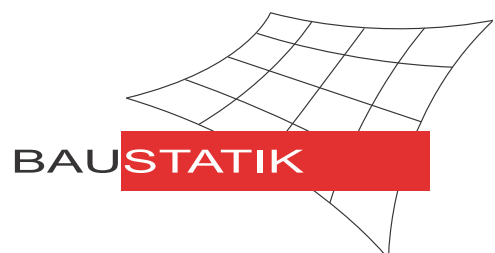


Modellierung effektiver Materialeigenschaften von viskoelastischen Schalenstrukturen mit künstlichen neuronalen Netzen

Jeremy Geiger



Modellierung effektiver Materialeigenschaften von viskoelastischen Schalenstrukturen mit künstlichen neuronalen Netzen

Zur Erlangung des akademischen Grades eines
DOKTORS DER INGENIEURWISSENSCHAFTEN (Dr.-Ing.)
von der KIT-Fakultät für
Bauingenieur-, Geo- und Umweltwissenschaften
des Karlsruher Instituts für Technologie (KIT)
genehmigte
DISSERTATION

von
Jeremy Geiger, M.Sc.
aus Stuttgart

Tag der mündlichen Prüfung: 05. Dezember 2025

Referent: Prof. Dr.-Ing. habil. Werner Wagner
Korreferenten: Prof. Dr.-Ing. Oliver Weeger
Prof. Dr.-Ing. Steffen Freitag

Karlsruhe 2026

Herausgeber:

Prof. Dr.-Ing. Steffen Freitag
Prof. Dr.-Ing. habil. Werner Wagner

Organisation und Verwaltung:

Institut für Baustatik
Karlsruher Institut für Technologie
Kaiserstr. 12
76131 Karlsruhe

Telefon: (0721) 608-42280
Telefax: (0721) 608-46015
E-Mail: info@ibs.kit.edu

© 2026 Jeremy Geiger
 Institut für Baustatik
 Karlsruher Institut für Technologie
 Kaiserstraße 12
 76131 Karlsruhe

Das vorliegende Werk darf von der KIT-Bibliothek frei im Internet angeboten werden. Die Nutzung erfolgt ausschließlich zu wissenschaftlichen Zwecken und zum Eigengebrauch. Die Urheberrechte liegen bei den Autorinnen bzw. Autoren. Für den Inhalt sind alleine die Autorinnen und Autoren verantwortlich. Jegliche Formen der kommerziellen Nutzung und Abänderung der Publikation sind ohne vorherige Zustimmung und Absprache mit den Autorinnen bzw. Autoren ausdrücklich verboten. Die Namen der Autorinnen und Autoren müssen stets genannt werden. Die Nutzerinnen und Nutzer sind für die Einhaltung der Rechtsvorschriften selbst verantwortlich und können bei Missbrauch haftbar gemacht werden.

ISBN 978-3-935322-31-7

Vorwort

Die vorliegende Dissertation entstand während meiner Zeit als wissenschaftlicher Mitarbeiter am Institut für Baustatik des Karlsruher Instituts für Technologie.

Bedanken möchte ich mich insbesondere bei meinem Hauptreferenten, Prof. Dr.-Ing. habil. Werner Wagner, dessen fachliche Führung als ehemaliger Institutsleiter sowie dessen Betreuung im Rahmen der Promotion ich immer sehr zu schätzen wusste. Seine Begeisterung für die Forschung und sein Interesse an dem Thema der Arbeit trugen maßgeblich zum Gelingen dieser bei. Sowohl die fachlichen als auch die persönlichen Diskussionen über Themen, die jenseits des Instituts stattfanden, waren eine große Bereicherung, ebenso wie die gemeinsamen Aktivitäten im Rahmen von Institutsausflügen.

Prof. Dr.-Ing. Steffen Freitag danke ich für das entgegengebrachte Vertrauen, für die Schaffung eines angenehmen Forschungs- und Arbeitsklimas als Institutsleiter sowie für die Übernahme eines der beiden Korreferate. Weiterhin danke ich Prof. Dr.-Ing. Oliver Weeger für die Bereitschaft, das Korreferat zu übernehmen, sowie für die fachlichen Diskussionen.

Ich konnte mich während meiner Zeit am Institut glücklich schätzen, mit Menschen zusammenarbeiten zu dürfen, die ich heute nicht nur als Kollegen, sondern als gute Freunde betrachte. Hervorheben möchte ich dabei Patrick und Lukas, mit denen unzählige fachliche Diskussionen – oft bis in die späten Stunden – den entscheidenden Fortschritt brachten. Auch den weiteren promovierenden Kolleginnen und Kollegen, die es mir leicht machten, die Freude an der Arbeit zu bewahren, gilt mein Dank. Mit allen weiteren Mitarbeiterinnen und Mitarbeitern war das Arbeitsklima stets von einem positiven Miteinander geprägt. Zudem war die Zusammenarbeit mit den Studierenden – sei es in der Lehre, bei der Betreuung von Abschlussarbeiten oder als wissenschaftliche Hilfskräfte – eine große persönliche und fachliche Bereicherung, wobei insbesondere Letzteres auch in Teilen in diese Arbeit einfließen konnte.

Außerhalb des Instituts habe ich in meinem persönlichen Umfeld großen Rückhalt erfahren. Ich danke allen, die sich Zeit für das Gegenlesen genommen haben, sowie jenen, die mich bei der Organisation der Feierlichkeiten entlastet haben. Mein besonderer Dank gilt meinem Bruder Joshua, der mir immer bedingungslos zur Seite stand, sowie meiner Freundin Leonie, die mich in den entscheidenden Phasen mit großem Zuspruch und moralischer Unterstützung begleitet hat.

Karlsruhe, im März 2026

Jeremy Geiger

Kurzfassung

Die vorliegende Arbeit befasst sich mit dem Einsatz künstlicher neuronaler Netze (KNN) als Materialmodell zur Approximation der effektiven konstitutiven Beziehungen von Schalen mit viskoelastischen Mikrostrukturen.

Für die Konstruktion des KNN-Materialmodells wird eine Strategie unter Verwendung synthetischer Materialversuche entwickelt, bei der auf Basis repräsentativer Volumenelemente (RVE) effektive Spannungs-Dehnungs-Paare komplexer viskoelastischer Mikrostrukturen generiert werden. Das KNN erlernt daraus die effektive konstitutive Beziehung und kann anschließend als Ersatzmodell für Mehrskalenlösungen eingesetzt werden.

Die vorgeschlagene Methodik ist grundsätzlich auch für den Aufbau des Modells anhand experimenteller Daten geeignet. Es wird gezeigt, dass durch die Verwendung von Relaxations-, Kriech- und zyklischen Versuchen die relevanten Zeitbereiche des viskoelastischen Materialverhaltens gut abgedeckt werden können. Zudem werden Vergleiche zu Strategien mit zufälliger Datengenerierung durchgeführt.

Um im mehrdimensionalen Fall die erforderliche Menge an Primärdaten zu reduzieren und gleichzeitig robuste Generalisierungsfähigkeiten zu erreichen, wird zusätzlich zur Information aus uniaxialen synthetischen Materialversuchen die aus der Homogenisierung berechnete Materialtangente über Nebenbedingungen im Rahmen des nichtlinearen Optimierungsproblems berücksichtigt. Es zeigt sich, dass diese Nebenbedingungen einen signifikanten Informationszuwachs darstellen, durch welchen ein stabiles Materialmodell mit sehr guten Approximationsfähigkeiten aufgebaut werden kann. Darüber hinaus führt ihre Einbeziehung zu einer verbesserten Approximation im Vergleich zu rein auf Primärdaten trainierten KNN sowie zu einer deutlichen Reduktion der benötigten Primärdaten und damit der notwendigen FE-Berechnungen auf dem RVE.

Neben Untersuchungen am Materialpunkt wird das volle Potenzial des KNN ausgeschöpft, indem es als Materialmodell in numerischen Beispielen eingesetzt wird. Dabei zeigt sich, dass das Modell in der Finite-Elemente-Methode (FEM) im Vergleich zu 3D-Modellen und Schalen-Mehrskalenlösungen eine deutlich höhere Recheneffizienz aufweist.

Abschließend wird die Möglichkeit zur Berücksichtigung geometrischer Parameter des RVEs im Eingangsvektor des KNN untersucht. Damit lassen sich mit einem einzigen Modell die effektiven Eigenschaften von Schalen-RVEs mit unterschiedlichen geometrischen Merkmalen abbilden. Dies stellt insbesondere im Designprozess ein vielversprechendes Werkzeug dar, um das makroskopische Verhalten der Gesamtstruktur in Abhängigkeit von den Eigenschaften der Mikrostruktur zu untersuchen.

Abstract

The present work investigates the use of artificial neural networks (ANN) as material models for approximating the effective constitutive behavior of shells with viscoelastic microstructures.

A strategy for constructing the ANN-based material model is developed using synthetic material tests, where based on representative volume elements (RVEs) effective stress–strain pairs of complex viscoelastic microstructures are generated. From these data, the ANN learns the effective constitutive relationship and can subsequently be used as a surrogate model for multiscale simulations.

The proposed methodology is, in principle, also suitable for model construction using experimental data. It is shown that by making use of relaxation, creep, and cyclic loading tests, the relevant time ranges of viscoelastic behavior can be well represented. Furthermore, the results are compared with alternative strategies based on randomly generated data.

To reduce the required amount of primary data in the multidimensional case and to achieve robust generalization capabilities, additionally to the information from uniaxial synthetic material tests the homogenized material tangent is considered, which is incorporated as constraints within the nonlinear optimization problem. These constraints provide a significant increase in information, enabling the construction of a stable material model with very good approximation capabilities. In addition, their inclusion improves the prediction accuracy compared to models trained solely on primary data and considerably reduces both the amount of required primary data and therefore the number of necessary FE-computations on the RVE.

Beyond material-point investigations, the full potential of the ANN is exploited by employing it as a constitutive model in numerical examples. When used in finite element simulations, the ANN model demonstrates a substantially higher computational efficiency compared to 3D models and shell-based multiscale approaches.

