stark“ bezeichnete) Konvergenz äquivalent, wie man mit Hilfe des RIESZschen Darstellungssatzes sofort zeigt.*
- (ii) *Auch sind natürlich stark konvergente Folgen aufgrund der Stetigkeit der x' insbesondere schwach konvergent.*
- (iii) *Relativ leicht zeigt man zudem, dass schwache Grenzwerte eindeutig sind.*
- (iv) *Nicht offensichtlich ist allerdings, dass schwach konvergente Folgen auch beschränkt sind. Dass dies dennoch der Fall ist, kann mit Hilfe des Satzes von BANACH-STEINHAUS nachgewiesen werden, siehe [1, S. 227].*

Die Koerzivität liefert uns nun beschränkte Folgen, aus denen wir schwach konvergente Teilfolgen auswählen können, sofern X reflexiv ist. Bevor wir uns darüber Gedanken machen, welche Räume reflexiv sind, müssen wir folglich noch die Definition der Unterhalbstetigkeit auf schwache Konvergenz verallgemeinern.⁶

Definition 2.11 (SCHWACHE FOLGENUNTERHALBSTETIGKEIT)

Eine Funktion $I : X \rightarrow \mathbb{R}$ heißt im Punkt $\bar{x} \in X$ „schwach von unten folgenhalbstetig“, bzw. „schwach unterhalbstetig“ oder auch kurz „SUHS“, falls gilt:

$$\forall (x_k)_{k \in \mathbb{N}} \subseteq X : \quad x_k \rightharpoonup \bar{x} \quad \Rightarrow \quad I(\bar{x}) \leq \liminf_{k \rightarrow \infty} I(x_k) . \quad (2.5)$$

Entsprechend nennt man I schwach unterhalbstetig, falls I in allen $x \in X$ schwach unterhalbstetig ist.

Zur Reflexivität: Als geeigneter Raum für das Variationsproblem hat sich $W^{1,p}(\Omega)$, also ein Teilraum von $L^p(\Omega)$ herauskristallisiert. Nun ist nach einem bekannten Resultat für $p \in [1, \infty)$ und dem durch $\frac{1}{p} + \frac{1}{q} = 1$ erklärten⁷ konjugierten Exponenten q die Abbildung

$$\begin{aligned} j : L^q(\Omega) &\longrightarrow (L^p(\Omega))' \\ v &\longmapsto jv \quad \text{mit} \quad jv(u) := \int_{\Omega} vu \, dx \end{aligned} \quad (2.6)$$

ein isometrischer Isomorphismus. Der Dualraum X' zu $X := L^p(\Omega)$ für $p \in [1, \infty)$ ist also vermöge des konjugierten Exponenten $q \in (1, \infty]$ durch $X' = L^q(\Omega)$ gegeben. Nochmalige Anwendung von (2.6) auf $L^q(\Omega)$ zeigt:

$$L^p(\Omega) \text{ ist reflexiv, falls } p \in (1, \infty) .$$

⁶ Dies wird deshalb nötig, da schwache Konvergenz nicht mehr die Konvergenz bezüglich der Norm in X bedeutet, sondern in der Topologie des Dualraumes X' erklärt ist. Im HILBERTraumfall wird dies besonders anschaulich: Die schwache Konvergenz wird über das Skalarprodukt bestimmt. Dies kann selbst für einfache Folgen unerwünschte Effekte nach sich ziehen. So stimmt i. Allg. das Quadrat des schwachen Grenzwertes einer Folge nicht mit dem schwachen Grenzwert des Quadrates der Folge überein, oder dergleichen. Stetigkeit bzw. UHS von I genügt also nicht.

⁷ $\frac{1}{\infty} := 0$, also $p = 1 \Leftrightarrow q = \infty$.

Damit ist schließlich die Begründung nachgeliefert, warum wir $W^{1,p}(\Omega)$ nur für $1 < p < \infty$ betrachten. Denn

$$\varphi_k \rightharpoonup \varphi \text{ in } W^{1,p}(\Omega)$$

bedeutet schwache Konvergenz der φ_k und ihrer schwachen Ableitungen in $L^p(\Omega)$ gegen φ und seine schwachen Ableitungen. Also

$$\begin{cases} \varphi_k \rightharpoonup \varphi & \text{in } L^p(\Omega), \\ D\varphi_k \rightharpoonup D\varphi & \text{in } L^p(\Omega, \mathbb{R}^n) \end{cases}$$

entsprechend (2.6) und Definition 2.9. Zum Beispiel gilt

$$\varphi_k \rightharpoonup \varphi \text{ in } L^p(\Omega) \iff \forall v \in L^q(\Omega) \quad \int_{\Omega} v\varphi_k dx \longrightarrow \int_{\Omega} v\varphi dx .$$

2.3.2. Ein Existenzsatz im skalaren Fall

Nachdem nun alle Grundlagen bereitgestellt sind, können wir eine Existenzaussage gemäß [8, S. 448] formulieren:

Satz 2.12 (EXISTENZSATZ IM SKALAREN FALL)

Seien $\Omega \subseteq \mathbb{R}^n$ offen und beschränkt mit glattem Rand $\partial\Omega$, $W = W(x, \varphi, \xi)$ glatt und $p \in (1, \infty)$. Mit $\varphi_0 \in W^{1,p}(\Omega)$ sei $I(\varphi_0) < \infty$ und

$$\mathcal{A} := \{ \varphi \in W^{1,p}(\Omega) : \varphi \in \varphi_0 + W_0^{1,p}(\Omega) \} ,$$

also $\mathcal{A} \neq \emptyset$ und $m = \inf_{\varphi \in \mathcal{A}} \{ I(\varphi) \} < \infty$.

$W : \Omega \times \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^\infty$ erfülle die Koerzivitätsbedingung

$$\exists a > 0, b \geq 0 \quad \forall (x, \varphi, \xi) \in \Omega \times \mathbb{R} \times \mathbb{R}^n \quad W(x, \varphi, \xi) \geq a \|\xi\|^p - b, \quad (2.7)$$

und das Funktional I sei SUHS.

Dann gibt es mindestens eine Lösung $\bar{\varphi} \in \mathcal{A}$ des Variationsproblems, sprich

$$I(\bar{\varphi}) = \min_{\varphi \in \mathcal{A}} \{ I(\varphi) \} .$$

Bemerkung 2.13

- (i) Wir haben den Satz nur für den Fall $N = 1$ formuliert, aber natürlich ist er auch für den skalaren Fall $n = 1$ gültig, d.h. $\varphi : \Omega \subseteq \mathbb{R} \longrightarrow \mathbb{R}^N$ mit $N \geq 1$.
- (ii) Dass (2.7) eine Koerzivitätsbedingung gemäß Definition 2.5 respektive (2.4) darstellt, ist nicht sofort ersichtlich, da keine Bedingung an x oder an φ selber gestellt wird. Der Grund liegt in der POINCARÉsche Ungleichung, welche es uns erlaubt die Funktionswerte von φ durch seine Ableitungen zu kontrollieren.

- (iii) Die Wahl des Exponenten p in der Koerzivitatsbedingung garantiert gerade, dass $I(\varphi)$ fur alle $\varphi \in \mathcal{A}$ definiert ist, den Fall $I(\varphi) = \infty$ jedoch nicht ausgeschlossen.
- (iv) Den Glattheitsanspruch an $\partial\Omega$ spezifiziert DACOROGNA als LIPSCHITZrand. Vereinfacht gesprochen muss also die bijektive volle Bilddarstellung eines jeden Randpunktes mitsamt der Umkehrfunktion LIPSCHITZ-stetig sein. Selbst wenn man nur fordert, dass W stetig ist, so lasst sich EVANS' Resultat unter diesen Randbedingungen auch dann noch auf den vektoriellen Fall verallgemeinern.
- (v) DACOROGNA formuliert einen weiteren Existenzsatz mit noch schwacheren Voraussetzungen an W als die Stetigkeit,⁸ dafur aber einer Koerzivitatsbedingung, welche sich auch auf x und φ bezieht. Auch dieser Existenzsatz ist ebenso im vektoriellen Fall gultig.
- (vi) Fragen die bisher ganzlich ausgeklammert wurden, sind die der Regularitat, Stabilitat und Eindeutigkeit von Losungen. Wir werden im abschließenden Kapitel uber nichtlineare Elastizitat kurz darauf zuruckkommen. Es sei bereits an dieser Stelle angemerkt, dass je nach Problemstellung oftmals gar keine Eindeutigkeit aufgrund der physikalischen Bedingungen zu erwarten und somit gewunscht ist.

2.3.3. Der vektorielle Fall

Wir wissen nun, dass **Satz 2.12** (nach notationeller Anpassung) auch im allgemeinen Fall $n, N \in \mathbb{N}$ gultig ist. Zu Recht kann man sich deshalb fragen, warum der vektorielle Fall uberhaupt gesondert betrachtet wird. Der Grund liegt in der notigen SUHS des Funktionals I . Diese „von Hand“ fur ein konkretes Funktional I nachzuweisen, stellt aufgrund der unhandlichen Definition der SUHS ein schwieriges Unterfangen dar. Abhilfe schafft das folgende Resultat:

Satz 2.14

Sei W nach unten beschrankt und die Abbildung $\xi \mapsto W(x, \varphi, \xi)$ fur alle $x \in \Omega$, $\varphi \in \mathbb{R}^N$ konvex. Dann ist das Funktional I SUHS auf $W^{1,p}(\Omega, \mathbb{R}^N)$.

Unter der Pramisse der Konvexitat von $\xi \mapsto W(x, \varphi, \xi)$ werden denn auch die Existenzsatze in der Regel formuliert. Obige Bemerkung bezuglich der leichten Abwandlung der Voraussetzungen gilt entsprechend. Aus gutem Grund haben wir den Existenzsatz dennoch anders formuliert. Denn obgleich man im skalaren Fall (nach TONELLI) auch die Umkehrung von Satz 2.14 zeigen kann, ist im vektoriellen Fall ($n, N \geq 2$) die Konvexitat bzgl. ξ keine notwendige Bedingung fur die SUHS des Funktionals I .

⁸ W muss eine CARATHÉODORY-Funktion sein, d.h. fur fast alle $x \in \Omega$ stetig als Funktion von $(\varphi, \xi) \in \mathbb{R}^N \times \mathbb{R}^{N \times n}$, und fur alle $(\varphi, \xi) \in \mathbb{R}^N \times \mathbb{R}^{N \times n}$ messbar als Funktion von x .

Und tatsächlich gibt es viele Anwendungen, für die diese Konvexitätsbedingung verletzt ist, da sie einfachsten physikalischen Gesetzmäßigkeiten widersprechen würde. So beschreibt in unserem Anwendungsfall der nichtlinearen Elastizität ($n = N = 3$) W die Energie eines Materials unter einer Deformation φ . Mit $F := \xi := \nabla\varphi$ sei der „Deformationsgradient“ bezeichnet. Untersucht man nun den Fall $\det F \searrow 0$, so bedeutet dies, dass der „Körper“ auf „Volumen 0“ zusammengepresst wird. Man sollte folglich erwarten, dass die Energie W „explodiert“: $W \rightarrow \infty$. Wäre nun W tatsächlich konvex bezüglich F , so zeigt CIARLET in [5, S. 170ff.], dass dieses natürliche Verhalten der Energie dann nicht mehr gegeben ist. Auch das Axiom der „frame-indifference“⁹ würde nicht mehr gelten. Man musste sich also über realistischere Bedingungen Gedanken machen, welche einem die SUHS des Funktionals I garantieren.

Auf der Suche nach solchen hinreichenden Bedingungen traten verschiedene Konzepte zutage. Die sogenannte „Quasikonvexität“ liefert eine äquivalente Bedingung zur SUHS von I (und damit das Pendant zur Konvexität bzgl. F im skalaren Fall), ist aber aufgrund einer punktwisen Definition nur schwer nachzuweisen. Eine leicht schwächere Variante stellt die „rank-1-Konvexität“ dar, welche das richtige Konzept im Zusammenhang mit den EULER-LAGRANGE-Gleichungen liefert: Sie ist äquivalent zur Elliptizität der (ELG). Es war jedoch John Ball im Jahre 1977, welcher mit einem weiteren und stärkeren Konzept, dem der „Polykonvexität“, in seinen wegweisenden Arbeiten [2, 3] das adäquate Mittel zur Behandlung vektorieller Variationsprobleme erkannte und bereitstellte. Da Quasi- und rank-1-Konvexität nicht Gegenstand dieser Arbeit sind, geben wir die Definitionen nicht an. Wir bemerken lediglich,¹⁰ dass für eine Funktion W die folgende Implikationskette allgemeine Gültigkeit besitzt:

$$W \text{ konvex} \quad \Rightarrow \quad W \text{ polykonvex} \quad \Rightarrow \quad W \text{ quasikonvex} \quad \Rightarrow \quad W \text{ rank-1-konvex}.$$

Im skalaren Fall sind all diese Konzepte äquivalent, wohingegen im vektoriellen Fall i. Allg. keine der Umkehrungen richtig ist ($N \geq 3$ und $n \geq 2$ im Falle W rank-1-konvex $\not\Rightarrow$ W quasikonvex, $n, N \geq 2$ sonst).

Man kann Polykonvexität gemäß DACOROGNA für allgemeine $n, N \in \mathbb{N}$ definieren, muss dann aber einen gewissen Aufwand für die Notation betreiben. Wir begnügen uns mit dem Anwendungsfall $n = N = 3$, und behalten die Schreibweise $F = \nabla\varphi$ bei. Die Definition der Polykonvexität lautet dann wie folgt:

Definition 2.15 (POLYKONVEXITÄT)

Mit $\Omega \subseteq \mathbb{R}^3$ sei nun also

$$W : \Omega \times \mathbb{R}^3 \times \mathbb{R}^{3 \times 3} \longrightarrow \mathbb{R}^\infty, \quad W = W(x, \varphi, F).$$

Für $\mathbb{R}^{3 \times 3}$ schreiben wir $\mathbb{M}^{3 \times 3}$, die Menge aller reellen 3×3 -Matrizen. Dann nennen wir W „polykonvex“, falls die auf $\mathbb{M}^{3 \times 3}$ definierte Abbildung $F \mapsto W(x, \varphi, F)$ für alle $(x, \varphi) \in \Omega \times \mathbb{R}^3$ als konvexe Funktion aller Unterdeterminanten von F geschrieben

⁹ Vereinfacht formuliert (und nicht ausschließlich) bedeutet dies: Die Energie ist rotationsinvariant.

¹⁰ Wie bereits in der Einleitung geschehen.

werden kann.

Genauer: Es gebe eine (i. Allg. nicht eindeutige) Funktion

$$P : \Omega \times \mathbb{R}^3 \times \mathbb{M}^{3 \times 3} \times \mathbb{M}^{3 \times 3} \times \mathbb{R} \longrightarrow \mathbb{R}^\infty$$

derart, dass $W(x, \varphi, F) = P(x, \varphi, F, \text{Adj } F, \det F)$ und die Abbildung

$$\begin{aligned} P_{x,\varphi} : \mathbb{R}^{19} &\longrightarrow \mathbb{R} \\ (A, B, C) &\longmapsto P_{x,\varphi}(A, B, C) := P(x, \varphi, A, B, C) \end{aligned}$$

für alle $(x, \varphi) \in \Omega \times \mathbb{R}^3$ konvex ist.

Bemerkung 2.16

- (i) Mit $\det F$ wird natürlich die Determinante einer Matrix beschrieben, und $\text{Adj } F$ bezeichne die Adjunkte einer Matrix F . Siehe [Definition 3.14](#) auf Seite 32.
- (ii) Der zunächst womöglich willkürlich erscheinenden Definition kann eine physikalische Motivation zugesprochen werden: Repräsentiert W die innere Energiedichte eines Körpers unter „Dehnung und Stauchung“, so mag W explizit von den lokalen Änderungen der Linien- ($\nabla\varphi$), Flächen- ($\text{Adj } \nabla\varphi$) und Raumelemente ($\det \nabla\varphi$) abhängen.
- (iii) Polykonvexität ist eine globale Eigenschaft. In der Tat: KRISTENSEN [9] hat eine glatte nicht-polykonvexe Funktion auf $\mathbb{M}^{2 \times 2}$ mit der Eigenschaft konstruiert, dass die Restriktion dieser Funktion auf beliebige Kugeln mit Radius 1 zu einer auf ganz $\mathbb{M}^{2 \times 2}$ glatten und polykonvexen Funktion fortgesetzt werden kann.

Nehmen wir also an, W sei polykonvex. Möchte man nun Existenzsätze formulieren, so funktioniert der Nachweis ganz ähnlich zu der Vorgehensweise im skalaren (bzw. vektoriellem) Fall unter der (womöglich unrealistischen) Annahme W sei konvex bezüglich $\xi = F$. Jedoch mit einem wichtigen Unterschied: Die Konvexität garantierte uns die schwache Unterhalbstetigkeit des Funktionals I . Im polykonvexen Fall wird die Konvexität aber über die Funktion P aus [Definition 2.15](#) realisiert, welche konvex von den Unterdeterminanten von F abhängt. Deren schwache Stetigkeit ist aber an die Raumdimension gebunden, wie die folgende Aussage aus [8, S. 454] zeigt:

Lemma 2.17 (SCHWACHE STETIGKEIT VON DETERMINANTEN)

Sei $n = N$ und $p \in (N, \infty)$ und

$$\varphi_k \rightharpoonup \varphi \quad \text{in } W^{1,p}(\Omega, \mathbb{R}^N).$$

Dann folgt

$$\det(D\varphi_k) \rightharpoonup \det(D\varphi) \quad \text{in } L^{p/N}(\Omega).$$

Das Aufweichen der Konvexitätsbedingung durch die schwächere Voraussetzung der Polykonvexität geht somit zu Lasten des Exponenten p in $W^{1,p}(\Omega, \mathbb{R}^N)$, dem Raum in dem wir (oder zumindest in Teilmengen hieraus) nach Minimierern suchen, da uns die Polykonvexität die SUHS von I garantiert. In unserer Anwendung müssten wir also schon $p > 3$ fordern, damit die Folge der Determinanten der Ableitungen der Minimalfolge schwach in $L^q(\Omega)$ mit $q > 1$ konvergiert (reflexiv!). Entsprechend müssen die Exponenten in den Koerzivitätsbedingungen diesen Zwängen unterworfen werden. Diese beiden Eigenschaften liefern dann im Wesentlichen die Existenzaussage. Als Beispiel zitieren wir für $n = N = 3$ einen Existenzsatz aus [12, S. 32f.], welcher trotz schwacher Voraussetzungen noch Exponenten p dazugewinnen kann:

Satz 2.18 (EXISTENZSATZ IM VEKTORIELLEN FALL $n = N = 3$)

Seien $\Omega \subseteq \mathbb{R}^3$ offen und beschränkt mit glattem Rand $\partial\Omega$ (z. B. ein LIPSCHITZrand) und $p \in [2, \infty)$. Für ein $\varphi_0 \in W^{1,p}(\Omega, \mathbb{R}^3)$ gelte

$$I(\varphi_0) = \int_{\Omega} W(x, \varphi_0(x), \nabla\varphi_0(x)) dx < \infty ,$$

also $\mathcal{A} \neq \emptyset$ und $m = \inf_{\varphi \in \mathcal{A}} \{I(\varphi)\} < \infty$ für

$$\mathcal{A} := \left\{ \varphi \in W^{1,p}(\Omega, \mathbb{R}^3) : \varphi \in \varphi_0 + W_0^{1,p}(\Omega, \mathbb{R}^3) \right\} .$$

Des Weiteren gelte für W :

- (i) $W : \Omega \times \mathbb{R}^3 \times \mathbb{M}^{3 \times 3} \rightarrow \mathbb{R}^\infty$ ist eine CARATHÉODORY-Funktion (vgl. *Bemerkung 2.13*).
- (ii) $W(x, \varphi, \cdot)$ ist polykonvex für alle $\varphi \in \mathbb{R}^3$ und fast alle $x \in \Omega$.
- (iii) Für $r > 1$ und $c \in \mathbb{R}_+$ erfülle W eine der folgenden Koerzivitätsbedingungen bzgl. einer Matrixnorm $\|\cdot\|$ (siehe (3.3) auf Seite 25 bzw. *Definition 3.16* auf Seite 32):

$$W(x, \varphi, F) \geq \begin{cases} c(\|F\|^p - 1) & \text{für } p \geq 3 , \\ c(\|F\|^p + \|\text{Adj } A\|^q - 1) & \text{für } p \in [2, 3), q > \frac{3}{2} , \\ c(\|F\|^p + \|\text{Adj } A\|^q + |\det F|^r - 1) & \text{für } p \in [2, 3), q > \frac{3p}{4p-3} . \end{cases}$$

Dann gibt es mindestens eine Lösung $\bar{\varphi} \in \mathcal{A}$ des Variationsproblems, sprich

$$I(\bar{\varphi}) = \min_{\varphi \in \mathcal{A}} \{I(\varphi)\} .$$

Da wir zur Motivation bereits hier und dort der nichtlinearen Elastizität vorweggegriffen haben, wissen wir bereits, dass eine freie Energiefunktion W , welche zu starke Kompressionen energetisch bestraft, nicht konvex in $F = \nabla\varphi$ sein kann. Formuliert man aber W als Funktion $\widetilde{W} = \widetilde{W}(C)$ des rechten CAUCHY-GREEN-Deformationstensors $C := F^T F$ um, so darf \widetilde{W} als konvexe Funktion bezüglich C sehr wohl eine Singularität für $\det C \searrow 0$ besitzen. Aus diesem Grunde überlegen wir uns im nächsten Kapitel, wie Funktionen $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ aussehen müssen, damit $f \circ \det$ konvex als Funktion von C ist.

3. Konvexität von $f \circ \det$ auf $\mathbb{P}Sym(3)$

PROBLEM

Wir betrachten eine Funktion $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ und fragen uns, welcher Gestalt f sein muss, damit

$$\begin{aligned} f \circ \det & : \mathbb{P}Sym(3) \longrightarrow \mathbb{R} \\ C & \longmapsto f(\det C) \end{aligned} \tag{3.1}$$

konvex ist. Hierbei werden wir voraussetzen, dass f hinreichend glatt, genauer: zweimal stetig differenzierbar ist.¹

In diesem Kapitel bezeichnen X, Y stets (reelle) normierte Räume, und es sei $a \in \Omega \subseteq X$.

3.1. Eine Differentialungleichung an die Funktion f

3.1.1. Vorbereitende Lemmata und Definitionen

Für $u, v \in X$ nennen wir

$$[u, v] := \{(1 - \tau)u + \tau v : \tau \in [0, 1]\}$$

die „Strecke mit den Endpunkten u, v “.

Definition 3.1 (KONVEXITÄT)

(i) $g : I(\text{Intervall}) \subseteq \mathbb{R} \rightarrow \mathbb{R}$ heißt „konvex“

$$:\Leftrightarrow \forall u, v \in I \quad \forall \tau \in [0, 1] : \quad g((1 - \tau)u + \tau v) \leq (1 - \tau)g(u) + \tau g(v) .$$

(ii) $K \subseteq X$ heißt „konvex“

$$:\Leftrightarrow \forall u, v \in K : \quad [u, v] \subseteq K .$$

¹ Wie eingangs erwähnt, handelt es sich bei C um den rechten CAUCHY-GREEN-Deformationstensor. Es ist $C := F^T F \in \mathbb{P}Sym(3)$, wie wir im nächsten Kapitel sehen werden.

(iii) $g : K(\text{konvex}) \subseteq X \rightarrow \mathbb{R}$ heißt „konvex“

$$:\Leftrightarrow \forall u, v \in K \quad \forall \tau \in [0, 1] : \quad g((1 - \tau)u + \tau v) \leq (1 - \tau)g(u) + \tau g(v)$$

$$\stackrel{(i)}{\Leftrightarrow} \forall u, v \in K : \quad g|_{[u, v]} : [0, 1] \subseteq \mathbb{R} \rightarrow \mathbb{R} \\ \lambda \mapsto g(u + \lambda(v - u)) \quad \text{ist}$$

konvex, denn für $\tilde{u}, \tilde{v} \in [u, v]$ ist $[\tilde{u}, \tilde{v}] \subseteq [u, v] \subseteq K$.

Definition 3.2

(i) $\mathbb{M}^{N \times N}$ bezeichne die Menge aller reellen $N \times N$ -Matrizen.

(ii) $\text{Sym}(N)$ bezeichne die Menge aller reellen symmetrischen $N \times N$ -Matrizen.

(iii) $\mathbb{P}\text{Sym}(N)$ bezeichne die Menge aller reellen symmetrischen positiv definiten $N \times N$ -Matrizen.

Hierbei nennen wir $A \in \mathbb{M}^{N \times N}$ „positiv definit“ genau dann, wenn

$$\langle Ax, x \rangle > 0 \quad \forall x \in \mathbb{R}^N \setminus \{0\}.$$

Insbesondere gilt dann, dass A regulär ist und nur positive Eigenwerte besitzt, d.h. es ist $\det A > 0$.

Lemma 3.3

$$A \in \mathbb{M}^{N \times N} \text{ ist positiv definit} \Leftrightarrow \exists c > 0 \quad \forall x \in \mathbb{R}^N : \quad \langle Ax, x \rangle \geq c \cdot \|x\|^2.$$

Beweis „ \Leftarrow “ ist klar.

„ \Rightarrow “

Wir nehmen das Gegenteil an und finden also zu jedem $n \in \mathbb{N}$ ein $x_n \in \mathbb{R}^N \setminus \{0\}$ mit

$$\langle Ax_n, x_n \rangle < \frac{1}{n} \cdot \|x_n\|^2.$$

Mit $y_n := x_n / \|x_n\|$ ist dies gleichbedeutend mit

$$\langle Ay_n, y_n \rangle < \frac{1}{n} \quad \forall n \in \mathbb{N}. \quad (3.2)$$

Da

$$\|y_n\| = 1 \quad \forall n \in \mathbb{N},$$

existiert demnach eine Teilfolge von $(y_n)_{n \in \mathbb{N}}$, welche gegen ein $y \in \mathbb{R}^N$ konvergiert. Die Stetigkeit der Norm liefert $\|y\| = 1$, also $y \neq 0$. Die Stetigkeit des Skalarproduktes durch Grenzübergang für die Teilfolge in (3.2) begründet schließlich

$$\langle Ay, y \rangle \leq 0,$$

ein Widerspruch zur positiven Definitheit von A . ■

Im Folgenden werden wir A als stetigen linearen Operator von \mathbb{R}^N nach \mathbb{R}^N auffassen. Schreibweise:

$$A \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N).$$

Mit der von einer beliebigen Norm $\|\cdot\|_\star$ im \mathbb{R}^N induzierten Operatornorm

$$\|A\|_\star := \sup \{ \|Ax\|_\star : \|x\|_\star \leq 1 \}, \quad (3.3)$$

gilt die bekannte Operatorungleichung:

$$\forall x \in \mathbb{R}^N \quad \|Ax\|_\star \leq \|A\|_\star \cdot \|x\|_\star. \quad (3.4)$$

Den \star zur Kennzeichnung der Norm werden wir in Zukunft vernachlässigen.

Lemma 3.4

$\mathbb{P}\text{Sym}(N)$ ist offene und konvexe Teilmenge des normierten Raumes $\text{Sym}(N)$.

Beweis „zu $\text{Sym}(N)$ ist normierter Raum“

$\text{Sym}(N)$ ist ein Untervektorraum von $\mathbb{M}^{N \times N}$, und die eben eingeführte Norm lässt sich auf diesen einschränken.

„zu $\mathbb{P}\text{Sym}(N)$ ist konvex“

Es gilt für beliebige $A, B \in \mathbb{P}\text{Sym}(N)$ und $\tau \in [0, 1]$:

$$((1 - \tau)A + \tau B)^T = (1 - \tau)A^T + \tau B^T = (1 - \tau)A + \tau B$$

und

$$\left\langle ((1 - \tau)A + \tau B)x, x \right\rangle = (1 - \tau) \underbrace{\langle Ax, x \rangle}_{>0} + \tau \underbrace{\langle Bx, x \rangle}_{>0} > 0 \quad \forall x \in \mathbb{R}^N \setminus \{0\}.$$

„zu $\mathbb{P}\text{Sym}(N)$ ist offene Teilmenge von $\text{Sym}(N)$ “

Wir betrachten ein beliebiges $A \in \mathbb{P}\text{Sym}(N)$. Hierzu $c = c_A > 0$ gemäß [Lemma 3.3](#).

Damit ist

$$B_{c/2}(A) = B_{c/2}^{\text{Sym}(N)}(A) \subseteq \mathbb{P}\text{Sym}(N),$$

denn für beliebiges $S = A + H \in B_{c/2}(A)$, also $\|H\| < c/2$, gilt $\forall x \in \mathbb{R}^N \setminus \{0\}$:

$$\langle Sx, x \rangle = \langle Ax, x \rangle + \langle Hx, x \rangle \stackrel{\text{L. 3.3} + (\text{CSU})^2}{\geq} c \cdot \|x\|^2 - \|H\| \cdot \|x\|^2 > \frac{c}{2} \|x\|^2.$$

$S \in \text{Sym}(N)$ ist also auch positiv definit. ■

Lemma 3.5 (DIAGONALISIERBARKEIT)

Jede Matrix $A \in \text{Sym}(N)$ lässt sich mithilfe einer orthogonalen Matrix Q – d.h. $Q^{-1} = Q^T$ – ähnlich auf eine Diagonalmatrix D transformieren:

$$Q^{-1}AQ = D = \text{diag}(\lambda_1, \dots, \lambda_N),$$

wobei $\lambda_1, \dots, \lambda_N$ die N reellen Eigenwerte von A bezeichnen. A besitzt also N linear unabhängige und zueinander orthogonale Eigenvektoren (nämlich genau die Spalten von Q).

Beweis [7, S. 233] ■

Lemma 3.6

$\text{span}(\mathbb{P}\text{Sym}(3)) = \text{Sym}(3)$, wobei span die lineare Hülle bezeichne.

Beweis Natürlich ist jede Linearkombination positiv definiter symmetrischer Matrizen eine symmetrische Matrix, d.h. „ \subseteq “ ist klar. Wir müssen also zeigen, dass jede symmetrische Matrix A als Linearkombination positiv definiter symmetrischer Matrizen geschrieben werden kann:

Dies ist aufgrund von [Lemma 3.5](#) äquivalent dazu, dass wir jede Diagonalmatrix D als solche Linearkombination schreiben können. Denn

$$D = \sum_{i=1}^m a_i \cdot A_i \quad \text{mit } a_i \in \mathbb{R}, A_i \in \mathbb{P}\text{Sym}(3) \quad \text{für } i = 1, \dots, m,$$

wobei $Q^{-1}AQ = D$, zieht

$$A = \sum_{i=1}^m a_i \cdot QA_iQ^{-1}$$

nach sich. Analog im umgekehrten Fall. Schließlich ist $A_i \in \mathbb{P}\text{Sym}(3)$ genau dann, wenn $QA_iQ^{-1} \in \mathbb{P}\text{Sym}(3)$, denn

$$\mathbb{R}^3 \setminus \{0\} = \{y := Q^{-1}x \mid x \in \mathbb{R}^3 \setminus \{0\}\}$$

und damit

$$\langle QA_iQ^{-1}x, x \rangle = \langle A_iQ^{-1}x, Q^{-1}x \rangle = \langle A_iy, y \rangle.$$

Sei also $D = \text{diag}(a, b, c)$ beliebig. Falls $a, b, c > 0$, gibt es nichts zu zeigen.

Falls D die Nullmatrix ist, schreiben wir $D = 0 \cdot \mathbb{1}$, wobei $\mathbb{1}$ die Einheitsmatrix bezeichne.

Von den übrigen Fällen betrachten wir stellvertretend den Fall $a, -c \in \mathbb{R}_+$, $b = 0$, an dem die allgemeine (selbstverständlich nicht eindeutige) Vorgehensweise deutlich wird:

$$D = \text{diag}(a, b, c) = \text{diag}(a, 0, c) = 2 \cdot \underbrace{\text{diag}(a, 1, 1)}_{\in \mathbb{P}\text{Sym}(3)} - \underbrace{\text{diag}(a, 2, 2 - c)}_{\in \mathbb{P}\text{Sym}(3)}$$

\Rightarrow „ \supseteq “ ■

Bemerkung 3.7

Natürlich gilt die Gleichheit im vorigen Lemma nicht nur im Falle $N = 3$. Da jedoch $N = 3$ der Fall ist, der uns letztlich nur interessiert, haben wir den Beweis der Übersichtlichkeit halber auf diese Dimension beschränkt.