Finally, the integration of geometric RVE parameters into the ANN input vector is investigated. This allows the effective properties of shell RVEs with varying geometrical characteristics to be represented by a single model. In particular, this provides a promising tool for design applications, enabling the analysis of the macroscopic structural behavior as a function of the underlying microstructural features.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.2	Stand der Forschung	3
1.3	Ziele und Gliederung der Arbeit	6
2	Mechanische Grundlagen: vom Materialpunkt zur FEM	9
2.1	Grundgleichungen der Kontinuumsmechanik	9
2.1.1	Kinematische Beziehungen	9
2.1.2	Spannungsmaße	12
2.2	Das materialunabhängige Randwertproblem der Festkörpermechanik	14
2.2.1	Die Impulserhaltung als starke Form des Gleichgewichts . .	14
2.2.2	Die schwache Form des Randwertproblems	16
2.2.3	Linearisierungen	19
2.3	Finite-Elemente-Methode	20
2.3.1	Isoparametrisches Konzept	20
2.3.2	Diskretisierung der Grundgleichungen	23
2.4	FEM für Schalenstrukturen	25
2.4.1	Kinematik und Variationsformulierung	26
2.4.2	Finite-Elemente Formulierung	28
3	Viskoelastizität	33
3.1	Eindimensionale rheologische Modelle	34
3.2	Thermodynamische Grundlagen und Einordnung der Modelle . . .	39
3.2.1	Energiefunktionen mit inneren Variablen	39
3.2.2	Eindimensionales verallgemeinertes MAXWELL-Modell . .	42
3.2.3	Mehrdimensionale Erweiterung	44
3.3	Algorithmische Umsetzung	45
3.3.1	Eindimensionale Implementierung	45
3.3.2	Mehrdimensionale Implementierung	48
3.4	Zeitabhängige viskoelastische Materialversuche	49

4	Einführung in die Mikromechanik	51
4.1	Das repräsentative Volumenelement als Einheitszelle	52
4.2	Dreidimensionaler Mikro-Makro Übergang	54
4.3	Herleitung mikroskopischer Randbedingungen	56
5	Mikro-Makro Übergang für Schalen	59
5.1	Mikroskopisches Homogenisierungsproblem	60
5.2	Rand- und Nebenbedingungen für Schalen-RVEs	62
5.3	Numerische Homogenisierung	65
6	Künstliche neuronale Netze	71
6.1	Das Feedforward-Netz	73
6.1.1	Transformation der KNN-Daten	75
6.1.2	Die Vorwärtsrechnung	76
6.1.3	Definition des Fehlers und Rückwärtsrechnung	80
6.1.4	Partielle Ableitungen der Ausgangsgrößen	84
6.2	Optimierungsverfahren - Das Training des KNN	86
6.2.1	Das Verfahren des steilsten Abstiegs	88
6.2.2	Der ADAM Optimierungsalgorithmus	89
6.2.3	Das Levenberg-Marquardt-Verfahren	91
6.2.4	Beispiel zu Funktionsapproximation	95
6.2.5	Abschließende Anmerkungen zum KNN-Training	99
6.3	KNN-Training mit Nebenbedingungen	100
6.3.1	Theorie nichtlinearer restringierter Optimierungsprobleme	101
6.3.2	SOBOLEV-Training: Die Tangenten-Nebenbedingung	104
6.3.3	SOBOLEV-Training mit dem Levenberg-Marquardt- Verfahren	106
6.3.4	Beispiel zur Funktionsapproximation mit SOBOLEV- Training	110

7	Konstruktion viskoelastischer Materialmodelle mit KNN	113
7.1	KNN-Architektur für Viskoelastizität	113
7.1.1	Modellierung mit Feedforward-Netzen	114
7.1.2	Konsistenz des Modells mit Grundsätzen der Materialmodellierung	116
7.1.3	Diskussion weiterer KNN Architekturen	117
7.2	Generierung von viskoelastischen Trainingsdaten	118
7.2.1	Definition expliziter Zeitpfade auf Basis von Materialversuchen	119
7.2.2	Eine adaptive Zeitschrittweiten-Strategie	122
7.2.3	Explizite Zeitpfade: Zufällige Datengenerierung	123
7.2.4	Grenzen des viskoelastischen Trainingsraums	125
7.2.5	Synthetische Materialversuche	127
7.2.6	Synthetische Materialversuche mit Schalen-RVEs	129
7.3	Nebenbedingungen für viskoelastisches Materialverhalten	131
7.3.1	Sobolev Training - Einbindung der Materialtangente	131
7.3.2	Diskussion weiterer Nebenbedingungen	132
8	KNN-Untersuchungen auf Materialpunkt- und Strukturebene	134
8.1	Eindimensionale Untersuchungen	134
8.1.1	Einfluss der Datengenerierung	134
8.1.2	Einfluss der adaptiven Zeitschrittweitenstrategie	138
8.1.3	Einfluss unterschiedlichen Verhaltens bei Kriechen und Relaxation	140
8.1.4	Berücksichtigung von Materialparametern als Eingangsvariablen	143
8.1.5	Numerisches Beispiel: Weicher Aktuator mit viskoelastischen Fasern	146
8.2	Dreidimensionale Untersuchungen	151
8.2.1	Konstruktion des Materialmodells auf Basis uniaxialer Versuche	151
8.2.2	Einfluss der Tangenten-Nebenbedingungen	154
8.2.3	Hydrostatische und deviatorische Materialantwort	156

8.2.4	Numerisches Beispiel: Ringförmige Platte mit Schlitz . . .	158
8.3	Modellierung von Schalenstrukturen	160
8.3.1	Schalen-RVE mit Schichtung: Voruntersuchungen am Materialpunkt	161
8.3.2	Numerisches Beispiel: Geschichtete Zylinderschale	165
8.3.3	Numerisches Beispiel: Glasfaserverstärkte Platte	171
8.3.4	Schalen-RVE mit Einschluss: Voruntersuchungen am Materialpunkt	177
8.3.5	Numerisches Beispiel: Hohlkörperdecke	180
9	Zusammenfassung und Ausblick	186
	Literaturverzeichnis	191

1 Einleitung

1.1 Motivation

Wäre zum Zeitpunkt des Verfassens dieser Arbeit die *künstliche Intelligenz* (KI) soweit fortgeschritten, dass sie alle Probleme des Menschen lösen könnte, würde der Prompt (dt.: *Eingabeaufforderung*): „Konstruiere mir ein Materialmodell für viskoelastische Schalenstrukturen“ vermutlich ausreichen, um den Umfang der vorliegenden Arbeit reduzieren zu können. Selbst wenn die KI an diesen Punkt kommen sollte, wären aus Sicht des Autors viele der nachfolgenden Ausführungen dennoch unerlässlich. Assistenzprogramme, wie sie heute in vielerlei Form zur Unterstützung in alltäglichen Aufgaben zu finden sind, entfalten ihr größtes Potenzial, wenn der Anwender das notwendige Hintergrundwissen aufweist, um den Ursprung und die Qualität der Ergebnisse einordnen zu können. Eine thematisch zugehörige Analogie dafür lässt sich direkt in der Baustatik verorten: Seit der Entwicklung der ersten Programme für statische Berechnungen in den späten 1950er und frühen 1960er Jahren haben sich diese als unabdingbare Werkzeuge im Ingenieuralltag etabliert. Die umfassenden Funktionen heutiger Programme können jedoch nur erfolgreich genutzt werden, wenn ein tiefes Verständnis der baustatischen Grundlagen vorhanden ist. Gerade weil moderne Ingenieurbauwerke ohne solche Funktionen kaum denkbar wären, wäre es ein ebenso großer Fehler, die Möglichkeiten der KI in der Baustatik zu ignorieren wie alle statischen Berechnungen ausschließlich händisch vorzunehmen.

Die Baustatik ist dabei heutzutage weitaus mehr als „nur“ statisches Berechnungswesen. Ingenieurbauwerke weisen aufgrund ihrer Größe und Komplexität meist intrinsisch nicht die Möglichkeit zur Untersuchung ihres Verhaltens in Experimenten auf, weshalb die Modellbildung unerlässlich ist. Nicht ohne Grund kamen die Entwickler der erste Pionierarbeiten zur Finite-Elemente-Methode (FEM) aus dem Bauingenieurwesen, wenngleich die Anwendung zunächst in der Luft- und Raumfahrt erfolgte. Getreu AGYRIS' Zitat „*the computer shapes the theory*“ [7] hat sich das Maß an Komplexität, das in FEM-Modellen abgebildet werden kann, stetig weiterentwickelt. Entsprechend vielseitig haben sich die Aufgabengebiete der Baustatik ausgeprägt. Neben der Numerik als Grundbaustein der FEM, sieht man sich heute mit einem großen Spektrum an Themen konfrontiert, welche von der Optimierung von Tragwerken über Kenntnisse moderner hochleistungsfähiger Werkstoffe bis zur Materialmodellierung reichen, um nur einige zu nennen. Gerade für die beiden zuletzt genannten hat sich der Bedarf an numerischen Möglichkeiten zunehmend erhöht, um nicht nur die makroskopischen (*mit bloßen Auge erkennbaren*), sondern auch die mikroskopischen Eigenschaften von Strukturen gezielt erfassen zu können.

Seit Ende der 1990er- und Anfang der 2000er-Jahre haben sich in diesem Kontext sogenannte FE^2 -Methoden entwickelt, wobei exemplarisch die Arbeiten von MICHEL ET AL. [93], FEYEL und CHABOCHE [31] oder MIEHE und KOCH [95] zu nennen sind. Das Quadrat impliziert dabei zwei gekoppelte Modelle: eines, welches die Makro- und ein weiteres, welches die Mikroebene abbildet. In der Theorie wird damit die Notwendigkeit eines vollaufgelösten, dreidimensionalen Modells der betrachteten makroskopischen Struktur umgangen, was bei komplexen Mikrostrukturen – wie beispielsweise Faserverbundwerkstoffen – bis heute eine in Bezug auf die benötigten Rechenzeiten unwirtschaftliche Modellierungsvariante darstellt. Für dünne, biegebeanspruchte Bauteile bietet sich dabei eine FE^2 -Formulierung für Schalen an, wobei hier insbesondere die Arbeiten von GEERS ET AL. [32], COENEN ET AL. [24], GRUTTMANN und WAGNER [45,46] sowie CONG ET AL. [25] hervorzuheben sind. Trotz vielversprechender Entwicklungen hat sich der FE^2 -Ansatz bislang jedoch nicht in der Ingenieurpraxis etabliert, da sich der Berechnungsaufwand zwar reduzieren lässt, dieser sich aber meist weiterhin außerhalb wirtschaftlicher Anwendbarkeit bewegt.

Gleichzeitig schreitet die Entwicklung von hochleistungsfähigen Verbundwerkstoffen, bei denen gezielt – und häufig auf Mikrostrukturebene – verschiedene Materialien eingesetzt werden, in Anbetracht der Nachfrage nach wirtschaftlicheren und nachhaltigeren Lösungen, stetig voran. Das Prinzip ist dabei längst nicht erst seit der Moderne bekannt; historisch wurden schon bei 3 000 Jahre alten ägyptischen Lehmziegeln Verstärkungen mit Stroh oder anderen Pflanzenfasern nachgewiesen (vgl. SCHÜRMANN [120]). Heute finden sich in allen Ingenieurdisziplinen, sei es der Luft- und Raumfahrt, dem Maschinenbau oder dem Bauingenieurwesen, Beispiele für derartige Konstruktionsprinzipien. Dabei können zwei Problemstellungen klassifiziert werden: Eine Mikrostruktur ist vorgegeben und von Interesse ist das resultierende makroskopische Verhalten der Gesamtstruktur. Alternativ können gewisse makroskopische Eigenschaften erwünscht sein, womit sich die Frage nach dem Aufbau der zugrundeliegenden Mikrostruktur stellt. Letzteres wird auch als inverser Designprozess bezeichnet und wird aktuell vor allem im Rahmen sogenannter Metamaterialien untersucht (vgl. z. B. KUMAR ET. AL. [75]). In beiden Fällen ist die Notwendigkeit einer effizienten Methode, den Mikro-Makro Übergang numerisch zu erfassen, nicht von der Hand zu weisen.

Ein möglicher Anwendungsfall im Bauwesen, für den eine effiziente Erfassung des Mikro-Makro-Übergangs von Interesse ist, sind Schalen, die Mikrostrukturen mit viskoelastischen Eigenschaften aufweisen. Dabei kann es sich beispielsweise um faserverstärkte Bauteile, geschichtete Lamine oder Deckenplatten mit Hohlräumen als verlorene Schalung handeln. Im Rahmen des Designprozesses ergeben sich für den Ingenieur daraus typische Fragestellungen wie: Welcher Fasergehalt und welche Steifigkeiten werden in den Schichten benötigt, um unerwünschte Ver-

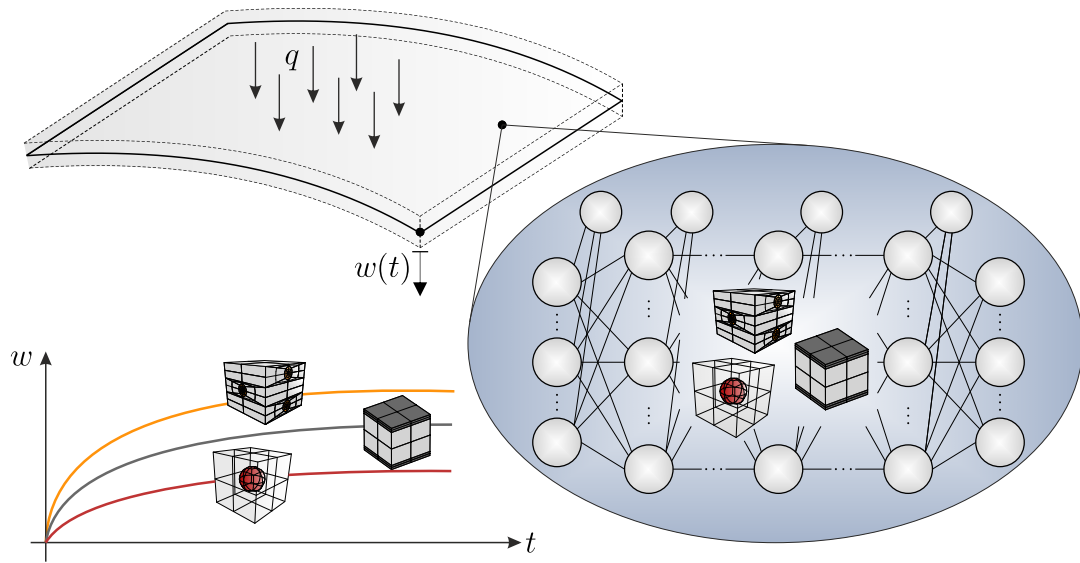


Abbildung 1.1: Schematische Darstellung einer Schale mit einem KNN als Materialmodell, welches die effektiven Eigenschaften der viskoelastischen Mikrostruktur approximiert

formungen der Verbundstruktur zu vermeiden? Wie verhält sich mein Laminat beim Einsatz von Schichten mit zeitabhängigem Materialverhalten? Welche Eigenschaften meiner Schalung reduzieren den Materialverbrauch und gewährleisten gleichzeitig ein gebrauchstaugliches Verhalten der Deckenplatte?

Für diese und weitere Fragestellungen wird im Folgenden eine Methodik entwickelt, die Ingenieuren zukünftig als ein zusätzliches Werkzeug zur Unterstützung des Entwurfsprozesses dienen soll. Im Sinne aktueller Forschungstrends, die zunehmend auf Methoden der künstlichen Intelligenz zurückgreifen, dient dabei ein künstliches neuronales Netz (KNN) als Ersatzmodell für den Mikro-Makro-Übergang. Dies ist schematisch in Abbildung 1.1 dargestellt. Auf Basis generierter Mikrostrukturdaten erlernt das KNN deren effektive Eigenschaften und stellt anschließend im Rahmen numerischer Simulationen ein effizientes Materialmodell dar, das das Materialverhalten in einem Bruchteil der von der FE^2 -Methode benötigten Zeit approximiert. Damit lässt sich unter anderem das Verhalten von Schalen mit viskoelastischen mikrostrukturellen Eigenschaften untersuchen.

1.2 Stand der Forschung

Auch wenn die Anwendung von künstlicher Intelligenz erst in den vergangenen Jahren in der breiten Öffentlichkeit an Relevanz gewonnen hat, beschäftigt sich die Forschung bereits seit mehreren Jahrzehnten mit dieser Thematik. Ein allgemeiner historischer Rückblick hierzu wird in Kapitel 6 gegeben. Erste Ansätze

zum **Einsatz künstlicher neuronaler Netze (KNN) in der Festkörper- und Strukturmechanik** lassen sich bis in die 1990er-Jahre zurückverfolgen. Zu den frühen Arbeiten zählen unter anderem die Beiträge von CHEN ET AL. [20] oder GHABOUSSI und JOGHATAIE [37], in denen KNN zur dynamischen Analyse von Gebäuden eingesetzt wurden. Einer der ersten umfassenden Überblicke über potenzielle Einsatzmöglichkeiten neuronaler Netze – von der Strukturanalyse bis hin zur Modellierung von Plastizität und Schädigung – liefert Waszczyszyn [144]. Die Pionierarbeiten von GHABOUSSI ET AL. [36] legten 1991 mit der Modellierung des biaxialen Spannungs-Dehnungsverhaltens von Beton den Grundstein für die Anwendung neuronaler Netze als Materialmodell. Darauf folgten Arbeiten zur Konstruktion solcher Modelle auf Basis globaler Last-Verschiebungskurven (GHABOUSSI ET AL. [38]) sowie zur Modellierung des Verhaltens von Sand auf Grundlage von Triaxialversuchen (GHABOUSSI und SIDARTA [39]). Für die Anwendung in der Finite-Elemente-Methode (FEM) sind insbesondere die Arbeiten von LEFIK und SCHREFLER [81] sowie von HASHASH ET AL. [50] zu nennen, in denen bereits inelastisches Materialverhalten behandelt wird.

Da die **Materialmodellierung** bis heute eine Disziplin darstellt, in der nach alternativen Ansätzen zu klassischen phänomenologischen Modellen gesucht wird, hat sich der Einsatz von KNN insbesondere in diesem Bereich der Festkörpermechanik zunehmend etabliert. Mit den rasanten Fortschritten in der Entwicklung neuartiger KI-Algorithmen stieg zugleich die Komplexität der für die Materialmodellierung eingesetzten Netzarchitekturen deutlich an. In MASI ET AL. [91] wird erstmals eine thermodynamisch konsistente Architektur vorgestellt, bei der auf Basis der freien Energie und der Dissipation die Zusammenhänge zu Spannungen und internen Variablen berücksichtigt werden. Ein ähnlicher Ansatz findet sich in LINKA ET AL. [84], die mit den sogenannten *Constitutive Artificial Neural Networks* (CANNs) eine vergleichbare Struktur vorschlagen. Für aktuelle Arbeiten, in denen die vielversprechenden *Kolmogorov-Arnold*-Netze verwendet werden wird auf ABDOLAZIZI ET AL. [1] verwiesen. KLEIN ET AL. [68] nutzen *Physics-Augmented Neural Networks* (PANNs) um thermodynamische Konsistenz im elektro-elastischen Materialverhalten sicherzustellen, während KALINA ET AL. [64] *Physics-Constrained Neural Networks* zur Modellierung anisotroper Hyperelastizität einführen (vgl. aktuelle Arbeiten in KALINA ET AL. [63]). In WEBER ET AL. [146] werden physikalische Grundprinzipien für hyperelastisches Materialverhalten durch die Formulierung geeigneter Nebenbedingungen schwach erzwungen. Eine Erweiterung für elasto-plastisches Material findet sich in [145] sowie für *Smart Material Alloys* in [148]. VLASSIS und- SUN [138] integrieren das sogenannte *Sobolev*-Training (CZARNECKI [26]), bei dem zusätzlich die Information der Materialtangente in das Training eingeht – ein Ansatz, der auch in dieser Arbeit Anwendung findet. Für eine umfassende Übersicht verschiedener

KNN-basierter Konstitutivmodelle wird auf DORNHEIM ET AL. [28] verwiesen.

Erste Anwendungen als **Ersatzmodell im Rahmen der Mehrskalenmodellierung** finden sich Ende der 2000er-Jahre in den Arbeiten von UNGER und KÖNKE [136, 137], die entsprechende Ansätze für Stahlbetonbauteile entwickelten. In LEFIK ET AL. [80] werden neuronale Netze zur Vorhersage effektiver Materialeigenschaften von Kompositstrukturen eingesetzt, während HAMBLI ET AL. [49] deren Anwendung im Rahmen der Mehrskalenmodellierung von menschlichem Knochengewebe demonstrieren. LE ET AL. [78] nutzen KNN zur Analyse nichtlinearer elastischer Mikrostrukturen. Einen ersten umfassenden Überblick über die Potenziale von KNN für physikalisch motivierte Mehrskalenprobleme liefern PENG ET AL. [106]. GHAVAMIAN und SIMONE [40] setzen ein rekurrentes neuronales Netz (RNN) als Ersatzmodell für Mikrostrukturen mit geschichtsabhängigem Materialverhalten ein, während ROCHA ET AL. [111] einen probabilistischen KNN-Ansatz für plastisches Verhalten und Risswachstum auf der Mikroebene vorstellen. Ein weiteres Ersatzmodell für heterogene Mikrostrukturen auf Basis von *Convolutional Neural Networks* wird von ALDAKHEEL ET AL. [5] vorgeschlagen. Für erste Untersuchungen zum Einsatz von KNN in der Topologieoptimierung elastischer Mikrostrukturen wird auf WHITE ET AL. [149] verwiesen. SCHOMMARTZ ET AL. [119] untersuchen den Einsatz der PANNs in aktuellen Arbeiten als Ersatzmodell für den Mikro-Makro Übergang von Balken mit hyperelastischen Mikrostrukturen. Eine ausführliche Diskussion aktueller Entwicklungen sowie des Standes der Forschung zur Anwendung neuronaler Netze in der Mehrskalenmodellierung und im Materialdesign findet sich in BISHARA [16].

Ein wesentlicher Bestandteil der vorliegenden Arbeit ist die Approximation **viskoelastischen Materialverhaltens**. In AL-HAIK ET AL. [4] werden KNN zur Approximation des temperaturabhängigen Kriechverhaltens von karbonfaserverstärkten Epoxidverbunden eingesetzt. Aufgrund der Raten- und Geschichtsabhängigkeit viskoelastischer Materialien müssen neben den Dehnungen zusätzliche Variablen im Eingangsvektor des KNN berücksichtigt werden, um allgemeine viskoelastische konstitutive Beziehungen abzubilden. Dies wurde erstmals von JUNG und GHABOUSSI [62] vorgeschlagen. Um die explizite Definition von Geschichtsvariablen zu vermeiden, wird von CHEN [20] ein rekurrentes neuronales Netz (RNN) mit einer Long Short-Term Memory (LSTM)-Struktur verwendet. Einen ersten Überblick über verschiedene KNN-Architekturen zur Modellierung viskoelastischer Materialien geben ROSENKRANZ ET AL. [113]. Die Erweiterung der CANNs auf nichtlineare Viskoelastizität wurde von ABDOLAZIZI [2] vorgenommen. Darüber hinaus schlagen AS'AD und FARHAT [9] eine Architektur vor, welche die thermodynamische Konsistenz über mechanische Zwangsbedingungen sicherstellt. Ein alternativer Modellierungsansatz auf Basis eines Potenzials der freien Energie und eines Dissipationspotenzials findet sich in BENADY ET AL. [14]. Eine aktu-

ell besonders vielversprechende Methodik wird in ROSENKRANZ ET AL. [114] vorgestellt. Hierbei handelt es sich um eine KNN-Architektur, die thermodynamische Konsistenz per Konstruktion gewährleistet und dabei ohne die explizite Definition interner Variablen auskommt. BOZKURT und TAGARIELLI [18] zeigen die Anwendung von KNN als Ersatzmodell zur Approximation effektiver viskoelastischer Eigenschaften in 3D-Mehrskalenberechnungen poröser Elastomere bei großen Verzerrungen.