3.1.2. Differenzierbarkeit in normierten Räumen

Das bekannte eindimensionale Resultat

$$g \text{ konvex} \Leftrightarrow g'' \geq 0 \quad (3.5)$$

für $g \in C^2(I, \mathbb{R})$ mit einem Intervall³ $I \subseteq \mathbb{R}$ wollen wir auf reellwertige Funktionen mit Definitionsbereich $K(\text{offen und konvex}) \subseteq X$ verallgemeinern. Hierfür benötigen wir:

Definition 3.8 (DIFFERENZIERBARKEIT)

Eine Funktion $g \in \mathcal{F}(\Omega, Y)$ nennen wir „in a differenzierbar mit der Ableitung $Dg(a) := g'(a) := T \in \mathcal{L}(X, Y)$ “, wenn eine ganze Umgebung $B_\rho(a)$ zu Ω gehört, und dort

$$g(x) = g(a) + T(x - a) + r(x - a)\|x - a\|$$

{ bzw. mit $x = a + h$ für $h \in B_\rho(0)$ äquivalent, aber übersichtlicher formuliert

$$g(a + h) = g(a) + Th + r(h)\|h\| \quad \}$$

gilt, wobei der „Differentiationsrest“

$$r : B_\rho(0) \longrightarrow Y$$

die Eigenschaft

$$\lim_{h \rightarrow 0} r(h) = 0$$

hat. Ist g in allen Punkten $a \in \Omega$ differenzierbar, so nennt man g „differenzierbar“. Ω muss dann also eine offene Menge sein.

³ Dies ist so zu verstehen, dass die zweite Ableitung von g nur im Innern $\overset{\circ}{I}$ des beliebigen Intervalls I zu bilden ist (und diese folglich dort stetig sein soll), wohingegen die Stetigkeit von g selbst auf ganz I gefordert ist. Man könnte die Voraussetzung noch abschwächen, denn es genügt bereits, dass g in $\overset{\circ}{I}$ zweimal differenzierbar ist (aber natürlich weiterhin stetig auf ganz I).

In letzterem Fall erhalten wir also eine Abbildung

$$Dg : \Omega \longrightarrow \mathcal{L}(X, Y)$$

von Ω in den Raum $\mathcal{L}(X, Y)$ der stetigen linearen Operatoren von X nach Y . Mit der unter (3.3) definierten Operatornorm ist $\mathcal{L}(X, Y)$ ein normierter Raum (und sogar ein BANACHraum, falls Y vollständig ist). Wir können und werden nun aber Dg als eine (in der zweiten Komponente lineare) Abbildung von $\Omega \times X$ nach Y auffassen

$$\begin{aligned} Dg : \Omega \times X &\longrightarrow Y \\ (a, h) &\longmapsto Dg(a, h) := Dg(a).h := Dg(a)h \end{aligned}$$

Welche Auffassung man zugrunde legt, ist beispielsweise im Fall $\dim X < \infty$ für die Stetigkeit von Dg nicht von Belang.⁴ Falls also Dg in einer der beiden Auffassungen (und damit im Falle $\dim X < \infty$ auch in der jeweils anderen) stetig ist, so nennen wir g „stetig differenzierbar“ und schreiben hierfür $g \in C^1(\Omega, Y)$.

Wir möchten nun die zweite Ableitung definieren, werden hierbei aber nicht **Definition 3.8** direkt auf $Dg \in \mathcal{F}(\Omega, \mathcal{L}(X, Y))$ anwenden, sondern anlehnend an eben angegebener Auffassung (wiederum im Falle $\dim X < \infty$ in äquivalenter Weise) formulieren⁵:

Definition 3.9 (2. ABLEITUNG)

Eine differenzierbare Funktion $g \in \mathcal{F}(\Omega, Y)$ nennen wir „in a zweimal differenzierbar“, wenn die für alle $k \in X$ definierten (und bzgl. k linearen) Funktionen

$$\begin{aligned} f_k : \Omega &\longrightarrow Y \\ x &\longmapsto f_k(x) := Dg(x).k \end{aligned}$$

alle in a differenzierbar sind. Als „zweite Ableitung von g im Punkt a “ bezeichnet man die durch

$$\begin{aligned} \underbrace{D^2g(a)}_{=:g''(a)} : \underbrace{X \times X}_{=:X^2} &\longrightarrow Y \\ (h, k) &\longmapsto D^2g(a)(h, k) := D^2g(a).(h, k) := Df_k(a)h = Df_k(a).h \end{aligned}$$

definierte bilineare Funktion.

Ist g in allen Punkten $a \in \Omega$ zweimal differenzierbar, so nennt man g „zweimal differenzierbar“.

⁴ Für die Richtung „ $Dg : \Omega \times X \longrightarrow Y$ stetig $\Rightarrow Dg : \Omega \longrightarrow \mathcal{L}(X, Y)$ stetig“ benötigt man die Kompaktheit der Einheitsphäre im normierten Raum X (bedenke (3.3)). Dies ist gleichbedeutend damit, dass X endlich-dimensional ist (siehe **Satz 2.6** auf Seite 14). An Y sind keinerlei zusätzliche Bedingungen zu stellen. Setzt man zweckmäßigerweise $\|\cdot\|_{\Omega \times X} := \|\cdot\|_X + \|\cdot\|_X$, so folgt die Behauptung nun durch einfaches Nachrechnen.

⁵ Siehe **Bemerkung 3.10** auf Seite 29.

In letzterem Fall fassen wir wieder D^2g als eine (in den letzten beiden Komponenten bilineare) Abbildung

$$\begin{aligned} D^2g &: \Omega \times X^2 \longrightarrow Y \\ (a, h, k) &\longmapsto D^2g(a, h, k) := D^2g(a).(h, k) \end{aligned}$$

von $\Omega \times X^2$ nach Y auf. Falls diese Abbildung stetig ist, nennen wir g entsprechend „zweimal stetig differenzierbar“ und schreiben $g \in C^2(\Omega, Y)$.

Im Sonderfall $X = \mathbb{R}^N$, $Y = \mathbb{R}$ ist $D^2g(a) \in \text{Sym}(N)$ nach dem Satz von SCHWARZ,⁶ und man erhält in Matrixschreibweise

$$D^2g(a).(h, k) = (k_1, \dots, k_N) \begin{pmatrix} \partial_{x_1^2}^2 g(a) & \cdots & \partial_{x_1 x_N}^2 g(a) \\ \vdots & \ddots & \vdots \\ \partial_{x_N x_1}^2 g(a) & \cdots & \partial_{x_N^2}^2 g(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_N \end{pmatrix} = \langle k, D^2g(a)h \rangle .$$

Bemerkung 3.10

Definition 3.8 angewandt auf $Dg \in \mathcal{F}(\Omega, \mathcal{L}(X, Y))$ würde für $h \in B_\rho(0)$

$$Dg(a+h) = Dg(a) + Th + r(h)\|h\| \quad (3.6)$$

liefern, wobei $r \in \mathcal{F}(B_\rho(0), \mathcal{L}(X, Y))$ mit $\lim_{h \rightarrow 0} r(h) = 0$ und

$$D^2g(a) := D(Dg)(a) := T \in \mathcal{L}(X, \mathcal{L}(X, Y)) .$$

Auf (3.6) kann man beliebige $k \in X$ anwenden, weshalb man mit

$$\forall k \in X : \quad \rho_k := \rho, \quad r_k(h) := r(h)k \quad \text{und} \quad Df_k(a) \quad \text{durch} \quad Df_k(a)h := Thk \quad \forall h \in X$$

unmittelbar die Formulierung aus *Definition 3.9* erhält:

$$\forall k \in X : \quad f_k(a+h) = f_k(a) + Df_k(a)h + r_k(h)\|h\| \quad (3.7)$$

für $h \in B_{\rho_k}(0)$, $r_k \in \mathcal{F}(B_{\rho_k}(0), Y)$ mit $\lim_{h \rightarrow 0} r_k(h) = 0$ und $Df_k(a) \in \mathcal{L}(X, Y)$.

Um umgekehrt (3.6) aus (3.7) zu folgern, definiert man die entsprechenden Terme natürlich dadurch, dass man die Funktionswerte für alle k kennt. Dass $\inf_{k \in X} \rho_k > 0$ gilt und T stetig ist, kann man aber nur im Falle $\dim X < \infty$ garantieren. Hierzu nutzt man die Linearität von f_k bzgl. k (es genügt (3.7) für eine Basis von X zu kennen!) sowie die Linearität der Ableitung aus, denn es gilt

$$\|T\| = \sup_{\|h\| \leq 1} \|Th\| = \sup_{\|h\| \leq 1} \sup_{\|k\| \leq 1} \|Thk\| \leq \sup_{\|k\| \leq 1} \|Df_k(a)\| .$$

⁶ In diesem Fall genügt es, $D^2g(a)$ auf der „Diagonalen“ zu kennen, denn für jede symmetrische Bilinearform $\alpha : X \times X \rightarrow \mathbb{R}$ gilt

$$\forall v, w \in X \quad \alpha(v, w) = \frac{1}{2} (\alpha(v+w, v+w) - \alpha(v, v) - \alpha(w, w)) . \quad (\text{Polarisationsformel})$$

Die n . Ableitung für $n \in \mathbb{N}_{\geq 3}$ werden wir nun (der Vollständigkeit halber) induktiv einführen:

Definition 3.11 (HÖHERE ABLEITUNGEN)

Eine $(n - 1)$ -mal differenzierbare Funktion $g \in \mathcal{F}(\Omega, Y)$ nennen wir „in a n -mal differenzierbar“, wenn die für alle $k = (k^1, \dots, k^{n-1}) \in X^{n-1}$ definierten (und bzgl. k^1, \dots, k^{n-1} $(n - 1)$ -linearen) Funktionen

$$\begin{aligned} f_k &: \Omega \longrightarrow Y \\ x &\longmapsto f_k(x) := D^{n-1}g(x).(k^1, \dots, k^{n-1}) \end{aligned}$$

alle in a differenzierbar sind. Als „ n . Ableitung von g im Punkt a “ bezeichnet man die durch

$$\begin{aligned} \underbrace{D^n g(a)}_{=: g^{(n)}(a)} &: \underbrace{X \times X^{n-1}}_{=: X^n} \longrightarrow Y \\ (h, k) &\longmapsto D^n g(a)(h, k) := D^n g(a).(h, k) := Df_k(a)h \end{aligned}$$

definierte n -lineare Funktion.

Ist g in allen Punkten $a \in \Omega$ n -mal differenzierbar, so nennt man g „ n -mal differenzierbar“.

In letzterem Fall fassen wir wieder $D^n g$ als (eine in den letzten n Komponenten n -lineare) Abbildung

$$\begin{aligned} D^n g &: \Omega \times X^n \longrightarrow Y \\ (a, h, k) &\longmapsto D^n g(a, h, k) := D^n g(a).(h, k^1, \dots, k^{n-1}) \end{aligned}$$

von $\Omega \times X^n$ nach Y auf. Falls diese Abbildung stetig ist, nennen wir g entsprechend „ n -mal stetig differenzierbar“ und schreiben $g \in C^n(\Omega, Y)$. Wenn g für alle $n \in \mathbb{N}$ n -mal stetig differenzierbar ist, so nennen wir g „unendlich oft differenzierbar“ und schreiben $g \in C^\infty(\Omega, Y)$.

Im Sonderfall $X = \mathbb{R}^N$, $Y = \mathbb{R}$ ist $D^n g(a)$ in den letzten n Komponenten total symmetrisch nach dem Satz von SCHWARZ (es gilt im Übrigen eine Verallgemeinerung der **Polarisationsformel**, siehe [11, S. 103]), und man erhält (wobei wir nun (h, k) ersetzen und als $h \in X^n$ auffassen)

$$D^n g(a).(h^1, \dots, h^n) = \sum_{j_1, \dots, j_n=1}^N \partial_{x_{j_1} \dots x_{j_n}}^n g(a) \cdot h_{j_1}^1 \cdot \dots \cdot h_{j_n}^n .$$

Nun formulieren wir unsere Verallgemeinerung:

Lemma 3.12

Für $g \in C^2(K, \mathbb{R})$ mit K (offen und konvex) $\subseteq X$ gilt:

$$g \text{ konvex} \iff D^2 g(x).(z, z) \geq 0 \quad \forall x \in K, z \in \text{span}(K) . \quad (3.8)$$

Für die Konvexität von g genügt es also nicht, dass $D^2 g(x).(z, z) \geq 0$ nur $\forall x, z \in K$ gilt, wie NEFFs (Gegen-)Beispiel [11, S. 188] zeigt.

Beweis

Die Konvexität von g ist (vgl. [Definition 3.1](#)) äquivalent zu der Konvexität der Einschränkung von g auf jeden Streckenabschnitt in K . Der Richtungsvektor z (entspricht $v - u$ in [Definition 3.1](#)) des Streckenabschnitts liegt i. Allg. jedoch nicht in K , sondern in $\text{span}(K)$. Die Konvexität von g ist also gleichbedeutend zur Konvexität der reellen Funktion

$$g_x(\lambda) := g(x + \lambda z)$$

auf dem offenen (da K n.V. offen) Intervall

$$I_x := \{\lambda \in \mathbb{R} \mid x + \lambda z \in K\}$$

um 0 für alle $x \in K$, $z \in \text{span}(K)$.

Nach unserem klassischen Resultat ist dies äquivalent dazu, dass $g_x''(\lambda) \geq 0$ für $\lambda \in I_x$ gilt. Die Kettenregel liefert schließlich

$$g_x'(\lambda) = Dg(x + \lambda z) \cdot z,$$

und damit

$$g_x''(\lambda) = D^2g(x + \lambda z) \cdot (z, z).$$

Die Behauptung folgt nun, indem man $\lambda = 0$ setzt. Dies ist deshalb möglich, da wir mit $\tilde{x} := x + \lambda z \in K$ für $\lambda \in I_x$

$$D^2g(x + \lambda z) \cdot (z, z) = D^2g(\tilde{x} + 0 \cdot z) \cdot (z, z)$$

schreiben können. ■

Speziell folgt unmittelbar:

Satz 3.13 (KONVEXITÄTSBEDINGUNG)

Für $g \in C^2(\mathbb{P}\text{Sym}(N), \mathbb{R})$ ist

$$g \text{ konvex} \iff D^2g(C) \cdot (H, H) \geq 0 \quad \forall C \in \mathbb{P}\text{Sym}(N), H \in \text{Sym}(N). \quad (3.9)$$

Beweis Wir setzen natürlich

$$K := \mathbb{P}\text{Sym}(N) \quad \text{und} \quad X := \text{Sym}(N),$$

im vorigen Resultat, denn nach [Lemma 3.4](#) ist $\mathbb{P}\text{Sym}(N)$ eine offene und konvexe Teilmenge des normierten Raumes $\text{Sym}(N)$, und nach [Lemma 3.6](#) (unter Beachtung von [Bemerkung 3.7](#)) gilt

$$\text{span}(K) = \text{span}(\mathbb{P}\text{Sym}(N)) = \text{Sym}(N) = X. \quad \blacksquare$$

3.1.3. Eine Konvexitätsbedingung

Bevor wir nun [Satz 3.13](#) auf unser [Problem \(3.1\)](#) anwenden, ist noch ein wenig Vorarbeit nötig.

Definition 3.14 (KOFAKTOR/ADJUNKTE)

- (i) Für $A \in \mathbb{M}^{N \times N}$ nennen wir $\text{Cof } A := (\tilde{a}_{ij})_{i,j=1,\dots,N} \in \mathbb{M}^{N \times N}$ den „Kofaktor von A “, wobei

$$\tilde{a}_{ij} := (-1)^{i+j} \cdot \Delta_{ij} .$$

Hierbei bezeichne Δ_{ij} die Determinante derjenigen Untermatrix, die durch Streichen der i . Zeile und j . Spalte entsteht.

- (ii) Für $A \in \mathbb{M}^{N \times N}$ nennen wir $\text{Adj } A := (\text{Cof } A)^T$ die „Adjunkte von A “.

Bemerkung 3.15

Für $A \in \text{Gl}(N, \mathbb{R})$ ($:=$ die Menge aller reellen invertierbaren $N \times N$ -Matrizen) folgt aus der CRAMERSchen Regel:

$$\text{Cof } A = \det A \cdot A^{-T} \quad , \quad \text{und damit} \quad \text{Adj } A = \det A \cdot A^{-1} . \quad (3.10)$$

Wir wollen nun auf $\mathbb{M}^{3 \times 3}$ ein Skalarprodukt definieren, indem wir für $A, B \in \mathbb{M}^{3 \times 3}$ setzen:

Definition 3.16 (SKALARPRODUKT AUF $\mathbb{M}^{3 \times 3}$)

$$\langle A, B \rangle_{\mathbb{M}^{3 \times 3}} := \langle A, B \rangle_{3 \times 3} := \text{tr}(AB^T) ,$$

wobei tr die Spur einer Matrix bezeichne.