1.3 Ziele und Gliederung der Arbeit

Ausgehend vom aktuellen Stand der Forschung werden in dieser Arbeit systematisch mehrere – nach bestem Wissen des Autors – bislang offene Fragestellungen untersucht. Das übergeordnete Ziel besteht darin, parametrisierbare viskoelastische Mikrostrukturen für Schalen auf Basis synthetisch gewonnener Daten aus repräsentativen Volumenelementen (RVE) zu approximieren. Die vorliegende Arbeit steht dabei im Zusammenhang mit zwei begutachteten Veröffentlichungen des Autors [34, 35], auf deren Ergebnissen sie teilweise aufbaut und die sie in einen erweiterten methodischen Rahmen einordnet. Im Hinblick dieses Ziels werden chronologisch die folgenden Aspekte behandelt:

- **Datengenerierung auf Basis experimenteller Versuche**
Bestehende Arbeiten verwenden meist zufällige Verzerrungspfade (engl.: *random strain walks*) zur Erzeugung von Trainingsdaten für KNN. Klassische analytische viskoelastische Materialmodelle werden jedoch typischerweise anhand von Kriech-, Relaxations- oder zyklischen Versuchen identifiziert. Im Folgenden wird untersucht, inwiefern sich diese Art von Versuchen zur Datengenerierung für KNN eignet, um sowohl die benötigte Datenmenge zur Modellkonstruktion zu reduzieren als auch die vorgeschlagene Methodik grundsätzlich auf experimentelle Daten anwenden zu können.
- **Verwendung synthetischer Materialversuche zur Konstruktion mehrdimensionaler Materialmodelle**
Im mehrdimensionalen Fall müssen bei der Erzeugung von Trainingsdaten die Verzerrungskomponenten variiert werden. Abhängig von der zugrunde liegenden Eingangsdimension ergeben sich hierbei beliebig viele Kombinationen. Untersucht wird, inwieweit rein uniaxiale synthetische Materialversuche genutzt werden können, um die Anzahl zu variierender Verzerrungskomponenten deutlich zu verringern.
- **Einbindung der Tangenteninformationen in das KNN-Training**
Werden synthetische Daten beispielsweise auf Basis numerischer Homogeni-

sierung mit RVEs erzeugt, liegen neben den effektiven Spannungs-Dehnungs-Daten auch Informationen über die Materialtangente vor. Es wird untersucht, wie sich diese Informationen über geeignete Nebenbedingungen in das KNN-Training integrieren lassen, um die Anzahl erforderlicher FE-Berechnungen zu reduzieren. Zudem wird der Einfluss auf die Robustheit des Trainingsprozesses und die Approximationsgüte des Modells analysiert.

- **Eignung des KNN-Materialmodells in komplexen numerischen Beispielen**

Ein wichtiger Aspekt ist nicht nur die Leistungsfähigkeit eines Materialmodells am Materialpunkt, sondern auch seine Stabilität und Genauigkeit bei der Einbindung in die Finite-Elemente-Analyse. Anhand anspruchsvoller numerischer Beispiele wird daher die Eignung der entwickelten KNN-Modelle überprüft.

- **Einsatz als effektives Materialmodell für Schalen**

Die in dieser Arbeit verwendete Schalenformulierung folgt einer Theorie mit acht arbeitskonformen Spannungs-Dehnungs-Paaren. Dadurch ergibt sich für das KNN ein hochdimensionaler Zusammenhang, insbesondere wenn zusätzliche Variablen zur Abbildung von Raten- und Geschichtsabhängigkeit berücksichtigt werden. Untersucht wird, inwiefern der sogenannte „Fluch der Dimensionalität“ (engl.: *curse of dimensionality*) die Leistungsfähigkeit des KNN begrenzt.

- **Berücksichtigung zusätzlicher Parameter als Eingangsvariablen**

Abschließend wird geprüft, ob der bislang ausschließlich auf physikalischen Größen basierende Eingangsvektor des KNN (Verzerrungen, Spannungen, Zeitinkremente) um geometrische Parameter des RVE erweitert werden kann. Dadurch ließe sich das Verhalten ganzer „Materialfamilien“ mit nur einem einzigen KNN abbilden. Die Anwendbarkeit dieses Ansatzes wird sowohl am Materialpunkt als auch auf Strukturebene untersucht.

Dabei sei angemerkt, dass das Ziel dieser Arbeit *nicht* im vollständigen Ersatz von Mehrskalenmodellen auf Basis konsistenter numerischer Homogenisierung liegt. Nach Auffassung des Autors stellen diese Verfahren nach wie vor ein unverzichtbares Werkzeug zur Analyse komplexer Mikrostrukturen dar, deren Wirtschaftlichkeit sich mit der fortschreitenden Entwicklung moderner Rechenkapazitäten weiter verbessern wird. Künstliche neuronale Netze können jedoch insbesondere im Rahmen des Designprozesses neuartiger Werkstoffe ein wertvolles Hilfsmittel sein, um den Einfluss mikrostruktureller Eigenschaften zunächst effizient und hinreichend genau abschätzen zu können.

Mit dem zuvor genannten Ziel folgt die Arbeit einem Aufbau, der den Leser systematisch von den theoretischen Grundlagen bis hin zur praktischen Umsetzung der Methodik führt.

In **Kapitel 2** werden die materialunabhängigen mechanischen und numerischen Grundlagen erläutert. Damit wird das theoretische Fundament geschaffen, auf dem sowohl die Beziehungen am Materialpunkt als auch die verwendeten Finite-Elemente-Formulierungen des Kontinuums und der Schale aufbauen.

Kapitel 3 widmet sich der Einführung in die Viskoelastizität. Die wesentlichen Aspekte des Materialverhaltens werden anhand einfacher rheologischer Modelle erläutert und in den thermodynamischen Kontext eingeordnet. Zudem werden die Algorithmen vorgestellt, die für die Datengenerierung der ein- und dreidimensionalen KNN-Materialmodelle erforderlich sind.

Kapitel 4 und **Kapitel 5** bilden gemeinsam die Grundlage zur Erläuterung des Mehrskalenzkonzepts für Schalen. Während ersteres die kontinuumsmechanischen Prinzipien der Homogenisierung behandelt, liegt der Schwerpunkt des zweiten Kapitels auf der Kopplung von Schalen auf der Makroebene mit Kontinua auf der Mikroebene. In diesem Zusammenhang wird das numerische Homogenisierungsschema im Detail diskutiert, das die Basis für die Gewinnung synthetischer Materialdaten mittels Schalen-RVEs bildet.

Eine umfassende und weitgehend anwendungsunabhängige Darstellung der Grundlagen künstlicher neuronaler Netze erfolgt in **Kapitel 6**. Aufbauend auf der Funktionsweise eines Feedforward-Netzes werden die implementierten Trainingsalgorithmen erläutert. Ergänzend wird eine theoretische Einführung in die restriktierte nichtlineare Optimierung gegeben, die für die Berücksichtigung der Nebenbedingungen im KNN-Training erforderlich ist. Alle Herleitungen werden durch entsprechende Implementierungsvorschriften ergänzt.

In **Kapitel 7** wird die entwickelte Methodik zur Generierung viskoelastischer Daten beschrieben. Das Kapitel enthält zudem eine Diskussion der verwendeten KNN-Architektur, eine kritische Einordnung gegenüber bestehenden Ansätzen sowie die Beschreibung der Algorithmen zur Erzeugung synthetischer Materialdaten auf Basis dreidimensionaler analytischer Stoffgesetze. Abschließend werden die notwendigen Erweiterungen bei der Verwendung von Schalen-RVEs vorgestellt.

Kapitel 8 fokussiert sich schließlich auf die systematische Beantwortung der in dieser Arbeit formulierten Forschungsfragen. Beginnend mit eindimensionalen Analysen wird zunächst der Einfluss der Datengenerierung untersucht, bevor die Methodik auf dreidimensionale und abschließend auf komplexe Schalenprobleme angewendet wird.

2 Mechanische Grundlagen: vom Materialpunkt zur FEM

Im Folgenden werden die notwendigen mechanischen und numerischen Prinzipien, die im Rahmen der Arbeit Verwendung finden, erläutert. Die Ausführungen sind auf die wesentlichen Aspekte beschränkt, der interessierte Leser wird an entsprechende Fachliteratur verwiesen. Die materialunabhängigen Gleichungen der *Kontinuumsmechanik* in Form der kinematischen Beziehungen und des Gleichgewichts werden zur Formulierung des *Randwertproblems der Festkörpermechanik* benötigt. Dessen allgemeine numerische Lösung ohne explizite Materialgleichung wird abschließend unter der Verwendung des Verfahrens der *Methode der Finiten Elemente* diskutiert.

2.1 Grundgleichungen der Kontinuumsmechanik

Die Kontinuumsmechanik befasst sich mit der Bewegung von Körpern unter der Einwirkung von Kräften und den daraus resultierenden Deformationen. Aus diesen wiederum lassen sich Beziehungen für die Verzerrungszustände sowie unterschiedliche Spannungsmaße ableiten. Im Folgenden werden lediglich Festkörper betrachtet, die massenerhaltenden und isothermen Prozessen unterliegen. Des Weiteren werden quasi-statische Einwirkungen vorausgesetzt, was die Vernachlässigung von Trägheitseffekten ermöglicht. Umfassende Beschreibungen, auf denen die nachfolgenden Ausführungen basieren, sind in OGDEN [103], MARSDEN und HUGHES [90], TRUESDELL und NOLL [134] und HOLZAPFEL [55] zu finden.

2.1.1 Kinematische Beziehungen

Abstrakt entspricht ein Körper einer abgeschlossenen Region $\mathcal{B}$ des dreidimensionalen Euklidischen Punktraumes. Die Punkte des Raumes, die dem Körper zugeordnet werden können, heißen *materielle Punkte* oder *Partikel*. Definiert sei X als ein Partikel des Körpers in einer Referenzkonfiguration $\mathcal{B}_0$ und $\{\mathbf{e}_i\}$, $i = 1, 2, 3$, eine zugehörige orthonormale Basis, dann ist $\mathbf{X} = X_i \mathbf{e}_i$ der Ortsvektor zum Partikel X . Mit der Bewegung des Körpers ändert sich seine Konfiguration und die Lage des Partikels. Sei $t \in I \subset \mathbb{R}$ die Zeit im Intervall I und $\mathcal{B}_0$ die Konfiguration des Körpers zum Zeitpunkt $t = t_0$, so beschreibt der Ortsvektor $\mathbf{x} = x_i \mathbf{e}_i$ die neue Lage des Partikels x auf einer Konfiguration $\mathcal{B}_t$ zum Zeitpunkt $t > t_0$. $\mathcal{B}_t$ wird im Folgenden als *Momentankonfiguration* des Körpers bezeichnet. Zur Vereinfachung wird im Weiteren davon ausgegangen, dass die Referenzkonfiguration der initialen

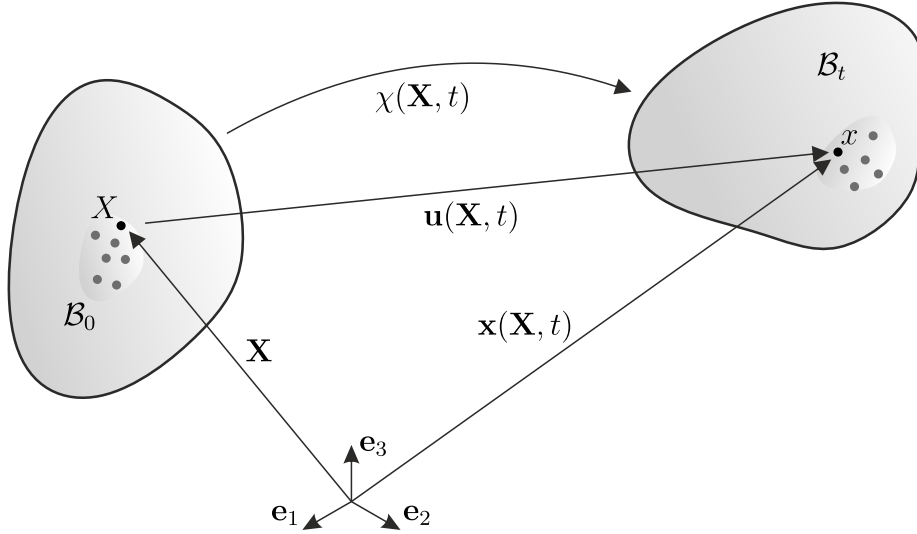


Abbildung 2.1: Materieller Körper und ausgewähltes Partikel in seiner Referenz- und deformierten Momentankonfiguration

und undeformierten Konfiguration des Körpers zum Zeitpunkt $t = 0$ entspricht¹. Die Abbildung von Punkten von der Referenz- in die Momentankonfiguration bezeichnet die Bahn $\chi(\mathbf{X}, t)$ womit folgt

$$\mathbf{x} = \chi(\mathbf{X}, t). \quad (2.1)$$

Alternativ zu der Bahn $\chi(\mathbf{X}, t)$ kann mittels des Verschiebungsfelds $\mathbf{u}(\mathbf{X}, t)$ jedem Partikel X eindeutig seine Position in der Momentankonfiguration zugeordnet werden:

$$\mathbf{u}(\mathbf{X}, t) = \mathbf{x}(\mathbf{X}, t) - \mathbf{X}. \quad (2.2)$$

Die Zusammenhänge sind graphisch in Abbildung 2.1 dargestellt.

Bislang wurde lediglich die Bewegung der einzelnen Partikel beschrieben. Von Interesse ist jedoch auch das Verhalten von benachbarten Partikeln in den jeweiligen Konfiguration B_0 und B_t , welches die Grundlage für die Definition von Verzerungsmaßen bildet. Hierfür werden die Linienlemente $d\mathbf{X}$ und $d\mathbf{x}$ eingeführt. $d\mathbf{X}$ und $d\mathbf{x}$ beschreiben damit die Differenzen der Ortsvektoren zweier infinitesimal benachbarter Partikel X und Y in den jeweiligen Konfigurationen. Die Betrachtung der Deformation in der Umgebung eines materiellen Punktes erfolgt über das totale Differential

$$d\mathbf{x} = \frac{\partial \mathbf{x}(\mathbf{X}, t)}{\partial \mathbf{X}} d\mathbf{X} = \mathbf{F}(\mathbf{X}, t) d\mathbf{X} \quad (2.3)$$

¹Theoretisch ist die Wahl der Referenzkonfiguration beliebig und muss auch nicht vom Körper eingenommen werden. Dies wird im Rahmen der Arbeit nicht weiter diskutiert.

Dabei ist

$$\mathbf{F}(\mathbf{X}, t) := \frac{\partial \mathbf{x}}{\partial \mathbf{X}} \quad (2.4)$$

der materielle Deformationsgradient, über den die Abbildung von Linienelementen von der Referenz- in die Momentankonfiguration erfolgt. Auf die explizite Benennung der Abhängigkeit von $\mathbf{F}$ von den räumlichen Koordinaten $\mathbf{X}$ und der Zeit t wird im Folgenden verzichtet. Aus der Bedingung $\mathbf{F} d\mathbf{X} \neq \mathbf{0}$ ergibt sich direkt für $d\mathbf{X} \neq 0$, dass gelten muss $J = \det \mathbf{F} \neq 0$, was die Elimination von Linienelementen ausschließt. Daraus folgt auch die Existenz der Inversen von $\mathbf{F}$, dem räumlichen Deformationsgradienten $\mathbf{F}^{-1}$. Aus der Ableitung der Gleichung 2.2 nach den räumlichen Koordinaten folgt der Verschiebungsgradient

$$\mathbf{H} := \frac{\partial \mathbf{u}}{\partial \mathbf{X}} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}} - \mathbf{1}. \quad (2.5)$$

Damit lässt sich der materielle Deformationsgradient auch schreiben als

$$\mathbf{F} = \mathbf{H} + \mathbf{1}. \quad (2.6)$$

Die Transformation von Flächenelementen erfolgt mit der Nanson Formel

$$\mathbf{n} da = J \mathbf{F}^{-T} \mathbf{N} dA, \quad (2.7)$$

wobei $\mathbf{n}$ und $\mathbf{N}$ die jeweils den Flächenelementen der Momentan- und Referenzkonfiguration da und dA zugeordneten Normalenvektoren darstellen. Letztlich erfolgt die Transformation von infinitesimalen Volumenelementen mit der Beziehung

$$dv = J dV. \quad (2.8)$$

Mit der Konvention, dass Volumina stets positiv sein müssen, folgt die Bedingung $J > 0$. Sonderfälle der Volumentransformation stellen Starrkörperverschiebungen und isochore, oder volumenerhaltende, Deformationen dar, für die gilt $J = 1$. Dies kommt beispielsweise bei inkompressiblem Werkstoffverhalten zum Tragen.

Für die Einführung eines geeigneten Verzerrungsmaßes wird die quadratische Längenänderung zweier Linienelemente unter Verwendung von Gleichung (2.3)

$$d\mathbf{x}^2 - d\mathbf{X}^2 = d\mathbf{X}(\mathbf{F}^T \mathbf{F} - \mathbf{1}) d\mathbf{X} \quad (2.9)$$

betrachtet. Die Anwendung des Quadrats garantiert Verzerrungsfreiheit sowohl für reine Starrkörperverschiebungen als auch -rotationen bei denen $\mathbf{C} := \mathbf{F}^T \mathbf{F} = \mathbf{1}$ gilt. $\mathbf{C}$ ist dabei der rechte CAUCHY-GREEN Verzerrungstensor. Im Rahmen der Arbeit wird meist der mit Gleichung (2.9) hergeleitete GREEN-LAGRANGE Verzerrungstensor

$$\mathbf{E} = \frac{1}{2}(\mathbf{F}^T \mathbf{F} - \mathbf{1}) \quad (2.10)$$

verwendet. Er ist symmetrisch und mit Gl. 2.6 folgt

$$\mathbf{E} = \frac{1}{2}(\mathbf{H} + \mathbf{H}^T + \mathbf{H}^T \mathbf{H}), \quad (2.11)$$

was eine Berechnung mit dem Verschiebungsgradienten ermöglicht. Im Rahmen kleiner Verformungen gilt die Annahme $\mathbf{H}^T \mathbf{H} \ll \mathbf{H}$, wodurch die Bedeutung des Faktors $1/2$ deutlich wird, um die Konsistenz zu den Ingenieurverzerrungen

$$\boldsymbol{\varepsilon} := \frac{1}{2}(\mathbf{H} + \mathbf{H}^T) \quad (2.12)$$

zu gewährleisten. Es sei bereits erwähnt, dass die Ingenieurverzerrungen im Rahmen der Arbeit kaum Anwendung finden und $\boldsymbol{\varepsilon}$ später für die nichtlinearen Schalenverzerrungen verwendet wird. Weiterhin werden alle Zusammenhänge im nichtlinearen Verzerrungsmaß $\mathbf{E}$ hergeleitet, da dies den allgemeinen Fall darstellt, welcher im weiteren Verlauf für die Schalenformulierung benötigt wird.

2.1.2 Spannungsmaße

Neben den zuvor betrachteten Verzerrungsfeldern führen äußere, auf den Körper einwirkende Kräfte zur Ausbildung von Spannungsfeldern. Diese können in gleicher Weise wie die kinematischen Beziehungen in den jeweiligen Konfigurationen definiert werden. Um das Konzept der Spannungen abzuleiten, wird der materielle Körper in den jeweiligen Konfigurationen $\mathcal{B}_0$ und $\mathcal{B}_t$ gedanklich durch eine Ebene in zwei gegenüberliegende Schnittflächen geteilt, s. Abbildung 2.2. Auf den Körper wirken auf den Rändern $\partial\mathcal{B}_0$ und $\partial\mathcal{B}_t$ äußere Kräfte, während durch den gedanklichen Schnitt innere Kräfte in den Schnittflächen freigelegt werden. Es folgt der Zusammenhang zwischen einem freigelegten räumlichen Kraftinkrement $d\mathbf{f}$ und den Spannungsvektoren $\mathbf{t}$ und $\mathbf{T}$

$$\mathbf{t} := \lim_{\Delta a \rightarrow 0} \frac{d\mathbf{f}}{\Delta a} = \frac{d\mathbf{f}}{da} \quad \text{und} \quad \mathbf{T} := \lim_{\Delta A \rightarrow 0} \frac{d\mathbf{f}}{\Delta A} = \frac{d\mathbf{f}}{dA} \quad (2.13)$$

mit den finiten Flächenelementen der Momentan- Δa und Referenzkonfiguration ΔA . Entsprechend gilt $\mathbf{t} = \mathbf{t}(\mathbf{x}, t, \mathbf{n})$ sowie $\mathbf{T} = \mathbf{T}(\mathbf{X}, t, \mathbf{N})$. Benötigt wird nachfolgend das von CAUCHY postulierte Theorem, welches besagt, dass der Spannungsvektor linear von den Normalenvektoren der entsprechenden Schnittflächen abhängt. Es existieren somit in den jeweiligen Konfigurationen Tensorfelder zweiter Stufe, sodass gilt

$$\mathbf{t}(\mathbf{x}, t, \mathbf{n}) = \boldsymbol{\sigma}^T(\mathbf{x}, t) \mathbf{n} \quad \text{und} \quad \mathbf{T}(\mathbf{x}, t, \mathbf{N}) = \mathbf{P}^T(\mathbf{X}, t) \mathbf{N}. \quad (2.14)$$

Dabei ist $\boldsymbol{\sigma}^T = \boldsymbol{\sigma}$ der räumliche und symmetrische CAUCHY Spannungstensor und $\mathbf{P}^T$ der nicht-symmetrische erste PIOLA-KIRCHOFF'sche Spannungstensor. $\mathbf{P}$

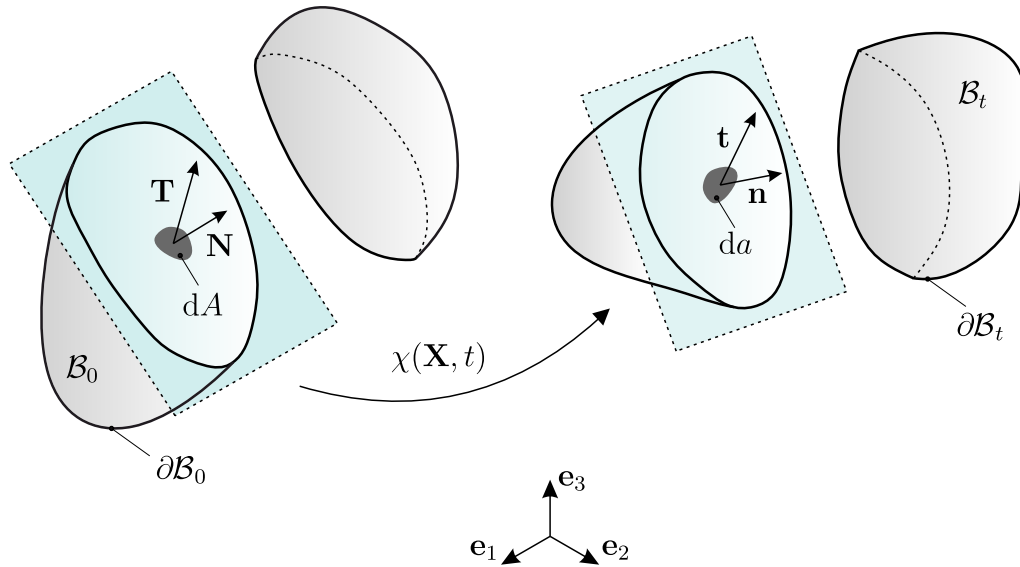


Abbildung 2.2: Darstellung der an den Schnittflächen wirkenden Spannungs- und zugehörigen Normalenvektoren in den jeweiligen Konfigurationen

wird auch der *nominale* Spannungstensor genannt, wobei im Ingenieurwesen häufig auch die Bezeichnungen *wahre Spannung* für $\boldsymbol{\sigma}$ und *technische Spannung* für $\mathbf{P}$ geläufig sind. Die Darstellung von der Gleichung (2.14) in den transponierten Größen lässt sich wie folgt erläutern: Mit $\mathbf{n}_i = \mathbf{e}_i$ sei der zugehörige Spannungsvektor $\mathbf{t}_i$ durch die Komponenten des Spannungsvektors $\boldsymbol{\sigma}$ gegeben, sodass der erste Index die Richtung der Normalen und der zweite Index die Richtung der Spannungskomponente vorgibt:

$$\mathbf{t}_i = \boldsymbol{\sigma}^T \mathbf{e}_i = \begin{bmatrix} \sigma_{i1} \\ \sigma_{i2} \\ \sigma_{i3} \end{bmatrix}. \quad (2.15)$$