Lemma 3.17 $\langle \cdot, \cdot \rangle_{3 \times 3}$ ist ein Skalarprodukt, und für $A, B, C \in \mathbb{M}^{3 \times 3}$ gilt:

- (i) $\text{tr } A = \langle A, \mathbf{1} \rangle_{3 \times 3}$
- (ii) $\langle A, BC \rangle_{3 \times 3} = \langle AC^T, B \rangle_{3 \times 3} = \langle B^T A, C \rangle_{3 \times 3} = \langle C, B^T A \rangle_{3 \times 3}$
- (iii) $\langle A, B \rangle_{3 \times 3} = \langle A^T, B^T \rangle_{3 \times 3}$

Beweis [11, S. 27ff.] ■

Lemma 3.18

Für $g, \tilde{g} \in \mathcal{F}(\Omega, Y)$ und einem beliebigen Skalarprodukt $\langle \cdot, \cdot \rangle$ auf $Y \times Y$ gilt:
Sind g, \tilde{g} in a differenzierbar, so ist auch

$$\begin{aligned} p &: \Omega \longrightarrow \mathbb{R} \\ x &\longmapsto p(x) := \langle g(x), \tilde{g}(x) \rangle \end{aligned}$$

in a differenzierbar mit

$$Dp(a).h = \langle Dg(a).h, \tilde{g}(a) \rangle + \langle g(a), D\tilde{g}(a).h \rangle \quad \forall h \in X.$$

Beweis

Wir wählen $\rho > 0$ derart, dass für g, \tilde{g} und $h \in B_\rho(0)$ nach Voraussetzung

$$\begin{aligned} g(a+h) &= g(a) + Dg(a).h + r(h)\|h\|, \\ \tilde{g}(a+h) &= \tilde{g}(a) + D\tilde{g}(a).h + \tilde{r}(h)\|h\|. \end{aligned}$$

Für diese h folgt dann:

$$\begin{aligned} p(a+h) &= \langle g(a) + Dg(a).h + r(h)\|h\|, \tilde{g}(a) + D\tilde{g}(a).h + \tilde{r}(h)\|h\| \rangle \\ &= p(a) + \underbrace{\langle Dg(a).h, \tilde{g}(a) \rangle + \langle g(a), D\tilde{g}(a).h \rangle}_{=: Th} + r_p(h)\|h\| \quad \text{mit} \end{aligned}$$

$$\begin{aligned} r_p(h) &:= \|h\|^{-1} \langle Dg(a).h, D\tilde{g}(a).h \rangle + \langle g(a+h), \tilde{r}(h) \rangle + \langle r(h), \tilde{g}(a) \rangle \\ &\quad + \langle r(h), D\tilde{g}(a).h \rangle \quad (\text{Trivialfall } h = 0 \text{ ausgeschlossen}). \end{aligned}$$

Offensichtlich ist T linear und auch stetig mit

$$\|T\|_{\mathcal{L}(X, \mathbb{R})} \leq \|Dg(a)\|_{\mathcal{L}(X, Y)} \|\tilde{g}(a)\|_Y + \|D\tilde{g}(a)\|_{\mathcal{L}(X, Y)} \|g(a)\|_Y,$$

was man mit Hilfe von (CSU) einsieht.

Ebenfalls (CSU) sowie die Eigenschaft, dass r, \tilde{r} Differentiationsreste sind, zeigen, dass r_p „Rest“ ist. ■

Definition 3.19 (TOPOLOGISCHER ISOMORPHISMUS)

Eine Abbildung $C \in \mathcal{L}(X, Y)$, welche bijektiv und deren Umkehrabbildung C^{-1} stetig ist, bezeichnen wir als „topologischen Isomorphismus“. Die Menge aller solcher Abbildungen bezeichnen wir mit $\mathcal{G}(X, Y)$.

Seien nun X, Y vollständig, also BANACHräume.

Lemma 3.20

(i) $\mathcal{G}(X, Y)$ ist eine offene Teilmenge von $\mathcal{L}(X, Y)$. Genauer:

$$C \in \mathcal{G}(X, Y) \quad \Rightarrow \quad B_{\|C^{-1}\|^{-1}}(C) \subset \mathcal{G}(X, Y)$$

(ii) Die „inverse Abbildung“ (oder auch „Kehrwertsfunktion“)

$$k : \mathcal{G}(X, Y) \longrightarrow \mathcal{L}(X, Y)$$

$$C \longmapsto C^{-1}$$

ist differenzierbar, und es gilt

$$Dk(C).H = -C^{-1}HC^{-1}.$$

Zum Beweis dieses Lemmas benötigen wir eine klassische Aussage über die sogenannte NEUMANNsche Reihe (in der wir $T^0 := \text{id}_X$ (Identität) für $T \in \mathcal{L}(X) := \mathcal{L}(X, X)$ setzen, und $T^n = T \circ T \circ \dots \circ T$ (n -mal, $n \in \mathbb{N}$) im Sinne der Hintereinanderausführung verstehen):

Satz 3.21 (NEUMANNSCHE REIHE)

Sei $T \in \mathcal{L}(X)$ mit

$$\limsup_{n \rightarrow \infty} \|T^n\|^{1/n} < 1$$

(hinreichend hierfür ist z. B. $\|T\| < 1$, da $\|T^n\|^{1/n} \leq \|T\| \forall n \in \mathbb{N}$). Dann konvergiert die Reihe

$$G := \sum_{n=0}^{\infty} T^n \quad \text{in } \mathcal{L}(X)$$

und es gilt $(\text{id} - T)G = \text{id} = G(\text{id} - T)$, d.h. $\text{id} - T \in \mathcal{G}(X, X)$ und $(\text{id} - T)^{-1} = G$. Insbesondere ist die Gleichung

$$x - Tx = z$$

für alle $z \in X$ eindeutig lösbar und es gilt $x = \sum_{n=0}^{\infty} T^n z$.

Beweis Einen Beweis findet man z. B. in [1, S. 146]. Die letzte Aussage gilt, da Konvergenz in $\mathcal{L}(X)$ punktweise Konvergenz nach sich zieht.

Denn mit $G_N := \sum_{n=0}^N T^n$ ist $G_N z = \sum_{n=0}^N T^n z$ für $z \in X$ beliebig und

$$\|Gz - G_N z\|_X = \|(G - G_N)z\|_X \leq \|G - G_N\|_{\mathcal{L}(X)} \|z\|_X \xrightarrow{N \rightarrow \infty} 0. \quad \blacksquare$$

Beweis des Lemmas

„zu (i)“

Für $C \in \mathcal{G}(X, Y)$ und $H \in \mathcal{L}(X, Y)$ mit $\|H\| < \|C^{-1}\|^{-1}$ ist $T := -C^{-1}H \in \mathcal{L}(X)$ mit $\|T\| < 1$. Nach der NEUMANNschen Reihe ist also $\text{id} - T \in \mathcal{G}(X, Y)$, und damit auch

$$C + H = C(\text{id} - (-C^{-1}H)) = C(\text{id} - T),$$

was zu zeigen war.

„zu (ii)“

Nach dem eben gezeigten ist $\mathcal{G}(X, Y)$ eine offene Menge, und für $H \in B_\rho(0)$ mit $\rho := \|C^{-1}\|^{-1}$ gilt:

$$\begin{aligned} k(C + H) &= (C(\text{id} - T))^{-1} = (\text{id} - T)^{-1}C^{-1} = \left(\sum_{n=0}^{\infty} T^n \right) C^{-1} = \sum_{n=0}^{\infty} T^n C^{-1} \\ &= C^{-1} + TC^{-1} + \sum_{n=2}^{\infty} T^n C^{-1} \\ &= k(C) - C^{-1}HC^{-1} + \underbrace{\left(\|H\|^{-1} \sum_{n=2}^{\infty} T^n C^{-1} \right)}_{=: r(H)} \|H\|, \end{aligned}$$

wobei wir im letzten Schritt wieder den Trivialfall $H = 0$ ausschließen.

$H \mapsto -C^{-1}HC^{-1}$ ist linear und stetig und $r(H)$ „Rest“, denn

$$\|r(H)\| \leq \|H\|^{-1} \|C^{-1}\| \cdot \frac{\|T\|^2}{1 - \|T\|} \leq \|C^{-1}\|^3 \cdot \frac{\|H\|}{1 - \|C^{-1}H\|} \xrightarrow{H \rightarrow 0} 0. \quad \blacksquare$$

Lemma 3.22 (DIFFERENTIALE)

Sei $g \in C^n(X, Y)$. Wenn stetige j -Linearformen

$$B_j : X \longrightarrow \mathcal{L}^j(X, Y), \quad j = 1, \dots, n, \quad 7$$

mit der Eigenschaft $\forall x, h \in X$

$$f(x+h) = f(x) + B_1(x) \cdot h + B_2(x) \cdot (h, h) + B_3(x) \cdot (h, h, h) + \dots + B_n(x) \cdot \underbrace{(h, \dots, h)}_{n\text{-fach}}$$

existieren, so gilt

$$D^j g(x) \cdot \underbrace{(h, \dots, h)}_{j\text{-fach}} = j! \cdot B_j(x) \cdot \underbrace{(h, \dots, h)}_{j\text{-fach}} \quad \text{für } j = 1, \dots, n.$$

Beweis [11, S. 101] ■

⁷ Allgemein spricht man von „Multilinearformen“. Speziell bedeutet $B_j(x) \in \mathcal{L}^j(X, Y)$ für ein $x \in X$, dass $B_j(x) \in \mathcal{F}(X^j, Y)$ stetig sowie in allen j Argumenten linear ist. In [Unterabschnitt 3.1.2](#) hätten wir also $D^n g(a) \in \mathcal{L}^n(X, Y)$ formulieren können.

Lemma 3.23 (ENTWICKLUNG DER DETERMINANTE FÜR $N = 3$)

$$\forall A, H \in \mathbb{M}^{3 \times 3} \quad (3.11)$$

$$\det(A + H) = \det A + \underbrace{\langle \text{Adj } A^T, H \rangle_{3 \times 3}}_{= \text{Cof } A} + \langle A, \underbrace{\text{Adj } H^T}_{= \text{Cof } H} \rangle_{3 \times 3} + \det H$$

Beweis

Wir setzen $A = (a_{ij})_{i,j=1,\dots,3}$, $H = (h_{ij})_{i,j=1,\dots,3}$ allgemein an, und bestimmen

$$\begin{aligned} \det(A + H) &= \begin{vmatrix} a_{11} + h_{11} & a_{12} + h_{12} & a_{13} + h_{13} \\ a_{21} + h_{21} & a_{22} + h_{22} & a_{23} + h_{23} \\ a_{31} + h_{31} & a_{32} + h_{32} & a_{33} + h_{33} \end{vmatrix} \\ &= (a_{11} + h_{11})(a_{22} + h_{22})(a_{33} + h_{33}) \\ &\quad - (a_{13} + h_{13})(a_{22} + h_{22})(a_{31} + h_{31}) \\ &\quad + (a_{12} + h_{12})(a_{23} + h_{23})(a_{31} + h_{31}) \\ &\quad - (a_{11} + h_{11})(a_{23} + h_{23})(a_{32} + h_{32}) \\ &\quad + (a_{13} + h_{13})(a_{21} + h_{21})(a_{32} + h_{32}) \\ &\quad - (a_{12} + h_{12})(a_{21} + h_{21})(a_{33} + h_{33}) \\ &= \det A + \det H \\ &\quad + h_{11}(a_{22}a_{33} - a_{32}a_{23}) + a_{11}(h_{22}h_{33} - h_{32}h_{23}) \\ &\quad - h_{12}(a_{21}a_{33} - a_{31}a_{23}) - a_{12}(h_{21}h_{33} - h_{31}h_{23}) \\ &\quad + h_{13}(a_{21}a_{32} - a_{31}a_{22}) + a_{13}(h_{21}h_{32} - h_{31}a_{22}) \\ &\quad - h_{21}(a_{12}a_{33} - a_{32}a_{13}) - a_{21}(h_{12}h_{33} - h_{32}h_{13}) \\ &\quad + h_{22}(a_{11}a_{33} - a_{31}a_{13}) + a_{22}(h_{11}h_{33} - h_{31}h_{13}) \\ &\quad - h_{23}(a_{11}a_{32} - a_{31}a_{12}) - a_{23}(h_{11}h_{32} - h_{31}a_{12}) \\ &\quad + h_{31}(a_{12}a_{23} - a_{22}a_{13}) + a_{31}(h_{12}h_{23} - h_{22}h_{13}) \\ &\quad - h_{32}(a_{11}a_{23} - a_{21}a_{13}) - a_{32}(h_{11}h_{23} - h_{21}h_{13}) \\ &\quad + h_{33}(a_{11}a_{22} - a_{21}a_{12}) + a_{33}(h_{11}h_{22} - h_{21}a_{12}) \\ &= \det A + \det H + \text{tr}(H \cdot \text{Adj } A) + \text{tr}(A \cdot \text{Adj } H) \\ &= \det A + \langle H, \text{Adj } A^T \rangle_{3 \times 3} + \langle A, \text{Adj } H^T \rangle_{3 \times 3} + \det H \\ &= \det A + \langle \text{Adj } A^T, H \rangle_{3 \times 3} + \langle A, \text{Adj } H^T \rangle_{3 \times 3} + \det H. \quad \blacksquare \end{aligned}$$

Korollar 3.24 (FORMALE ABLEITUNG DER DETERMINANTENFUNKTION FÜR $N = 3$)

Die Determinantenfunktion ist unendlich oft differenzierbar, und es gilt

$$\forall A, H \in \mathbb{M}^{3 \times 3}$$

$$D^j \det(A) \cdot \underbrace{(H, \dots, H)}_{j\text{-fach}} = \begin{cases} \langle \text{Adj } A^T, H \rangle_{3 \times 3} & \text{für } j = 1, \\ 2 \cdot \langle A, \text{Adj } H^T \rangle_{3 \times 3} & \text{für } j = 2, \\ 6 \cdot \det H & \text{für } j = 3, \\ 0 & \text{für } j \geq 4. \end{cases} \quad (3.12)$$

Beweis

Lemma 3.22 angewandt auf **Lemma 3.23** (\det ist ein Polynom dritten Grades, also unendlich oft differenzierbar). ■

Satz 3.25 (EINE KONVEXITÄTSBEDINGUNG FÜR DAS **PROBLEM (3.1)**)

Sei $f \in C^2(\mathbb{R}_+, \mathbb{R})$. Dann ist die Funktion

$$f \circ \det : \mathbb{P}\text{Sym}(3) \longrightarrow \mathbb{R}, \quad C \longmapsto f(\det C)$$

konvex genau dann, wenn

$$\boxed{\begin{aligned} &\forall C \in \mathbb{P}\text{Sym}(3) \quad \forall H \in \text{Sym}(3) : \\ &\left[f''(\det C) \det C + f'(\det C) \right] \langle C^{-1}, H \rangle^2 - f'(\det C) \langle HC^{-1}, C^{-1}H \rangle \geq 0. \end{aligned}} \quad (3.13)$$

Beweis

Wir möchten natürlich **Satz 3.13** auf

$$g := f \circ \det$$

anwenden. Vorstehende Resultate bezüglich der Differenzierbarkeit der Determinantenfunktion gelten natürlich insbesondere für die Einschränkung auf den (endlich-dimensionalen) normierten (Untervektor-)Raum $\text{Sym}(3)$. Damit gilt für die in diesem Raum nach **Lemma 3.4** offene Menge

$$\mathbb{P}\text{Sym}(3) \subseteq \mathcal{G}(\text{Sym}(3), \mathbb{R}) \subseteq \text{Gl}(3, \mathbb{R}),$$

dass $\det \in C^\infty(\mathbb{P}\text{Sym}(3), \mathbb{R})$, und deshalb $g \in C^2(\mathbb{P}\text{Sym}(3), \mathbb{R})$.