Mit den Gleichungen (2.13) und (2.14) sowie der NANSON Formel aus Gleichung (2.7) folgen die Transformationsbeziehungen

$$\mathbf{P} = J \boldsymbol{\sigma} \mathbf{F}^{-1} \quad \text{und} \quad \boldsymbol{\sigma} = \frac{1}{J} \mathbf{P} \mathbf{F}. \quad (2.16)$$

Da im Rahmen der Festkörpermechanik die Momentankonfiguration unbekannt ist, ist es wünschenswert, die Zusammenhänge auf Größen der Referenzkonfiguration zu beziehen. Weiterhin werden symmetrische Tensoren in der Regel bevorzugt. Da die zu dem Tensor $\mathbf{P}$ gehörige Basis, analog zu $\mathbf{F}$, sowohl räumliche, also auch materielle Koordinaten beschreibt, ist dieser nicht symmetrisch. Dies ist der Tatsache geschuldet, dass die Herleitung der Spannungsvektoren in Gleichung (2.13) auf der Basis eines räumlichen Kraftinkrements $d\mathbf{f}$ erfolgte. Wird

dieses zurück auf die Referenzkonfiguration transformiert, folgt der symmetrische zweite PIOLA-KIRCHOFFSCHE Spannungstensor

$$\mathbf{S}^T := \mathbf{P}^T \mathbf{F}^{-1}, \quad \mathbf{S}^T = \mathbf{S}. \quad (2.17)$$

Weiterhin gelten die Beziehungen

$$\mathbf{S} = J \mathbf{F}^{-1} \boldsymbol{\sigma} \mathbf{F}^{-T}, \quad \boldsymbol{\sigma} = \frac{1}{J} \mathbf{F} \mathbf{S} \mathbf{F}^T. \quad (2.18)$$

Die Bedeutung von $\mathbf{S}$ - dem keine physikalische Interpretation zugeordnet werden kann - wird im weiteren Verlauf der Arbeit noch deutlich. Abschließend sei angemerkt, dass im Falle infinitesimaler Verzerrungen keine Unterscheidung zwischen den Spannungsmaßen notwendig ist, d. h. $\mathbf{S} \approx \mathbf{P} \approx \boldsymbol{\sigma}$. Insbesondere bei strukturmekanischen Problemstellungen treten jedoch häufig große Verschiebungen und Verdrehungen auf, sodass geometrisch nichtlineare Verzerrungsmaße erforderlich sind, obwohl die tatsächlichen Verzerrungen klein bleiben.

2.2 Das materialunabhängige Randwertproblem der Festkörpermechanik

Mit den zuvor erläuterten kontinuumsmechanischen Zusammenhängen wird im Folgenden das Randwertproblem des materiellen Körpers hergeleitet. Dieses ist zunächst unvollständig, da dessen Lösung eine konsitutive Beziehung voraussetzt. Die starke Form der Impulserhaltung - die *Gleichgewichtsbedingung* - liefert zunächst die dem Randwertproblem zugrundeliegende Differentialgleichung in den Spannungen. Diese wird unter Verwendung des Variationsprinzips in die schwache Form überführt und als Randwertproblem formuliert, welches die Grundlage für die numerische Lösung mittels der Finiten-Elemente-Methode in Kapitel 2.3 darstellt. Abschließend wird ein iteratives Lösungsverfahren eingeführt, da es sich im Allgemeinen um eine nichtlineare Gleichung in den gesuchten Feldgrößen $\mathbf{u}(\mathbf{X}, t)$ handelt.

2.2.1 Die Impulserhaltung als starke Form des Gleichgewichts

Neben den kinematischen Beziehungen stellen die *physikalischen Erhaltungssätze*² den zweiten Satz an materialunabhängigen Gleichungen dar. In dieser Arbeit werden ausschließlich isotherme und geschlossene Systeme betrachtet. Geschlossen

²Erhaltungssätze stellen Forderungen aus Symmetriegruppen des Systems dar, bzw. gilt dieser Zusammenhang auch umgekehrt s. SALETAN [118]. So kann die Impulserhaltung als translatorische Symmetrie und die Drehimpulserhaltung als Rotationssymmetrie gesehen werden. Die Invarianz des Systems bezüglich der Zeit führt zur Energieerhaltung.

impliziert dabei, dass die skalarwertige Masse $m(\mathcal{B})$ unabhängig von der Konfiguration des Körpers $\mathcal{B}$ ist. Des Weiteren wird davon ausgegangen, dass die Masse kontinuierlich verteilt ist, was die Existenz eines Feldes $\rho(\mathbf{x}, t)$ voraussetzt, sodass

$$m(\mathcal{B}) = \int_{\mathcal{B}} \rho(\mathbf{x}, t) dv. \quad (2.19)$$

$\rho(\mathbf{x}, t)$ ist dabei die *Massendichte* einer Momentankonfiguration. Die Voraussetzung der Unabhängigkeit der Masse bezüglich der Konfiguration führt auf die Forderung

$$\int_{\mathcal{B}_t} \rho(\mathbf{x}, t) dv = \int_{\mathcal{B}_0} \rho_0(\mathbf{X}) dV. \quad (2.20)$$

mit der *Massendichte* $\rho_0(\mathbf{X})$ einer entsprechenden Referenzkonfiguration des Körpers $\mathcal{B}_0$. Dies wird als *Massenerhaltung* bezeichnet. Damit gilt für die Beziehung der Massendichten

$$\rho_0(\mathbf{X}) = J\rho(\mathbf{x}, t). \quad (2.21)$$

Für die Herleitung der Bewegungsgleichung des Körpers ist die *Impulserhaltung* notwendig. Sei $\mathbf{v}(\mathbf{x}, t)$ das Geschwindigkeitsfeld des Körpers $\mathcal{B}_t$ und $\mathcal{L}(t)$ dessen Impuls mit

$$\mathcal{L}(t) = \int_{\mathcal{B}_t} \rho(\mathbf{x}, t) \mathbf{v}(\mathbf{x}, t) dv, \quad (2.22)$$

dann folgt für die Erhaltung des Impulses

$$\dot{\mathcal{L}}(t) = \frac{d}{dt} \int_{\mathcal{B}_t} \rho(\mathbf{x}, t) \mathbf{v}(\mathbf{x}, t) dv = \int_{\mathcal{B}_t} \rho(\mathbf{x}, t) \dot{\mathbf{v}}(\mathbf{x}, t) dv = \mathbf{F}(t). \quad (2.23)$$

$\mathbf{F}(t)$ ist dabei die Summe der auf den Körper wirkenden Kräfte, die im Gleichgewicht mit der Änderung des Impulses stehen. Diese setzt sich diese aus den Körperlasten $\mathbf{b} = \rho \bar{\mathbf{b}}$ und den auf dem Rand wirkenden Oberflächenlasten $\mathbf{t}$ zusammen

$$\mathbf{F}(t) = \int_{\mathcal{B}_t} \mathbf{b} dv + \int_{\partial \mathcal{B}_t} \mathbf{t} da, \quad (2.24)$$

Ein Beispiel für $\bar{\mathbf{b}}$ ist die Gravitationsbeschleunigung, sodass die Körperlast gerade dem Eigengewicht entspricht. Bei Betrachtung von quasi-statischen Prozessen kann die Änderung der Beschleunigung $\dot{\mathbf{v}}(\mathbf{x}, t)$ vernachlässigt werden, womit für Gleichung (2.23) mit Gleichung (2.24) folgt

$$\int_{\mathcal{B}_t} \mathbf{b} dv + \int_{\partial \mathcal{B}_t} \mathbf{t} da = \mathbf{0}. \quad (2.25)$$

Angemerkt sei jedoch, dass quasi-statisch nur die Annahme für die Beschleunigung $\dot{\mathbf{v}} \approx 0$ trifft, aber $\mathbf{v} \neq 0$ gilt, was z.B. bei ratenabhängigem Materialverhalten

relevant wird. Weiterhin lässt sich mit dem CAUCHY Theorem (2.14) und der Anwendung des Gaußschen-Integralsatzes das Oberflächen- in ein Volumenintegral überführen

$$\int_{\mathcal{B}_t} \mathbf{b} \, dv + \int_{\mathcal{B}_t} \operatorname{div} \boldsymbol{\sigma}^T \, dv = \mathbf{0}. \quad (2.26)$$

Da Gleichung (2.26) für beliebige Volumen des Körpers dv gelten muss, folgt die lokale Form der Impulsbilanz

$$\mathbf{b} + \operatorname{div} \boldsymbol{\sigma}^T = \mathbf{0}. \quad (2.27)$$

Analog lässt sich diese auch in der Referenzkonfiguration des Körpers formulieren

$$\mathbf{B} + \operatorname{div} \mathbf{P}^T = \mathbf{0}, \quad (2.28)$$

mit $\mathbf{B} = \rho_0 \bar{\mathbf{b}}$. Gleichung (2.28) wird im Folgenden als *starke Form des Gleichgewichts* bezeichnet und stellt die zugrundeliegende Differentialgleichung des Randwertproblems in den Spannungen auf der Referenzkonfiguration dar. Die Lösung des Problems unter Beachtung von Randbedingungen setzt Kenntnis über die Materialgleichungen voraus, welche die kinematischen Beziehungen mit den Kräften in Relation setzt. Selbst dann stellen geschlossene Lösungen im linearen, und im nichtlinearen ohnehin, Sonderfälle dar. Daher wird die Differentialgleichung im Folgenden in eine geeignete Form überführt, um die Lösung dieser mittels der Numerik zu ermöglichen.

Der Vollständigkeit halber sei abschließend angemerkt, dass weitere Erhaltungssätze existieren, die an dieser Stelle nicht ausführlich beschrieben werden. So führt beispielsweise die globale Form der *Drehimpulsbilanz* in der Momentanconfiguration auf die Symmetrie des CAUCHY Spannungstensors $\boldsymbol{\sigma}^T = \boldsymbol{\sigma}$. Die *Energieerhaltung* wird im späteren Verlauf der Arbeit diskutiert.

2.2.2 Die schwache Form des Randwertproblems

Im Zentrum dieses Kapitels steht der materielle Körper in seiner Referenzkonfiguration mit Körper- und Oberflächenlasten, s. Abbildung 2.3. Für die Formulierung eines Randwertproblems sind Randbedingungen erforderlich. Unterschieden wird dabei in DIRICHLET Randbedingungen für das Verschiebungsfeld $\mathbf{u}(\mathbf{X}, t)$ und NEUMANN Randbedingungen auf denen Oberflächenlasten $\mathbf{T}(\mathbf{X}, t, \mathbf{N})$ vorgegeben sind. Erstere wirken auf dem Verschiebungsrand $\partial\mathcal{B}_{0u}$ und letztere auf dem Spannungsrand $\partial\mathcal{B}_{0\sigma}$ für die gilt

$$\partial\mathcal{B}_{0u} \cup \partial\mathcal{B}_{0\sigma} = \mathcal{B}_0 \quad \text{und} \quad \partial\mathcal{B}_{0u} \cap \partial\mathcal{B}_{0\sigma} = \emptyset. \quad (2.29)$$

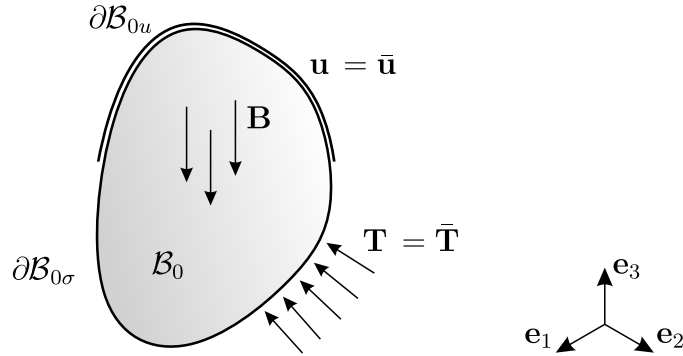


Abbildung 2.3: Darstellung des Randwertproblems auf der Referenzkonfiguration

Die starke Form des Randwertproblems lautet damit

$$\begin{aligned} \operatorname{div} \mathbf{P}^T + \mathbf{B} &= \mathbf{0}, \\ \mathbf{u} &= \bar{\mathbf{u}} \quad \text{auf } \partial \mathcal{B}_{0u}, \\ \mathbf{T} = \mathbf{P}^T \mathbf{N} &= \bar{\mathbf{T}} \quad \text{auf } \partial \mathcal{B}_{0\sigma}, \end{aligned} \quad (2.30)$$

wobei die vorgegebenen Größen auf dem Rand mit $(\bullet)$ gekennzeichnet sind. Unabhängigkeit der Tatsache, dass eine Definition des konstitutiven Zusammenhangs aussteht, muss zur numerischen Lösung die Gleichgewichtsbedingung $(2.30)_1$ in die schwache Form überführt werden. Für vertiefende Ausführungen zu den Vorteilen der schwachen Formulierung wird auf MARSDEN [90] und WASHIZU [143] verwiesen. Die nachfolgenden Herleitungen orientieren sich an HOLZAPFEL [55] und WRIGGERS [152]. Für die Überführung in die schwache Form wird die Impulsbilanz aus Gleichung (2.30) skalar mit einer Testfunktion multipliziert und über das Gebiet integriert

$$\int_{\mathcal{B}_0} (\mathbf{B} + \operatorname{div} \mathbf{P}^T) \cdot \delta \mathbf{u} \, dV = \mathbf{0}. \quad (2.31)$$

Die Wahl der Testfunktion ist grundsätzlich beliebig, wobei die Einhaltung der DIRICHLET Randbedingungen auf $\partial \mathcal{B}_{0u}$ beachtet werden muss. Die Anwendung partieller Integration auf den ersten Term

$$\operatorname{div} \mathbf{P}^T \cdot \delta \mathbf{u} = \operatorname{div}(\mathbf{P}^T \delta \mathbf{u}) - \mathbf{P}^T : \operatorname{grad} \delta \mathbf{u} \quad (2.32)$$

und nachfolgender Anwendung des Divergenztheorems unter Berücksichtigung der Spannungsrandbedingungen

$$\int_{\mathcal{B}_0} \operatorname{div}(\mathbf{P}^T \delta \mathbf{u}) \, dV = \int_{\partial \mathcal{B}_0} (\mathbf{P}^T \delta \mathbf{u}) \cdot \mathbf{N} \, dA = \int_{\partial \mathcal{B}_{0\sigma}} \bar{\mathbf{T}} \cdot \delta \mathbf{u} \, dA \quad (2.33)$$

liefert die schwache Form des Gleichgewichts

$$\delta\Pi(\mathbf{u}, \delta\mathbf{u}) := \underbrace{\int_{\mathcal{B}_0} \mathbf{P}^T : \text{grad } \delta\mathbf{u} \, dV}_{\delta\Pi_i} - \underbrace{\int_{\mathcal{B}_0} \mathbf{B} \cdot \delta\mathbf{u} \, dV - \int_{\partial\mathcal{B}_{0\sigma}} \bar{\mathbf{T}} \cdot \delta\mathbf{u} \, dA}_{\delta\Pi_{\text{ext}}} \stackrel{!}{=} 0, \quad (2.34)$$

auch *Prinzip der virtuellen Verschiebungen* genannt. Hierbei lassen sich die Anteile der virtuellen inneren $\delta\Pi_i$ und äußeren $\delta\Pi_{\text{ext}}$ Arbeiten identifizieren. Falls $\Pi(\mathbf{u})$ Potentialcharakter hat, lässt sich Gleichung (2.34) auch aus dem Prinzip des Minimums des Gesamtpotentials $\Pi(\mathbf{u}) \rightarrow \min$ ableiten. Mathematisch wird hierfür die erste Variation, auch verallgemeinerte Richtungsableitung, in Richtung von $\delta\mathbf{u}$

$$\delta\Pi = D\Pi(\mathbf{u}) \cdot \delta\mathbf{u} = \left. \frac{d}{d\alpha} \Pi(\mathbf{u} + \alpha \cdot \delta\mathbf{u}) \right|_{\alpha=0} \stackrel{!}{=} 0 \quad (2.35)$$

zu Null gesetzt, was äquivalent zur schwachen Form des Gleichgewichts ist. Während Gleichung (2.35) nicht zwingend für die Herleitung der schwachen Form notwendig ist, ist diese dennoch für die Herleitung der variationellen Beziehungen der verschiedenen Tensoren vonnöten. So gilt beispielsweise

$$\delta\mathbf{F} = \text{grad } \delta\mathbf{u}. \quad (2.36)$$

Dies lässt sich wiederum für eine Formulierung der schwachen Form in den symmetrischen Tensoren $\mathbf{S}$ und $\mathbf{E}$, bzw. $\delta\mathbf{E}$ nutzen. Mit der Gleichung (2.36) erhält man die 1. Variation des GREEN-LAGRANGE Verformungstensors

$$\delta\mathbf{E} = \frac{1}{2}(\delta\mathbf{F}^T \mathbf{F} + \mathbf{F}^T \delta\mathbf{F}) = \frac{1}{2}((\text{grad } \delta\mathbf{u})^T \mathbf{F} + \mathbf{F}^T \text{grad } \delta\mathbf{u}). \quad (2.37)$$

Unter Berücksichtigung von der Gleichung (2.17) lässt sich damit die Umformung

$$\mathbf{P}^T : \text{grad } \delta\mathbf{u} = \mathbf{S} : \mathbf{F}^T \text{grad } \delta\mathbf{u} = \mathbf{S} : \frac{1}{2}(\mathbf{F}^T \text{grad } \delta\mathbf{u} + (\text{grad } \delta\mathbf{u})^T \mathbf{F}) = \mathbf{S} : \delta\mathbf{E} \quad (2.38)$$

vornehmen, unter Ausnutzung der Tatsache, dass das Produkt eines symmetrischen Tensors mit dem antisymmetrischen Teil eines beliebigen Tensors 2. Stufe gleich Null ist. Es folgt durch Einsetzen in Gleichung (2.34)

$$\delta\Pi(\mathbf{u}, \delta\mathbf{u}) = \int_{\mathcal{B}_0} \mathbf{S} : \delta\mathbf{E} \, dV - \int_{\mathcal{B}_0} \mathbf{B} \cdot \delta\mathbf{u} \, dV - \int_{\partial\mathcal{B}_{0\sigma}} \bar{\mathbf{T}} \cdot \delta\mathbf{u} \, dA = 0. \quad (2.39)$$

Zusammen mit den Verschiebungs- und Spannungsrandbedingungen stellt diese Formulierung der Gleichgewichtsbedingung die Grundlage für die numerische Lösung dar. Im linearen Fall kann im Rahmen derer durch Einsetzen eines geeigneten Stoffgesetzes sowie der Kinematik und anschließend Einbau der Randbedingungen direkt nach den Verschiebungen $\mathbf{u}$ aufgelöst werden. Im nichtlinearen Fall muss zunächst eine Linearisierung der schwachen Form erfolgen, deren Lösung iterativ bestimmt werden kann.

2.2.3 Linearisierungen

Die Lösung nichtlinearer Gleichungen erfordert die Linearisierung eben dieser, um anschließend eine iterative Lösung zu ermöglichen. Die Vorteile der Linearisierung von Funktionen sind dabei nicht beschränkt auf die Lösung etwaiger. Gerade bei Anwendungen im Ingenieurwesen kommen Linearisierungen häufig zum Einsatz. So z. B. unter Annahme von kleinen Verformungen, was die Vereinfachung der trigonometrischen Beziehungen zulässt. Die grundsätzliche Idee besteht darin, dass sich eine stetige und einmal stetig differenzierbare nichtlineare Funktion $f(\mathbf{x})$ mit $\mathbf{x} \in \mathbb{R}^n$ in ihrer unmittelbaren Umgebung $\bar{\mathbf{x}} + \Delta\mathbf{x}$ mit $\lim_{\Delta\mathbf{x} \rightarrow 0}$ allein durch die Auswertung der Funktion selbst und ihrer Tangente approximativ linear in $\Delta\mathbf{x}$ darstellen lässt. $\bar{\mathbf{x}}$ ist dabei die bekannte Entwicklungsstelle und $\Delta\mathbf{x}$ der inkrementelle Zuwachs, oder auch die unabhängige Variable. Damit sei die Definition der Linearisierung von $f(\mathbf{x})$

$$L[f(\mathbf{x})]_{\mathbf{x}=\bar{\mathbf{x}}} = f(\bar{\mathbf{x}}) + \Delta f(\mathbf{x}) = f(\bar{\mathbf{x}}) + Df(\bar{\mathbf{x}}) \cdot \Delta\mathbf{x} \quad (2.40)$$

wobei

$$\Delta f(\mathbf{x}) = Df(\bar{\mathbf{x}}) \cdot \Delta\mathbf{x} = \left. \frac{d}{d\epsilon} f(\bar{\mathbf{x}} + \epsilon \cdot \Delta\mathbf{x}) \right|_{\epsilon=0} \quad (2.41)$$

gerade der Ableitung von $f(\mathbf{x})$ in Richtung von $\Delta\mathbf{x}$ entspricht. Angewendet auf Gleichung (2.39) wird

$$L[\delta\Pi(\mathbf{u}, \delta\mathbf{u})]_{\mathbf{u}=\bar{\mathbf{u}}} = \delta\Pi(\bar{\mathbf{u}}, \delta\mathbf{u}) + \Delta\delta\Pi(\bar{\mathbf{u}}, \delta\mathbf{u}) \quad (2.42)$$

erhalten, mit der linearisierten virtuellen Arbeit

$$\Delta\delta\Pi(\bar{\mathbf{u}}, \delta\mathbf{u}) = \Delta\delta\Pi_i(\bar{\mathbf{u}}, \delta\mathbf{u}) = \int_{B_0} (\Delta\mathbf{S} : \delta\mathbf{E} + \mathbf{S} : \Delta\delta\mathbf{E}) dV \quad (2.43)$$

unter der Annahme, dass der Anteil der virtuellen äußeren Arbeiten $\delta\Pi_{\text{ext}}$ unabhängig von dem Verschiebungsfeld ist. Die Linearisierung der Materialgleichung liefert

$$\Delta\mathbf{S} = \frac{\partial\mathbf{S}}{\partial\mathbf{E}} \Delta\mathbf{E} = \mathbb{C}_T \Delta\mathbf{E} \quad (2.44)$$

mit dem von der konstitutiven Beziehung abhängigen Materialtensor 4. Stufe³ $\mathbb{C}_T$. Auf Basis der Linearisierung erfolgt die Diskretisierung im Rahmen der Finite-Elemente Methode, womit final verschiedene Lösungsstrategien zur Berechnung der Gleichgewichtspfade des Systems diskutiert werden können.