Wir können somit die Adjunkte gemäß **(3.10)** umschreiben und es genügt wegen $\det C > 0$ zu zeigen, dass

$$D^2g(C).(H, H) = \det C \cdot \left\{ \left[f''(\det C) \cdot \det C + f'(\det C) \right] \cdot \langle C^{-1}, H \rangle_{3 \times 3}^2 - f'(\det C) \cdot \langle HC^{-1}, C^{-1}H \rangle_{3 \times 3} \right\}$$

für $C \in \mathbb{P}\text{Sym}(3)$ und $H \in \text{Sym}(3)$. Hierbei werden wir Produkt- und Kettenregel sowie den Großteil der Resultate dieses Abschnitts benutzen, diese aber aus Gründen der Übersichtlichkeit nicht zitieren:

$$\begin{aligned} Dg(C).H &= Df(\det C) D \det(C).H \\ &= f'(\det C) \cdot \langle \text{Adj } C^T, H \rangle_{3 \times 3} \\ &= f'(\det C) \cdot \det C \cdot \langle C^{-1}, H \rangle_{3 \times 3} \end{aligned}$$

und damit

$$\begin{aligned}
D^2g(C).(H, H) &= f''(\det C) \cdot (\det C)^2 \cdot \langle C^{-1}, H \rangle_{3 \times 3}^2 \\
&\quad + f'(\det C) \cdot \det C \cdot \langle C^{-1}, H \rangle_{3 \times 3}^2 \\
&\quad + f'(\det C) \cdot \det C \cdot \langle -C^{-1}HC^{-1}, H \rangle_{3 \times 3} \\
&= f''(\det C) \cdot \det C^2 \cdot \langle C^{-1}, H \rangle_{3 \times 3}^2 \\
&\quad + f'(\det C) \cdot \det C \cdot \langle C^{-1}, H \rangle_{3 \times 3}^2 \\
&\quad - f'(\det C) \cdot \det C \cdot \langle HC^{-1}, C^{-1}H \rangle_{3 \times 3} . \quad \blacksquare
\end{aligned}$$

3.1.4. Die Differentialungleichung

Ziel ist es nun, **Bedingung (3.13)** aus **Satz 3.25** in eine für die Funktion f „handhabbarere“ Form zu bringen. Mit wenig Mühe sieht man ein:

Lemma 3.26

$$f'(s) \leq 0 \quad \forall s \in \mathbb{R}_+ \quad (3.14)$$

ist notwendig für **(3.13)**.

Beweis

Denn angenommen es gibt ein $s \in \mathbb{R}_+$ mit $f'(s) > 0$, so wähle

$$C = \text{diag}(1, 1, s) \in \mathbb{P}\text{Sym}(3) \quad \text{und} \quad H = \text{diag}(1, -1, 0) \in \text{Sym}(3) .$$

Damit ist $\det C = s$, $\langle C^{-1}, H \rangle = 1 - 1 = 0$ und

$$\langle HC^{-1}, C^{-1}H \rangle = \langle \text{diag}(1, -1, 0), \text{diag}(1, -1, 0) \rangle = 2 .$$

Eingesetzt in **(3.13)** ergibt sich:

$$-2f'(s) \geq 0 . \quad \text{!} \quad \blacksquare$$

Also muss f für **(3.13)** monoton fallend sein. Hierbei haben wir Folgendes ausgenutzt:

Wie wir in dem folgenden Lemma sehen werden, können wir die symmetrisch positiv definite Matrix C in **(3.13)** durch eine Diagonalmatrix D (natürlich mit positiven Einträgen auf der Diagonalen) ersetzen. Eine kleine Rechnung zeigt,⁸ dass daraus

$$\langle HC^{-1}, C^{-1}H \rangle > 0 \quad \forall C \in \mathbb{P}\text{Sym}(3) \quad \forall H \in \text{Sym}(3) \setminus \{0\}$$

folgt, wohingegen im Allgemeinen nur

$$\langle C^{-1}, H \rangle^2 \geq 0 \quad \forall C \in \mathbb{P}\text{Sym}(3) \quad \forall H \in \text{Sym}(3) \setminus \{0\}$$

wegen $\langle C^{-1}, H \rangle^2 = (\text{tr}(C^{-1}H))^2$ gilt.

Zusammen mit dem Resultat aus [Lemma 3.26](#) ergibt sich daraus aber auch, dass

$$f''(s) \cdot s + f'(s) \geq 0 \quad \forall s \in \mathbb{R}_+$$

hinreichend für [\(3.13\)](#) ist. Es stellt sich jedoch heraus, dass dies **keine notwendige Bedingung** für [\(3.13\)](#) ist. Bevor wir dies näher erläutern, formulieren wir [\(3.13\)](#) noch etwas um:

Lemma 3.27

Es gilt [\(3.13\)](#) genau dann, wenn

$$\begin{aligned} & \forall H \in \text{Sym}(3) \quad \forall D^{-1} = \text{diag}(d_1, d_2, d_3), \\ & \text{wobei } d_1, d_2, d_3 \in \mathbb{R}_+ \quad \wedge \quad s^{-1} := \det D^{-1} = d_1 d_2 d_3 \in \mathbb{R}_+ : \\ & \left(f''(s) + \frac{f'(s)}{s} \right) \langle D^{-1}, H \rangle_{3 \times 3}^2 - \frac{f'(s)}{s} \langle D^{-1}H, HD^{-1} \rangle_{3 \times 3} \geq 0. \end{aligned} \tag{3.15}$$

Beweis

Wir betrachten in [\(3.13\)](#) ein beliebiges $C \in \mathbb{P}\text{Sym}(3)$. Nach [Lemma 3.5](#) gibt es eine orthogonale Matrix Q derart, dass

$$C = QDQ^T \quad \Leftrightarrow \quad C^{-1} = QD^{-1}Q^T$$

mit $D = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$ und positiven λ_i , da C positiv definit ist. $d_i := \lambda_i^{-1}$ und $s := \det C = \det D = \prod_{i=1}^3 \lambda_i$ begründet einen Teil der Aussage.

Mit Hilfe von [Lemma 3.17](#) folgt

$$\begin{aligned} \langle C^{-1}, H \rangle_{3 \times 3} &= \langle QD^{-1}Q^T, H \rangle_{3 \times 3} \\ &= \langle QD^{-1}, HQ \rangle_{3 \times 3} \\ &= \langle D^{-1}, \underbrace{Q^T H Q}_{=: \tilde{H} \in \text{Sym}(3)} \rangle_{3 \times 3} \end{aligned}$$

und in dieser Schreibweise entsprechend

$$\begin{aligned} \langle HC^{-1}, C^{-1}H \rangle_{3 \times 3} &= \langle HQD^{-1}Q^T, QD^{-1}Q^T H \rangle_{3 \times 3} \\ &= \langle \tilde{H}D^{-1}, D^{-1}\tilde{H} \rangle_{3 \times 3} \\ &= \langle D^{-1}\tilde{H}, \tilde{H}D^{-1} \rangle_{3 \times 3}. \end{aligned}$$

Variiert nun H in ganz $\text{Sym}(3)$, so auch \tilde{H} und vice versa. Anstelle von \tilde{H} schreiben wir deshalb wieder H , teilen [\(3.13\)](#) durch $s \in \mathbb{R}_+$ und erhalten die behauptete Darstellung. ■

⁸ Diese werden wir im Beweis der zweiten Aussage von [Lemma 3.28](#) (Seite 41) auch noch führen.

Zurück zur Frage, ob $f''(s) + \frac{f'(s)}{s} \geq 0 \quad \forall s \in \mathbb{R}_+$ notwendig für (3.13) ist:

Es ist $\langle D^{-1}, H \rangle^2 \stackrel{(\text{CSU})}{\leq} \|D^{-1}\|^2 \|H\|^2$.⁹ Möchte man nun die Annahme

$$\exists s \in \mathbb{R}_+ : f''(s) + \frac{f'(s)}{s} < 0$$

zum Widerspruch führen, sollte man den Vorfaktor $\langle D^{-1}, H \rangle^2$ möglichst stark gewichten. In (CSU) steht aber genau dann ein Gleichheitszeichen, wenn linkes und rechtes Argument linear abhängig sind, d.h.

$$H = k \cdot D^{-1} \quad \text{mit } k \in \mathbb{R} \setminus \{0\}$$

gilt.

Gleichzeitig gilt für den Vorfaktor $\langle D^{-1}H, HD^{-1} \rangle$ des zweiten, nicht-negativen Summanden nach [11, S. 31] die Abschätzung

$$\langle D^{-1}H, HD^{-1} \rangle \geq \lambda_{\min}^2(D^{-1}) \cdot \|H\|^2$$

nach unten. Wollen wir auch hier das Gleichheitszeichen erzwingen, so müssen wir alle Eigenwerte identisch wählen, was man z. B. anhand des Beweises der zweiten Aussage des folgenden Lemmas erkennt. Die einzige Möglichkeit hierzu ist gemäß (3.15)

$$d_1 = d_2 = d_3 = s^{-1/3}.$$

Setzen wir dies in (3.15) ein, so erhalten wir

$$\begin{aligned} & \left(f''(s) + \frac{f'(s)}{s} \right) \langle D^{-1}, H \rangle^2 - \frac{f'(s)}{s} \langle D^{-1}H, HD^{-1} \rangle \\ = & k^2 \left(\left(f''(s) + \frac{f'(s)}{s} \right) \langle D^{-1}, D^{-1} \rangle^2 - \frac{f'(s)}{s} \langle (D^{-1})^2, (D^{-1})^2 \rangle \right) \\ = & k^2 \left(\left(f''(s) + \frac{f'(s)}{s} \right) \text{tr}^2(D^{-1})^2 - \frac{f'(s)}{s} \cdot \text{tr}(D^{-1})^4 \right) \\ = & k^2 \left(\left(f''(s) + \frac{f'(s)}{s} \right) (3s^{-2/3})^2 - \frac{f'(s)}{s} \cdot 3s^{-4/3} \right) \\ = & 3k^2 s^{-4/3} \left(3f''(s) + 2\frac{f'(s)}{s} \right) \stackrel{!}{<} 0 \quad \Leftrightarrow \quad f''(s) < -\frac{2}{3} \cdot \frac{f'(s)}{s}. \end{aligned}$$

Dies ist eine stärkere Forderung an f'' , da wir bereits wissen, dass f monoton fallend sein muss. Wir können also im Allgemeinen an dieser Stelle keinen Widerspruch erwarten,

⁹ Ab hier werden wir zwecks besserer Lesbarkeit nur noch $\langle \cdot, \cdot \rangle$ für das Matrixskalarprodukt schreiben. Die Matrixnorm ist die von dem Matrixskalarprodukt induzierte, d.h. für $A \in \mathbb{M}^{3 \times 3}$ ist

$$\|A\|_{\mathbb{M}^{3 \times 3}} := \|A\|_{3 \times 3} := \|A\| := \sqrt{\langle A, A \rangle_{3 \times 3}} = \sqrt{\text{tr}(AA^T)} = \sqrt{\sum_{i,j=1}^3 a_{ij}^2}.$$

Dies entspricht gerade der Operator- bzw. Matrixnorm in (3.3), falls man den \mathbb{R}^3 mit der euklidischen, auch genannt 2-Norm ausstattet.

wenn wir bloß $f''(s) < -f'(s)/s$ für ein positives s annehmen. Umgekehrt sehen wir aber natürlich, dass

$$f''(s) \geq -\frac{2}{3} \cdot \frac{f'(s)}{s} \quad \forall s \in \mathbb{R}_+ \quad (3.16)$$

eine **notwendige Bedingung** für (3.15) und damit (3.13) darstellt. Es stellt sich sogar heraus, dass dies zusammen mit (3.14) ($f' \leq 0$) eine **hinreichende Bedingung** ist, wir also eine zu unserem **Problem (3.1)** äquivalente Forderung an die Funktion f in Form einer (zweier) Differentialungleichung(en) gefunden haben.

Zum Nachweis benötigen wir noch ein weiteres Lemma. Hierbei bezeichne für eine Matrix A , $\text{diag } A$ diejenige Matrix, die aus A durch Nullsetzen der Nicht-Diagonalelemente hervorgeht, und $\text{diag}_{\mathbb{M}^{N \times N}}$ die Menge aller $N \times N$ -Diagonalmatrizen.

Lemma 3.28

Für alle $P \in \text{diag}_{\mathbb{M}^{3 \times 3}}$ und $A \in \mathbb{M}^{3 \times 3}$ gilt

$$\langle P, A \rangle = \langle P, \text{diag } A \rangle =: \sigma(P, A), \quad (3.17)$$

$$\langle PA, AP \rangle \geq \langle P \text{diag } A, \text{diag } A P \rangle =: \tilde{\sigma}(P, A) \quad (3.18)$$

für den Fall, dass alle Diagonaleinträge von P nicht-negativ sind, sowie

$$\sigma^2(P, A) \leq 3 \cdot \tilde{\sigma}(P, A). \quad (3.19)$$

Beweis

Wir setzen allgemein $P = \begin{pmatrix} p_1 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & p_3 \end{pmatrix}$, $A = \begin{pmatrix} a_1 & a_4 & a_5 \\ a_6 & a_2 & a_7 \\ a_8 & a_9 & a_3 \end{pmatrix}$ an, und berechnen

$$\langle P, A \rangle = \sum_{i=1}^3 p_i a_i = \langle P, \text{diag } A \rangle. \quad \Rightarrow (3.17)$$

Und wegen

$$PA = \begin{pmatrix} p_1 a_1 & p_1 a_4 & p_1 a_5 \\ p_2 a_6 & p_2 a_2 & p_2 a_7 \\ p_3 a_8 & p_3 a_9 & p_3 a_3 \end{pmatrix}, \quad PA^T = \begin{pmatrix} p_1 a_1 & p_1 a_6 & p_1 a_8 \\ p_2 a_4 & p_2 a_2 & p_2 a_7 \\ p_3 a_5 & p_3 a_7 & p_3 a_3 \end{pmatrix}$$

folgt

$$\begin{aligned} \langle PA, AP \rangle &= \text{tr} \left(PA (AP)^T \right) = \text{tr} \left(P A P A^T \right) \\ &= \underbrace{p_1^2 a_1^2 + p_2^2 a_2^2 + p_3^2 a_3^2}_{= \langle P \text{diag } A, \text{diag } A P \rangle} + \underbrace{p_1 p_2 (a_4^2 + a_6^2) + p_1 p_3 (a_5^2 + a_8^2) + p_2 p_3 (a_7^2 + a_9^2)}_{\geq 0}, \end{aligned}$$

also (3.18).

„zu (3.19)“

Wie wir gesehen haben, ist

$$\tilde{\sigma}(P, A) = \sum_{i=1}^3 (p_i a_i)^2 \quad , \quad \sigma^2(P, A) = \left(\sum_{i=1}^3 p_i a_i \right)^2 .$$

Falls $\tilde{\sigma}(P, A)$ verschwindet, so auch $\sigma^2(P, A)$ und die Ungleichung ist erfüllt. Wir können also oBdA $\tilde{\sigma} \neq 0$ annehmen und müssen in diesem Fall zeigen, dass $\sigma^2/\tilde{\sigma} \leq 3$. Hierzu substituieren wir $x := p_1 a_1$, $y := p_2 a_2$ und $z := p_3 a_3$ und untersuchen stattdessen wegen

$$\mathbb{R}^3 = \left\{ (p_1 a_1, p_2 a_2, p_3 a_3)^T \mid p_i, a_i \in \mathbb{R} \text{ für } i = 1, \dots, 3 \right\}$$

die Funktion

$$g : \mathbb{R}^3 \setminus \{0\} \longrightarrow \mathbb{R}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto g(x, y, z) := \frac{(x + y + z)^2}{x^2 + y^2 + z^2} .$$

Hierzu schreiben wir g in Kugelkoordinaten

$$\mathbb{R}^3 \setminus \{0\} \ni \begin{pmatrix} x \\ y \\ z \end{pmatrix} \hat{=} r \cdot \begin{pmatrix} \cos \vartheta \cos \varphi \\ \cos \vartheta \sin \varphi \\ \sin \vartheta \end{pmatrix} \quad \text{mit } r \in \mathbb{R}_+, \vartheta \in [-\pi/2, \pi/2], \varphi \in [0, 2\pi]$$

um, wobei dieser Übergang nicht injektiv ist, was uns aber nicht zu stören braucht. Interessant ist, dass wir surjektiv von $\mathbb{R}^3 \setminus \{0\}$ nach $\mathbb{R}^3 \setminus \{0\}$ abbilden.

$$\frac{(x + y + z)^2}{x^2 + y^2 + z^2} \rightsquigarrow (\cos \vartheta (\cos \varphi + \sin \varphi) + \sin \vartheta)^2 .$$

Die Funktionswerte von g sind folglich unabhängig von der Länge r der Argumente. Wir vereinfachen noch weiter und diskutieren die Funktion

$$\tilde{g} : [-\pi/2, \pi/2] \times [0, 2\pi] \longrightarrow \mathbb{R}$$

$$\begin{pmatrix} \vartheta \\ \varphi \end{pmatrix} \longmapsto \tilde{g}(\vartheta, \varphi) := \cos \vartheta (\cos \varphi + \sin \varphi) + \sin \vartheta .$$

Da \tilde{g} stetig und der Definitionsbereich kompakt ist, besitzt \tilde{g} globale Extrema. Kandidaten sind neben den Randpunkten des Definitionsbereiches die kritischen Punkte im Innern $(-\pi/2, \pi/2) \times (0, 2\pi)$:

$$D\tilde{g}(\vartheta, \varphi) = \begin{pmatrix} -\sin \vartheta (\cos \varphi + \sin \varphi) + \cos \vartheta & \cos \vartheta (\cos \varphi - \sin \varphi) \end{pmatrix} \stackrel{!}{=} 0$$

ergibt in der zweiten Komponente $\cos \varphi = \sin \varphi \Leftrightarrow \varphi \in \{\pi/4, 5\pi/4\}$. Eingesetzt in die erste Komponente erhält man wegen $\cos \pi/4 = -\cos 5\pi/4 = 1/\sqrt{2}$

$$\cos \vartheta = \pm \sqrt{2} \sin \vartheta \quad \Leftrightarrow \quad \vartheta = \arctan \left(\pm 1/\sqrt{2} \right) \in (-\pi/2, \pi/2) ,$$

und somit die kritischen Punkte

$$\left(\arctan\left(1/\sqrt{2}\right), \pi/4\right) \text{ und } \left(\arctan\left(-1/\sqrt{2}\right), 5\pi/4\right).$$

Um den Funktionswert an diesen Stellen berechnen zu können, benötigen wir noch die Gleichheit

$$\sin y = \frac{\tan y}{\sqrt{1 + \tan^2 y}} \quad \text{für } y \in (-\pi/2, \pi/2),$$

welche man unmittelbar einsieht. Die Wahl $y := \arctan x \in (-\pi/2, \pi/2)$ für $x \in \mathbb{R}$ ergibt deshalb die Beziehung

$$\sin(\arctan x) = \frac{x}{\sqrt{1 + x^2}} \quad \forall x \in \mathbb{R}.$$

Analog gewinnt man die Beziehung

$$\cos(\arctan x) = \frac{1}{\sqrt{1 + x^2}} \quad \forall x \in \mathbb{R}$$

aus der Gleichheit

$$\cos y = \frac{1}{\sqrt{1 + \tan^2 y}} \quad \text{für } y \in (-\pi/2, \pi/2).$$

Wir bestimmen also

$$\tilde{g}\left(\arctan \frac{1}{\sqrt{2}}, \frac{\pi}{4}\right) = -\tilde{g}\left(\arctan \frac{-1}{\sqrt{2}}, \frac{5\pi}{4}\right) = \frac{\sqrt{2}}{\sqrt{3}} \cdot \sqrt{2} + \frac{1}{\sqrt{2}} \cdot \frac{\sqrt{2}}{\sqrt{3}} = \sqrt{3}.$$

Für die Randwerte gilt zum Einen

$$|\tilde{g}(\pm\pi/2, \varphi)| = |\pm 1| < \sqrt{3},$$

und zum Anderen

$$|\tilde{g}(\vartheta, 0)| = |\tilde{g}(\vartheta, 2\pi)| = |\cos \vartheta + \sin \vartheta| \leq \sqrt{3},$$

denn $(\cos \vartheta + \sin \vartheta)^2 = 1 + 2 \cos \vartheta \sin \vartheta \leq 3$. Dies liefert die gewünschte Behauptung, denn wir haben gezeigt

$$-\sqrt{3} \leq \tilde{g} \leq \sqrt{3} \quad \Leftrightarrow \quad \tilde{g}^2 \leq 3 \quad \Leftrightarrow \quad g \leq 3 \quad \Leftrightarrow \quad \frac{\sigma^2(P, A)}{\tilde{\sigma}(P, A)} \leq 3. \quad \blacksquare$$

Bemerkung 3.29

Natürlich gelten (3.17) und (3.18) nicht nur für 3×3 -Matrizen. Genauso wie in Lemma 3.6 auf Seite 26 haben wir den Beweis aber auf diesen Fall beschränkt, da er uns letztlich nur interessiert.

Auch die Ungleichung (3.19) lässt sich auf andere Dimensionen verallgemeinern: Für alle $P \in \text{diag}_{\mathbb{M}^{N \times N}}$ und $A \in \mathbb{M}^{N \times N}$ gilt

$$\sigma^2(P, A) \leq N \cdot \tilde{\sigma}(P, A). \quad (3.20)$$

Um dies nachzuweisen muss man nur bedenken, dass P und $\text{diag } A$ kommutieren, also

$$\tilde{\sigma}(P, A) = \langle P \text{diag } A, P \text{diag } A \rangle = \|P \text{diag } A\|^2$$

gilt. Mit Hilfe von (CSU) folgt daraus sofort

$$\sigma^2(P, A) = \langle P, \text{diag } A \rangle^2 = \langle P \text{diag } A, \mathbf{1} \rangle^2 \leq \|P \text{diag } A\|^2 \cdot \|\mathbf{1}\|^2 = N \cdot \tilde{\sigma}(P, A).$$

Wir haben also auch einen wesentlich einfacheren und vor allem kürzeren Beweis für Ungleichung (3.19) gefunden.

Theorem 3.30 (EINE DIFFERENTIALUNGLEICHUNG FÜR DAS **PROBLEM (3.1)**)¹⁰

Sei $f \in C^2(\mathbb{R}_+, \mathbb{R})$. Dann ist die Funktion

$$f \circ \det : \mathbb{P}\text{Sym}(3) \longrightarrow \mathbb{R}, \quad C \longmapsto f(\det C)$$

konvex genau dann, wenn

$$\boxed{f''(s) + \frac{2}{3s} \cdot f'(s) \geq 0 \quad \wedge \quad f'(s) \leq 0 \quad \forall s \in \mathbb{R}_+ .} \quad (3.21)$$

Beweis

Dass $f''(s) + 2f'(s)/3s \geq 0$ für alle positiven s notwendig für die Konvexität von $f \circ \det$ ist, haben wir in (3.16) auf Seite 41 bereits gezeigt. Und nach **Lemma 3.26** ist auch die Eigenschaft, dass f monoton abnimmt hierfür notwendig. Es genügt also zu zeigen, dass (3.21) hinreichend für die Konvexität von $f \circ \det$ ist.

Hierzu wollen wir (3.15) nachweisen. Entsprechend seien

$$H \in \text{Sym}(3) \quad \text{und} \quad D^{-1} = \text{diag}(d_1, d_2, d_3)$$

mit $d_1, d_2, d_3 \in \mathbb{R}_+$ und $s := (d_1 d_2 d_3)^{-1} = \det D$ beliebig. $P := D^{-1}$ und $A := H$ erfüllen also die Voraussetzungen aller Aussagen des vorigen **Lemmas 3.28**. In der Schreibweise dieses Lemmas sei oBdA $\sigma(D^{-1}, H) \neq 0$, denn sonst ist (3.15) wegen

¹⁰ Eine andere Vermutung könnte auf die (3.21) entsprechende(n) Differentialungleichung(en)

$$-\frac{f'(s)}{3} \leq s \cdot f''(s) \leq -f'(s) \quad \wedge \quad f'(s) \leq 0 \quad \forall s \in \mathbb{R}_+$$

führen. Dies widerlegen wir im **Anhang** mit Hilfe eines Gegenbeispiels.

der Voraussetzung $f' \leq 0$ in jedem Fall erfüllt.

$$\begin{aligned}
\Rightarrow & \left(f''(s) + \frac{f'(s)}{s} \right) \underbrace{\langle D^{-1}, H \rangle}_{\stackrel{(3.17)}{=} \sigma(D^{-1}, H)}^2 - \underbrace{\frac{f'(s)}{s}}_{\substack{\text{n.V.} \\ \geq 0}} \underbrace{\langle D^{-1}H, HD^{-1} \rangle}_{\stackrel{(3.18)}{\geq} \tilde{\sigma}(D^{-1}, H)} \\
& \geq \left(f''(s) + \frac{f'(s)}{s} \right) \cdot \sigma^2(D^{-1}, H) - \frac{f'(s)}{s} \cdot \tilde{\sigma}(D^{-1}, H) \\
& = \underbrace{\sigma^2(D^{-1}, H)}_{>0} \cdot \left[f''(s) + \frac{f'(s)}{s} \cdot \underbrace{\left(1 - \frac{\tilde{\sigma}(D^{-1}, H)}{\sigma^2(D^{-1}, H)} \right)}_{\substack{\stackrel{(3.19)}{\leq} 3\tilde{\sigma}(D^{-1}, H)} \\ \leq 2/3}} \right] \\
& \geq \sigma^2(D^{-1}, H) \cdot \underbrace{\left(f''(s) + \frac{2}{3} \cdot \frac{f'(s)}{s} \right)}_{\substack{\text{n.V.} \\ \geq 0}} \geq 0. \quad \Rightarrow (3.15) \quad \blacksquare
\end{aligned}$$

3.2. Lösungen der Differentialungleichung(en)

In diesem Abschnitt möchten wir uns ein Bild von denjenigen Funktionen verschaffen, die die Differentialungleichung(en) (3.21) aus **Theorem 3.30** lösen.

3.2.1. Lineare Differentialgleichungen und der Grenzfall

Wir zitieren [13, S. 22ff.] zunächst ein klassisches Resultat:

Satz 3.31 (LINEARE DIFFERENTIALGLEICHUNG)

Seien $\xi \in J(\text{Intervall}) \subseteq \mathbb{R} \ni \eta$ und $g, h \in C(J)$.

Dann hat das „(lineare) Anfangswertproblem“

$$\left\{ \begin{array}{l} Ly := y' + g(x)y = h(x) \\ \wedge \\ y(\xi) = \eta \end{array} \right\} \quad (\text{LAWP})$$

genau eine Lösung

$$y(x) = \exp\left(-\int_{\xi}^x g(t) dt\right) \cdot \left\{ \eta + \int_{\xi}^x h(t) \exp\left(\int_{\xi}^t g(\tau) d\tau\right) dt \right\}. \quad (3.22)$$

Diese existiert in ganz J .

Aufgrund des linearen Operators L bezeichnet

$$y' + g(x)y = h(x)$$

eine „lineare Differentialgleichung“. Im Falle $h \equiv 0$ nennen wir diese „homogen“, andernfalls „inhomogen“. Übertragen auf die Differentialungleichung (3.21) gilt dort

$$J = \mathbb{R}_+ \quad , \quad y = f' \quad , \quad g(x) = \frac{2}{3x} \quad , \quad h \equiv 0 . \quad (3.23)$$

Nun möchten wir aber Lösungen von $Ly \geq 0$ bestimmen. Diese in geschlossener Form anzugeben wird nicht möglich sein. Zunächst aber sehen wir ein, dass zumindest die Lösungen von $Ly = 0$ auch (3.21) lösen - vorausgesetzt für die Lösungen gilt $y \leq 0$. Dazu muss natürlich der „Anfangswert“ $y(\xi) = \eta$ nicht-positiv gewählt werden. Für die dann aber eindeutig bestimmten Lösungen gilt tatsächlich $y \leq 0$. Da $Ly = 0$ einen „Grenzfall“ für (3.21) darstellt, bezeichnen wir die Lösungen mit y_{limit} respektive f'_{limit} . Dies formulieren wir in dem folgenden

Lemma 3.32 (GRENZFALL FÜR (3.21))

Die Lösungsmenge für den Grenzfall

$$f''_{\text{limit}}(s) + \frac{2}{3s} \cdot f'_{\text{limit}}(s) = 0 \quad \wedge \quad f'_{\text{limit}}(s) \leq 0 \quad \forall s \in \mathbb{R}_+$$

von (3.21) ergibt sich zu

$$\begin{aligned} f'_{\text{limit}} &: \mathbb{R}_+ \longrightarrow \mathbb{R} \\ s &\longmapsto c \cdot s^{1/3} + d \end{aligned} \quad \text{mit } c \leq 0, d \in \mathbb{R} . \quad (3.24)$$

Beweis

In der oben unter (3.23) angegebenen Schreibweise ergibt sich die nach Satz 3.31 eindeutig bestimmte und auf ganz \mathbb{R}_+ existente Lösung y_{limit} des (LAWP) mit $\xi > 0 \geq \eta$ gemäß (3.22) zu

$$y_{\text{limit}}(x) = \eta \cdot \exp\left(-\int_{\xi}^x \frac{2}{3t} dt\right) = \eta \cdot \exp\left(-\frac{2}{3} \ln \frac{x}{\xi}\right) = \eta \xi^{2/3} \cdot x^{-2/3} , \quad (3.25)$$

da $h \equiv 0$. Wegen $\eta \xi^{2/3} \leq 0$ folgt $y_{\text{limit}} \leq 0$, d.h. (3.21) ist erfüllt. Integration von $f'_{\text{limit}} = y_{\text{limit}}$ mit der Abkürzung $c := 3\eta \xi^{2/3}$ und der Integrationskonstanten d ergibt die Behauptung. ■

Die additive Konstante d verwundert nicht, denn f tritt in (3.21) nicht explizit auf. Dies war deshalb zu erwarten, da die Eigenschaft der Konvexität unabhängig von einer vertikalen Verschiebung ist.

3.2.2. Ober- und Unterfunktionen

Um uns nun ein Bild von der „Struktur“ der Lösungen von (3.21) zu machen, benötigen wir die Begrifflichkeit der sogenannten Ober- bzw. Unterfunktion:

Definition & Satz 3.33 (OBER- BZW. UNTERFUNKTION)

Mit $D(\text{bel.}) \subseteq \mathbb{R}^2$ sei $F = F(x, y)$ in D erklärt.

Unter der Annahme, dass $\text{graph } v, \text{graph } w \subseteq D$, nennen wir die Funktion v „Unterfunktion“ (bzw. w „Oberfunktion“) bezüglich des Anfangswertproblems“

$$\left\{ \begin{array}{l} y' = F(x, y) \text{ in } J := [\xi, \xi + a] \\ \wedge \\ y(\xi) = \eta, \end{array} \right\} \quad (\text{AWP})$$

wenn sie in J differenzierbar ist und dort

$$v' < F(x, v), \quad v(\xi) \leq \eta \quad (\text{bzw. } w' > F(x, w), \quad w(\xi) \geq \eta)$$

erfüllt. Eine Unterfunktion verläuft unterhalb, eine Oberfunktion oberhalb einer Lösung. Genauer:

Ist y eine Lösung des (AWP), so gilt

$$v(x) < y(x) < w(x) \text{ in } J_0 := (\xi, \xi + a].$$

Falls das (AWP) sowohl eindeutig lösbar, als auch F im „Rechteck“

$$R := J \times [\eta - \varepsilon_1, \eta + \varepsilon_2]$$

stetig ist, so ist sogar (natürlich unter denselben Voraussetzungen an die Differenzierbarkeit und die Graphen von v und w)

$$v' \leq F(x, v) \text{ in } J, \quad v(\xi) \leq \eta \quad (\text{bzw. } w' \geq F(x, w) \text{ in } J, \quad w(\xi) \geq \eta)$$

bereits hinreichend dafür, dass v Unterfunktion (bzw. w Oberfunktion) ist. Und zwar in dem Sinne, dass dann

$$v(x) \leq y(x) \leq w(x)$$

in ganz J gilt.

Betrachten wir ein links an die Stelle ξ anschließendes Intervall $\tilde{J} := [\xi - a, \xi]$, so lauten die Bedingungen für eine Unterfunktion (bzw. für eine Oberfunktion) entsprechend

$$v' \left\{ \begin{array}{l} > \\ \geq \end{array} \right\} F(x, v) \text{ in } \tilde{J}, \quad v(\xi) \leq \eta \quad \left(\text{bzw. } w' \left\{ \begin{array}{l} < \\ \leq \end{array} \right\} F(x, w) \text{ in } \tilde{J}, \quad w(\xi) \geq \eta \right).$$

Beweis [13, S. 71ff.]

Bezogen auf die Differentialungleichung (3.21) ist in obiger Schreibweise

$$F_{2/3}(x, y) := -\frac{2}{3x} \cdot y \in C^\infty(\mathbb{R}_+ \times \mathbb{R}),$$

und wir wissen bereits, dass $y' = F_{2/3}(x, y)$ mit entsprechender Anfangsbedingung eindeutig lösbar ist. Es ergibt sich also unmittelbar die

Folgerung 3.34

Sei y differenzierbar in \mathbb{R}_+ und erfülle dort

$$y' \begin{cases} > \\ \geq \end{cases} F_{2/3}(x, y), \quad y(\xi) \geq \eta$$

für $\xi > 0 \geq \eta$.

$$\implies y(x) \begin{cases} > \\ \geq \end{cases} y_{\text{limit}}(x) = \eta \xi^{2/3} \cdot x^{-2/3} \quad \text{auf} \quad \begin{cases} (\xi, \infty) \\ [\xi, \infty) \end{cases}.$$

Analog:

$$y' \begin{cases} > \\ \geq \end{cases} F_{2/3}(x, y), \quad y(\xi) \leq \eta \implies y(x) \begin{cases} < \\ \leq \end{cases} y_{\text{limit}}(x) \quad \text{auf} \quad \begin{cases} (0, \xi) \\ (0, \xi] \end{cases}.$$

Dies ermöglicht uns den qualitativen Verlauf der Lösungen von (3.21) zu zeichnen. Hierzu bemerken wir:

Bemerkung 3.35

- (i) Unsere Aussagen beziehen sich auf $y = f'$! Für f stelle man sich entsprechend eine Stammfunktion vor. Wir werden aber später auch noch Funktionen f angeben.
- (ii) Anstelle von einer Differentialungleichung sollte man in (3.21) besser von Differentialungleichungen sprechen. Denn es ist auch $y \leq 0$ zu erfüllen. Wollen wir also (3.21) lösen, so müssen wir in **Folgerung 3.34** auch $0 \geq y(\xi) \geq \eta$ fordern.

Für $\eta = 0$ ist $y_{\text{limit}} \equiv 0$ nach (3.25) die eindeutige Lösung, und schneidet also y auf jeden Fall in ξ . Denn es bleibt für y nur $y(\xi) = 0$ übrig.

Im Falle $0 \geq y(\xi) > \eta$ wird y natürlich von der Lösung y_{limit} mit dem Anfangswert $y_{\text{limit}}(\xi) = y(\xi)$ geschnitten.

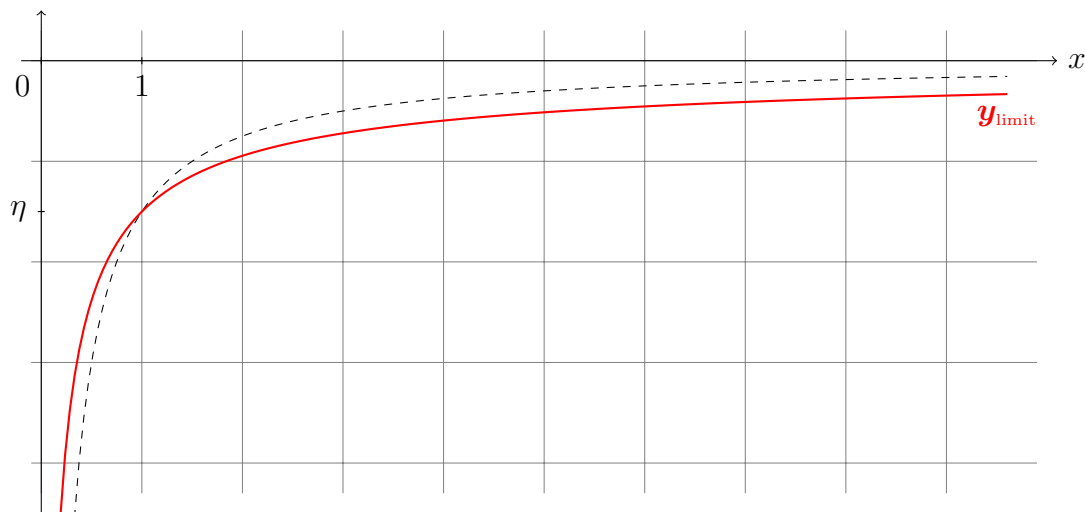
Kurz und gut: Um uns den qualitativen Verlauf besser vorstellen zu können, betrachten wir in Folgerung 3.34 fortan oBdA nur noch $y(\xi) = \eta = y_{\text{limit}}(\xi)$. Die beiden Fälle lassen sich also zu einem Fall zusammenfassen:

$$\underbrace{y' \begin{cases} > \\ \geq \end{cases} - \frac{2}{3x} \cdot y = F_{2/3}(x, y), \quad y(\xi) = \eta}_{\Downarrow}$$

$$y \begin{cases} > \\ \geq \end{cases} y_{\text{limit}} \quad \text{auf} \quad \begin{cases} (\xi, \infty) \\ [\xi, \infty) \end{cases} \quad \wedge \quad y \begin{cases} < \\ \leq \end{cases} y_{\text{limit}} \quad \text{auf} \quad \begin{cases} (0, \xi) \\ (0, \xi] \end{cases}$$

- (iii) Die Graphen der Lösungen y_{limit} verlaufen durch die Punkte $(1, \eta \xi^{2/3})$. Es genügt also für die Anfangswerte nur $y(1) = \eta = y_{\text{limit}}(1)$ für $\eta \leq 0$ zu betrachten.

Existieren also zu (3.21) noch andere Lösungen als y_{limit} , so müssen diese qualitativ den in der Skizze gestrichelten Verlauf nehmen,



denn wegen $y \leq 0$ müssen diese ebenfalls im vierten Quadranten verlaufen. Auch folgt aus (3.21) unmittelbar $y' = f'' \geq 0$, d.h. y muss monoton wachsen.

Des Weiteren erkennt man, dass f und f_{limit} an der Stelle 1 dieselbe Steigung besitzen. In $(0, 1)$ hingegen fällt f stärker als f_{limit} , in $(1, \infty)$ schwächer. Weiter unten werden wir den Verlauf von verschiedenen Funktionen f in einem Koordinatensystem sehen, wodurch dieser Effekt deutlich wird.

Die Frage lautet nun natürlich: Gibt es außer den y_{limit} noch weitere Lösungen von (3.21)? Eine Möglichkeit dies herauszufinden besteht darin, die „rechte Seite“ $F_{2/3}(x, y)$ gemäß

$$\tilde{F}(x, y) \left\{ \begin{array}{l} > \\ \geq \end{array} \right\} F_{2/3}(x, y)$$

geringfügig abzuändern und zu versuchen, die Differentialgleichung

$$y' = \tilde{F}(x, y)$$

explizit zu lösen. Wählt man z. B. für $\varepsilon > 0$

$$\tilde{F}(x, y) := F_{2/3}(x, y) + \varepsilon,$$

so ist die lineare inhomogene Differentialgleichung $y' = \tilde{F}(x, y)$ zwar stets auf ganz \mathbb{R}_+ lösbar, aber für die Lösungen y zeigt sich, dass $y \leq 0$ stets nur in einer beschränkten Umgebung von ξ und nicht auf ganz \mathbb{R}_+ gilt.

Die Modifikation

$$\tilde{F}(x, y) := F_{(2/3)+a}(x, y) = -\frac{y}{x} \cdot \left(\frac{2}{3} + a \right) = -\underbrace{\frac{3a+2}{3x}}_{=: g_a(x)} y$$

für $a \geq 0$ ($a = 0$ liefert natürlich y_{limit}) führt jedoch zum Erfolg. Wobei

$$F_{(2/3)+a}(x, y) \left\{ \begin{array}{l} > \\ \geq \end{array} \right\} F_{2/3}(x, y)$$

für $a > 0$ nur dann erfüllt sein kann, wenn die Lösungen, welche wir mit y_a bezeichnen wollen, auch $y_a \leq 0$, also die zweite Ungleichung in (3.21) erfüllen. Dies ist aber gegeben, denn die eindeutig bestimmte und auf ganz \mathbb{R}_+ existente Lösung y_a des linearen homogenen Anfangswertproblems

$$\left\{ \begin{array}{l} y'_a = -\frac{3a+2}{3x} y_a \Leftrightarrow y'_a + g_a(x) y_a = 0 \\ \wedge \quad y(1) = \eta \leq 0 \end{array} \right\} \quad (3.26)$$

ergibt sich gemäß (3.22) zu

$$y_a(x) = \eta \cdot \exp\left(-\int_1^x g_a(t) dt\right) = \eta \cdot x^{-(\frac{2}{3}+a)} \leq 0. \checkmark$$

Durch Integration erhalten wir also eine ganze Schar an Funktionen die (3.21) lösen, wobei dies nicht alle Lösungen sind. Natürlich gilt $f_{\text{limit}} = f_0$.

Dies fassen wir in dem folgenden Satz zusammen, und liefern anschließend eine Skizze für $a \in \{0, 1/6, 1/3, 2/3\}$:

Satz 3.36 (LÖSUNGSSCHAR FÜR DAS **PROBLEM (3.1)**)

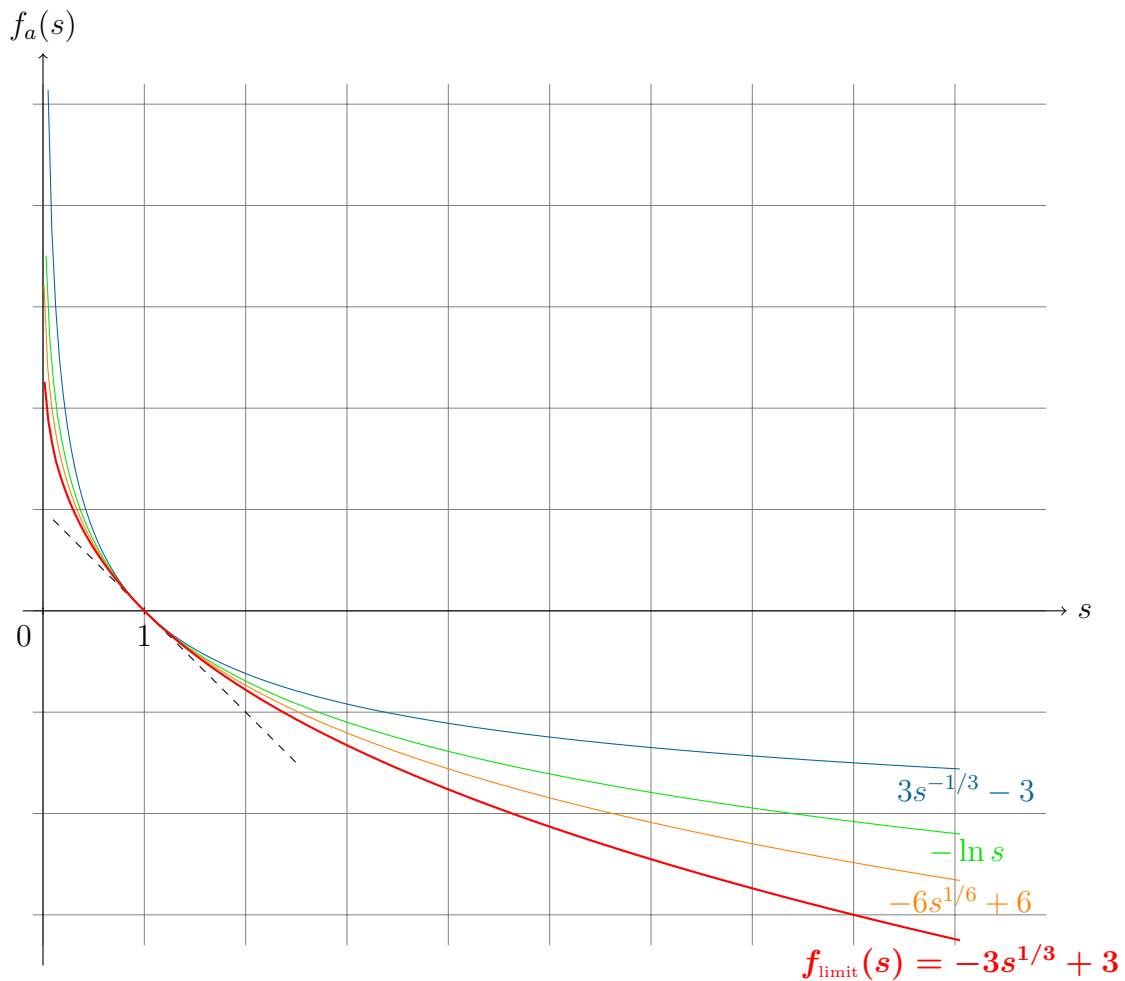
Für $c \leq 0$, $d \in \mathbb{R}$ beliebig, besitzt die für $a \in [0, \infty)$ auf \mathbb{R}_+ definierte Funktionenschar

$$f_a(s) := \begin{cases} d + c \cdot s^{\frac{1}{3}-a} & , \text{ für } a \in [0, 1/3) \\ d + c \cdot \ln s & , \text{ für } a = 1/3 \\ d - c \cdot s^{\frac{1}{3}-a} & , \text{ für } a \in (1/3, \infty) \end{cases} \quad (3.27)$$

die Eigenschaft, dass

$$\begin{aligned} f_a \circ \det & : \mathbb{P}\text{Sym}(3) \longrightarrow \mathbb{R} \\ C & \longmapsto f_a(\det C) \end{aligned}$$

konvex ist.



Diese anschauliche Darstellung kann man aus genannten Gründen für jedes f_{limit} durch geeignete Wahl von c und d erzwingen, d.h. für alle f_{limit} kann der qualitative Verlauf weiterer Lösungen unseres Problems in dieser Form angegeben werden. Die spezielle Wahl von c (und damit d) in der Skizze ergibt sich aus der Integration von

$$f'_a(x) = y_a(x) = \eta \cdot x^{-\left(\frac{2}{3}+a\right)},$$

denn man erhält

$$c = \frac{3}{1-3a}\eta \quad \text{für } a \in [0, \infty) \setminus \{1/3\} \quad \wedge \quad c = \eta \quad \text{für } a = 1/3.$$

Wie in der Skizze durch die gestrichelte Linie angedeutet, wurde außerdem $\eta = -1$ gewählt.

4. Nichtlineare Elastizität

In der Einleitung haben wir zum Einen motiviert, warum die Simulation/Modellierung von Bauteilen und Materialien wichtig ist, und zum Anderen, warum man Elastomere besser anhand nichtlinearer Modelle untersucht. Wir wollen nun einen kurzen Überblick über die wichtigsten Ideen, Bezeichnungen und Annahmen einer physikalisch sinnvollen Materialmodellierung geben, wobei wir uns an [11], [10] und [12] orientieren:

Unter der „Deformation“ eines Körpers $\Omega \subseteq \mathbb{R}^3$ aus einem gewissen (in unserem Fall: elastischen) Material, verstehen wir nichts anderes als eine Abbildung

$$\varphi : \Omega \longrightarrow \varphi(\Omega) \subseteq \mathbb{R}^3 .$$

Ω nennt man auch „Referenzkonfiguration“ und $\varphi(\Omega)$ einen „deformierten Körper“. Hierbei sei die Deformation φ nicht explizit von der Zeit t abhängig: Wir beschränken uns also auf den einfacheren Fall nichtlinearer *Elastostatik*.

Diese allgemeine Definition einer Deformation schließt aber natürlich auch solch physikalisch unsinnige Deformationen mit ein, in denen ein dreidimensionaler Körper im Extremfall auf einen einzigen Punkt zusammengedrückt wird. Um dies zu verhindern, müssen wir weitere Forderungen an unsere Deformation stellen:

Zunächst einmal sei unser Körper Ω von endlicher Größe und besitze einen hinreichend glatten Rand $\partial\Omega$. Auch wollen wir verhindern, dass unser Körper lokal wie eine zweidimensionale Fläche oder dergleichen aussieht. Mathematisch bedeuten diese Forderungen nichts anderes, als dass Ω offen und beschränkt sein soll und z. B. einen LIPSCHITZrand besitze. Zudem sei die Deformation φ hinreichend glatt. Eine weitere oft gestellte Forderung ist die Bijektivität von φ , um „Überlappungen“ von Teilen des Materials durch die Deformation zu vermeiden. Mit

$$\begin{aligned} F & : \Omega \longrightarrow \mathbb{M}^{3 \times 3} \\ x & \longmapsto F(x) := \nabla \varphi(x) = D\varphi(x) \end{aligned} , \tag{4.1}$$

kurz $F = \nabla \varphi$, sei eine wesentliche kinematische Größe bezeichnet: Der „Deformationsgradient“. Um Spiegelungen oder dgl. auszuschließen, fordert man außerdem eine Orientierungserhaltende Deformation, i. e.

$$\det(F(x)) > 0 \quad \forall x \in \Omega . \tag{4.2}$$

Daraus folgt unmittelbar, dass für alle $x \in \Omega$ $C(x) := F^T(x)F(x) \in \mathbb{P}\text{Sym}(3)$. Wie bereits des Öfteren erwähnt, nennen wir C den „rechten CAUCHY-GREEN-Deformationstensor“.

Für $\varphi \in W^{1,p}(\Omega) = W^{1,p}(\Omega, \mathbb{R}^3)$ wie in [Abschnitt 2.3](#) gilt alles Gesagte entsprechend fast überall. $\varphi_0 \in W^{1,p}(\Omega)$ aus diesem Abschnitt stellt eine globale Randbedingung dar, d. h. wir geben jedem Randpunkt eine feste Verschiebung vor.¹

Den Bogen zu den Variationsformulierungen spannt nun die Kontinuumsmechanik, welche ein Kräftegleichgewicht formuliert, die sogenannte „Gleichgewichtskonfiguration“. Dies ist aber gerade die EULER-LAGRANGE-Gleichung² zu dem Energiefunktional I mit

$$I(\varphi) := \int_{\Omega} W(x, \varphi(x), \nabla \varphi(x)) dx, \quad (4.3)$$

identisch zu unserer bisherigen Schreibweise. Das Energiefunktional beschreibt die zur Deformation φ nötige Energie. W nennt man „freie Energiefunktion“ oder auch „gespeicherte Energiedichte“. Ein Minimierer von I aus der Klasse zulässiger Deformationen \mathcal{A} , beschreibt dann die energetisch günstigste Deformation, welche die vorgegebenen Randwerte realisiert.³

Hyperelastische Werkstoffe sind dadurch gekennzeichnet, dass sich Deformationen in der Regel spontan, oder aber zeitverzögert vollständig zurückbilden. Es stellt sich heraus, dass die freie Energiefunktion W in diesem Falle nicht von der Deformation φ abhängt. Falls das Material homogen ist, so ist W zudem nicht explizit von der Variable x abhängig. Von nun an sei also $W = W(F)$. An diese freie Energiefunktion stellt man weitere Bedingungen:

- „frame-indifference“

$$W(QF) = W(F) \quad (4.4)$$

für alle $F \in \mathbb{M}^{3 \times 3}$ und alle Rotationsmatrizen $Q \in \text{SO}(3)$.

- „Verhalten für extreme Deformationen“

$$W(F) \rightarrow \infty \quad \text{für } \det F \searrow 0 \quad \wedge \quad W(F) = \infty \quad \text{für } \det F \leq 0. \quad (4.5)$$

Für weitere Anmerkungen verweisen wir auf den [Unterabschnitt 2.3.3](#).

4.1. Das NEO-HOOKE-Modell

Je nach Material stehen zahlreiche Modelle zur Modellierung in Form der freien Energiefunktion W zur Verfügung. Für gummiartige Materialien ist das Modell von OGDEN

¹ Wir wollen uns auf diesen Fall beschränken. Man könnte auch für verschiedene Teile des Randes verschiedene Bedingungen stellen, z. B. Randwertvorgaben auf $\Gamma_0 \subseteq \partial\Omega$ und Zugbedingungen auf $\Gamma_1 \subseteq \partial\Omega$, wobei $\Gamma_0 \cap \Gamma_1 \neq \emptyset$ nicht ausgeschlossen ist.

² Womöglich in einer schwachen Formulierung, siehe [Bemerkung 2.13](#) auf Seite 17

³ Dass solch ein Minimierer nicht immer eindeutig ist, führe man sich z.B. anhand einer dünnen länglichen Platte aus Metall vor Augen. Übt man in Längsrichtung Druck aus, so ist nicht klar, ob sich die Platte nach oben oder unten biegt.

mit

$$\begin{aligned} W(F) &= \sum_{i=1}^r a_i \operatorname{tr} \left((F^T F)^{\alpha_i/2} \right) + \sum_{j=1}^s b_j \operatorname{tr} \left((\operatorname{Adj} (F^T F))^{\beta_j/2} \right) + f(\det F) \\ &= \sum_{i=1}^r a_i \operatorname{tr} (C^{\alpha_i/2}) + \sum_{j=1}^s b_j \operatorname{tr} \left((\operatorname{Adj} C)^{\beta_j/2} \right) + f(\det F) \end{aligned} \quad (4.6)$$

ein bekannter Vertreter. Hierbei seien $r, s \in \mathbb{N}$, $a_i, b_j \in \mathbb{R}_+$, $\alpha_i, \beta_j \geq 1$, f eine konvexe Funktion und die Potenz einer positiv definiten symmetrischen Matrix über die entsprechenden Potenzen der positiven Eigenwerte gemäß der üblichen Zerlegung in [Lemma 3.5](#) erklärt.

Setzt man $r = 1$, $\alpha_1 = 2$, $a = a_1 \in \mathbb{R}_+$ und alle b_j gleich Null, so erhält man als Spezialfall bzw. leichte Abwandlung das „NEO-HOOKE-Modell“

$$W(F) = a \|F\|^2 + f(\det F), \quad (4.7)$$

denn $\operatorname{tr} (F^T F) = \langle F^T F, \mathbf{1} \rangle_{3 \times 3} = \langle F, F \rangle_{3 \times 3}$.

Untersucht man nun die Konvexität der freien Energiefunktion bzgl. F , so ist Vorsicht geboten: CIARLET [5, S. 162] hat gezeigt, dass $\{F \in \mathbb{M}^{3 \times 3} : \det F > 0\} \subseteq \mathbb{M}^{3 \times 3}$ nicht konvex ist. Eine realistische Energiefunktion sollte aber unsere Forderung (4.5) für das Verhalten extremer Deformationen erfüllen. Wie man sofort anhand von [Definition 3.1](#) auf Seite 23 sieht, kann man auf konvexen Teilmengen K definierte konvexe Funktionen, wobei K Teilmenge eines normierten Raumes X sei, durch den „Wert“ $+\infty$ auf $X \setminus K$ problemlos zu auf ganz X konvexen Funktionen mit Werten in \mathbb{R}^∞ fortsetzen. Die Frage ist also, wie man mit dem Problem umgeht, dass $\{F \in \mathbb{M}^{3 \times 3} : \det F > 0\}$ nicht konvex ist. Glücklicherweise brauchen wir daran aber keinen Gedanken verschwenden, denn wir wissen bereits nach [Unterabschnitt 2.3.3](#), dass eine bezüglich F konvexe Energiefunktion W nicht die Singularität in (4.5) widerspiegeln kann. Aus diesem Grunde sind wir ja gerade zu dem Konzept der Polykonvexität übergegangen. Wegen $W = W(F)$ und der Tatsache, dass $\operatorname{Adj} F$ in (4.7) nicht auftritt, vereinfacht sich der Nachweis der Polykonvexität von W in der Schreibweise gemäß [Definition 2.15](#) zum Nachweis der Konvexität von

$$\begin{aligned} P : \quad \mathbb{R}^{10} &\longrightarrow \mathbb{R} \\ (A, B) &\longmapsto P(A, B) := a \|A\|^2 + f(B) \end{aligned}$$

Wegen $D^2(A \mapsto \langle A, A \rangle) \cdot (H, H) = 2 \cdot \|H\|^2$ für $H \in \mathbb{M}^{3 \times 3}$ beliebig und $a > 0$, ist der erste Summand sogar strikt konvex. Der zweite Term ist trivialerweise konvex in B , da dies Voraussetzung an die Funktion f war. Somit ist P (strikt) konvex in (A, B) und mithin W polykonvex.

Nehmen wir an, es gebe Minimierer des Energiefunktional I . Im Allgemeinen ist es schwierig nachzuweisen, dass diese Minimierer auch die (ELG) lösen, da weitere Wachstumsbedingungen an W gestellt werden müssten. Genauso ist nicht unmittelbar klar, wie viel Regularität solch ein Minimierer besitzt, d.h. wie glatt er ist. Eine weitere interessante

Frage ist die der Stabilität. Kann es sein, dass sich der Körper gänzlich anders deformiert, obgleich man die Randbedingungen nur geringfügig ändert? Dies ist natürlich zu vermeiden. Um noch etwas mehr Informationen zu gewinnen, schreiben wir unsere freie Energiefunktion W zunächst als Funktion \widetilde{W} in Abhängigkeit von C . Wegen $\det F > 0$ nach (4.2) und $\det F = \det F^T$ ist

$$\det F = (\det F^T \cdot \det F)^{1/2} = (\det C)^{1/2} .$$

Wie wir bereits gesehen haben, ist $\|F\|^2 = \langle F^T F, \mathbf{1} \rangle$. Somit gilt also $W(F) = \widetilde{W}(C)$ mit

$$\widetilde{W}(C) := a \langle C, \mathbf{1} \rangle + \frac{1}{2} f(\det C) . \quad (4.8)$$

Der erste Summand ist offensichtlich konvex in C . Wann dies auch für den zweiten Summanden gilt, war Gegenstand des vorigen Kapitels im Falle $f \in C^2(\mathbb{R}_+)$. In **Theorem 3.30** haben wir gezeigt, dass dies genau dann der Fall ist, wenn f (3.21) erfüllt:

$$f''(s) + \frac{2}{3s} \cdot f'(s) \geq 0 \quad \wedge \quad f'(s) \leq 0 \quad \forall s \in \mathbb{R}_+ .$$

Insbesondere gilt also, dass $f'' \geq 0$ und f mithin konvex ist. Um also eine bezüglich C konvexe freie Energiefunktion zu erhalten, müssen wir die die Auswahl an konvexen Funktionen f einschränken. Wie bereits in **Unterabschnitt 2.3.3** erwähnt und in **Satz 3.36** anschaulich dargestellt, schließt dies auch unser gefordertes Verhalten für extreme Deformationen nicht aus. Als Funktion von C darf unsere freie Energiefunktion W eine Singularität für $\det C \rightarrow 0 \Leftrightarrow \det F \rightarrow 0$ besitzen.

Die Hoffnung ist nun, aus der Konvexität in C weitere Erkenntnisse über die Regularität und Stabilität von Lösungen des Variationsproblems gewinnen zu können, und stellt einen möglichen Anknüpfungspunkt zur weiteren Vertiefung der Thematik dar. Zum Abschluss noch eine weitere Bemerkung:

Bemerkung 4.1

Ein anderes klassisches nichtlineares Modell, z. B. für Metalle, ist das sogenannte „Saint VENANT-KIRCHHOFF-Modell“. Eine typische Energiefunktion formuliert in Abhängigkeit des rechten CAUCHY-GREEN-Deformationstensors ist

$$\widetilde{W}(C) = \frac{1}{2} \|C - \mathbf{1}\|^2 - \ln(\det C) .$$

Auch diese freie Energiefunktion ist konvex in C und könnte von weiteren Erkenntnissen profitieren.

Literaturverzeichnis

- [1] H. W. Alt. *Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung*. Springer, Berlin, 5., überarbeitete Auflage, 2006. ISBN 978-3-540-34186-4. URL <http://www.springerlink.de/content/g05242>. 16, 34
- [2] J. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rat. Mech. Anal.*, 63:337–403, 1977. 19
- [3] J. Ball. Constitutive inequalities and existence theorems in nonlinear elastostatics. In R. Knops, editor, *Herriot Watt Symposium: Nonlinear Analysis and Mechanics.*, volume 1, pages 187–238. Pitman, London, 1977. 19
- [4] J. M. Borwein and J. D. Vanderwerff. *Convex Functions: Constructions, Characterizations and Counterexamples*. Cambridge University Press, 2010. ISBN 978-0-521-85005-6. 8
- [5] P. G. Ciarlet. *Mathematical Elasticity, Volume I: Three-Dimensional Elasticity*. 19, 55
- [6] B. Dacorogna. *Direct Methods in the Calculus of Variations*, volume 78 of *Applied Mathematical Sciences*. Springer, New York, 2. Auflage, 2008. ISBN 978-0-387-35779-9. 6, 7, 10
- [7] W. Dahmen and A. Reusken. *Numerik für Ingenieure und Naturwissenschaftler*. Springer, Berlin Heidelberg, 2., korrigierte Auflage, 2008. ISBN 978-3-540-76492-2. URL <http://www.springerlink.de/content/x5737x>. 26
- [8] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, 1998. ISBN 978-0-821-80772-9. 7, 10, 17, 20
- [9] J. Kristensen. On conditions for polyconvexity. *Proceedings of the American Mathematical Society*, 128. 20
- [10] J. Marsden and J. Hughes. *Mathematical Foundations of Elasticity*. Prentice-Hall, Englewood Cliffs, New Jersey, 1983. 53
- [11] P. Neff. *Mathematische Analyse multiplikativer Viskoplastizität*. PhD thesis, Technische Universität Darmstadt, Aachen, 2000. URL <http://tubiblio.ulb.tu-darmstadt.de/1147/>. 30, 32, 35, 40, 53
- [12] P. Pedregal. *Variational Methods in Nonlinear Elasticity*. SIAM, Philadelphia, 2000. ISBN 978-0-898-71452-4. 6, 21, 53

- [13] W. Walter. *Gewöhnliche Differentialgleichungen: eine Einführung*. Springer Lehrbuch. Springer, Berlin, 4., überarbeitete u. ergänzte Auflage, 1990. ISBN 978-3-540-52017-7. 45, 47
- [14] D. Werner. *Funktionalanalysis*. Springer, Berlin, 6., korrigierte Auflage, 2007. ISBN 978-3-540-72533-6. URL <http://www.springerlink.com/content/t1166x/>. 14, 15

A. Ein Gegenbeispiel

Eine andere Vermutung zur Lösung unseres Problems könnte auf

$$f(\det C) \text{ konvex in } C = C^T > 0$$

$$\Updownarrow$$

$$-\frac{f'(s)}{3} \stackrel{\textcircled{1}}{\leq} s \cdot f''(s) \stackrel{\textcircled{2}}{\leq} -f'(s) \quad \wedge \quad f'(s) \leq 0 \quad \forall s \in \mathbb{R}_+$$

mit den beiden Grenzfällen

$$\textcircled{1} f(s) = -s^{2/3} \quad , \quad \textcircled{2} f(s) = -\ln s$$

für $s \in \mathbb{R}_+$ führen. Der Grenzfall $\textcircled{2}$ zählt auch zu unserer Lösungsschar, aber wir haben in (3.21) gezeigt, dass $s \cdot f''(s) + f'(s)$ auf \mathbb{R}_+ nicht nach oben beschränkt ist.

Den „unteren“ Grenzfall $\textcircled{1}$ kann man durch die Wahl

$$D^{-1} := s^{-1/3} \cdot \mathbf{1} \quad , \quad H := \mathbf{1}$$

sofort widerlegen, denn dann ist die Konvexitätsungleichung (3.13) verletzt. Wir wollen aber zusätzlich noch ein Gegenbeispiel bzgl. der konkreten Konvexitätsdefinition angeben:

Wählen wir

$$C_1 := \mathbf{1} \quad , \quad C_2 := 2 \cdot \mathbf{1} \in \mathbb{P}\text{Sym}(3) \quad , \quad \tau := \frac{1}{2} \quad ,$$

so folgt

$$f \circ \det((1 - \tau)C_1 + \tau C_2) \stackrel{!}{\leq} (1 - \tau) \cdot (f \circ \det(C_1)) + \tau \cdot (f \circ \det(C_2))$$

$$\Leftrightarrow f(\det(\frac{3}{2} \cdot \mathbf{1})) \leq \frac{1}{2} \cdot f(\det \mathbf{1}) + \frac{1}{2} \cdot f(\det(2 \cdot \mathbf{1}))$$

$$\Leftrightarrow f\left(\frac{27}{8}\right) \leq \frac{1}{2}f(1) + \frac{1}{2}f(8)$$

$$\Leftrightarrow -\left(\frac{27}{8}\right)^{2/3} \leq -\frac{1}{2} - \frac{1}{2}8^{2/3}$$

$$\Leftrightarrow -\frac{9}{4} \leq -\frac{1}{2} - 2$$

$$\Leftrightarrow \frac{1}{4} \leq 0 \quad .$$

Ein Widerspruch.