³Verkettungen in der partiellen Ableitung, z.B. für $\mathbf{S} = \mathbf{S}(\mathbf{E}, \boldsymbol{\xi}(\mathbf{E}))$, sind dabei zu berücksichtigen.

2.3 Finite-Elemente-Methode

Zur Lösung von Differentialgleichungen mittels der Numerik hat sich in weiten Teilen der Ingenieurdisziplinen, insbesondere im Bauingenieurwesen und dem Maschinenbau, die Methode der Finiten-Elemente (FEM) durchgesetzt. Auf einen umfassenden historischen Überblick wird an dieser Stelle verzichtet, nicht unerwähnt bleiben sollten allerdings einige Meilensteine, wie z.B. die erste Anwendung des Verfahrens zur Strukturberechnung von Tragflügeln durch TURNER ET. AL. [135] und die wohl erste offizielle Benennung der Methode durch den Bauingenieur RAY W. CLOUGH. Ein Rückblick auf die Anfangszeit der FEM ist in CLOUGH [22] zu finden. Die folgenden Jahrzehnte wurden geprägt durch die Weiterentwicklung der Methode, wobei hier stets Ingenieure (vgl. ZIENKIEWICZ ET AL. [155]), als auch Mathematiker (vgl. STRANG [125]) gleichermaßen beteiligt waren. Nennenswert in der 90er Jahren ist die Einführung der sogenannten h - und p -Verfeinerung durch BABUŠKA und SURI [10] sowie das Konzept der *Iso-geometrischen Analyse* (IGA) durch HUGHES ET. AL [61] in den 2000er-Jahren. Dem Zeitgeist folgend wird heutzutage vermehrt an dem Einsatz von neuronalen Netzen zur Lösung partieller Differentialgleichungen als Alternative zur FEM geforscht, s. z.B. RAISSI ET. AL [108]. Trotz der Fortschritte hierbei sind die grundlegenden Konzepte, die nachfolgend beschrieben werden, bis heute fester Bestandteil kommerzieller Software zur Berechnung von Strukturen.

In den folgenden Kapiteln wird zusammenfassend auf die im Rahmen der Arbeit benötigten Aspekte eingegangen. Ausgehend von der Herleitung der Formfunktionen wird die linearisierte schwachen Form räumlich diskretisiert. Letztlich wird die Formulierung für Schalenelemente diskutiert. Die allgemeinen Herleitungen orientieren sich an ZIENKIEWICZ ET AL. [155] und WRIGGERS [152]. Auf gesonderte Literatur wird an den entsprechenden Stellen verwiesen.

2.3.1 Isoparametrisches Konzept

Der Grundgedanke der FEM besteht in der Diskretisierung eines materiellen Körpers $\mathcal{B}_0$ mit einer endlichen Anzahl von Elementen e und zugeordneten Gebieten $\mathcal{B}^e$, deren Vereinigung *numel* wiederum eine Approximation des Körpers $\mathcal{B}_0^h$ ergibt

$$\mathcal{B}_0 \approx \mathcal{B}_0^h = \bigcup_{e=1}^{numel} \mathcal{B}^e \quad (2.45)$$

Das Vorgehen ist exemplarisch in Abbildung 2.4 für zwei verschiedene Elementtypen dargestellt. In Abhängigkeit der gewählten Elementformulierung sowie der Elementanzahl lässt sich das Gebiet unterschiedlich gut mit den dargestellten Elementen approximieren. Sowohl die Geometrie als auch die gesuchten Feldgrößen

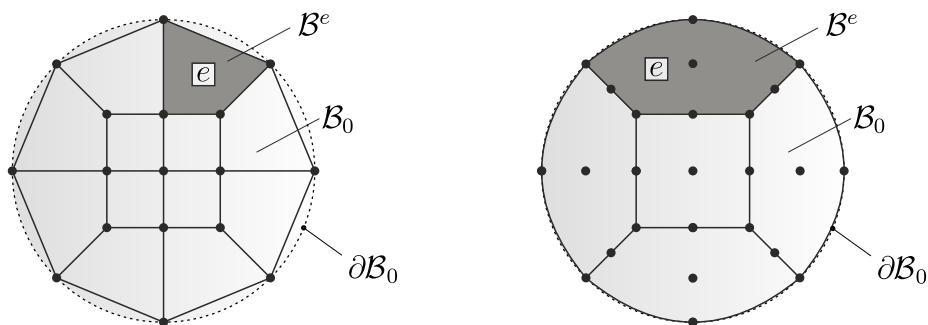


Abbildung 2.4: Mögliche Diskretisierungen eines kreisförmigen zweidimensionalen Gebiets mit finiten Elementen; lineare 4-Knoten Elemente (links) und quadratische 9-Knoten Elemente (rechts).

werden dabei mit Ansatzfunktionen innerhalb des Gebiets auf Basis diskreter Knotenwerte interpoliert. Für die Beschreibung der Geometrie eines Elements $\mathbf{X}_e$ gilt mit dem Vektor der natürlichen Koordinaten $\boldsymbol{\xi} = [\xi, \eta, \zeta]^T \in [-1, 1]^3$ die Approximation

$$\mathbf{X} \approx \mathbf{X}_e = \sum_{I=1}^{nel} N_I(\boldsymbol{\xi}) \mathbf{X}_I \quad \forall \quad \mathbf{X} \in \mathcal{B}^e. \quad (2.46)$$

$N_I(\boldsymbol{\xi})$ sind hierbei die gewählten Ansatzfunktionen und $\mathbf{X}_I = [X_1, X_2, X_3]^T_I$ ist der Ortsvektor des Knotens I . Die Ansatzfunktionen werden auf einem ein-, zwei- oder dreidimensionalen Referenzelement Ω_e mit natürlichen Koordinaten definiert und auf das physikalische Gebiet transformiert, s. Abbildung 2.5. Die Konstruktion der n Ansatzfunktionen mit Polynomgrad $n - 1$ erfolgt mit dem Prinzip der LAGRANGE-Interpolation

$$l_I(\boldsymbol{\xi}) = \prod_{\substack{J=1 \\ J \neq I}}^n \frac{\xi_J - \xi}{\xi_J - \xi_I}, \quad (2.47)$$

s. z.B. ZIENKIEWICZ ET AL. [155]. So lauten exemplarisch die Ansatzfunktionen für ein Element mit zwei Knoten

$$N_1 = l_1(\xi) = \frac{1 - \xi}{1 - (-1)} = \frac{1}{2}(1 - \xi), \quad N_2 = l_2(\xi) = \frac{-1 - \xi}{-1 - 1} = \frac{1}{2}(1 + \xi). \quad (2.48)$$

Mittels des Produktansatzes

$$N_I(\boldsymbol{\xi}) \equiv N_{IJK} = l_I(\xi) l_J(\eta) l_K(\zeta) \quad (2.49)$$

können entsprechend die dreidimensionalen Ansatzfunktionen bestimmt werden. Analog zu den geometrischen Eigenschaften werden unter Verwendung des isoparametrischen Konzepts die Feldgrößen mit den gleichen Ansätzen approximiert:

$$\mathbf{u} \approx \mathbf{u}_e = \sum_{I=1}^{nel} N_I(\boldsymbol{\xi}) \mathbf{u}_I \quad \forall \quad \mathbf{u} \in \mathcal{B}^e. \quad (2.50)$$

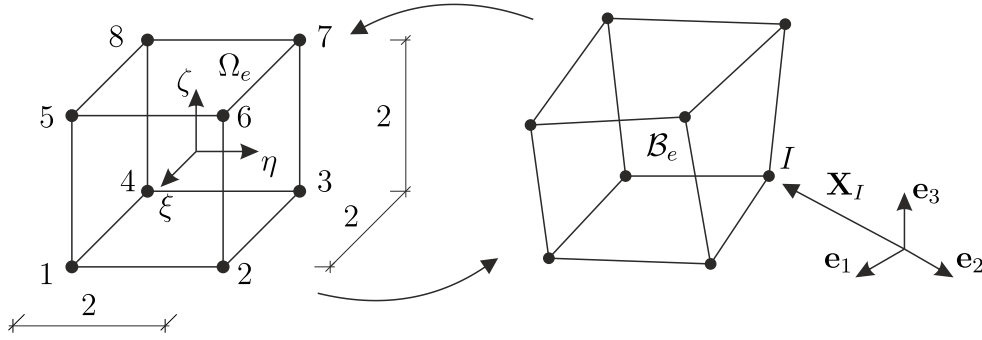


Abbildung 2.5: Isoparametrische Abbildung des physikalischen Gebiets auf das Referenzelement

Dabei ist $\mathbf{u}_I = [u_1, u_2, u_3]^T$ der Vektor der diskreten Knotenverschiebungen. Im Rahmen der FEM werden Ableitungen der Feldgrößen nach den physikalischen Koordinaten benötigt. Da die Verschiebungen selbst als diskrete Größe vorliegen, werden nur die Ableitungen der Ansatzfunktion benötigt. Unter Anwendung der Kettenregel folgt

$$N_{I,\xi} = \begin{bmatrix} N_{I,\xi} \\ N_{I,\eta} \\ N_{I,\zeta} \end{bmatrix} = \begin{bmatrix} X_{1,\xi} & X_{2,\xi} & X_{3,\xi} \\ X_{1,\eta} & X_{2,\eta} & X_{3,\eta} \\ X_{1,\zeta} & X_{2,\zeta} & X_{3,\zeta} \end{bmatrix} \begin{bmatrix} N_{I,X_1} \\ N_{I,X_2} \\ N_{I,X_3} \end{bmatrix} = \mathbf{J}_e^T N_{I,\mathbf{x}}. \quad (2.51)$$

Hierbei wurde die elementweise zu berechnenden Jacobi-Matrix⁴ $\mathbf{J}_e = \mathbf{X}_{e,\xi}$ eingeführt. Das Invertieren der Beziehung (2.51)

$$N_{I,\mathbf{x}} = \mathbf{J}_e^{-T} N_{I,\xi}. \quad (2.52)$$

liefert die gesuchten Ableitungen nach den physikalischen Koordinaten. Letztlich werden im Rahmen der FEM die Integrale der Form

$$\int_{B_e} (...) dV_e \quad (2.53)$$

numerisch gelöst. Diese Integration wird bevorzugterweise auf dem Referenzelement durchgeführt, um sich direkt den standardisierten Vorschriften numerischer Integrationsverfahren bedienen zu können. Die Transformation erfolgt mit der Determinanten der Jacobi-Matrix⁵ $\det \mathbf{J}_e$

$$\int_{B_e} (...) dV_e = \int_{-1}^1 \int_{-1}^1 \int_{-1}^1 (...) \det \mathbf{J}_e d\Omega_e. \quad (2.54)$$

⁴Die Definition der Jacobi-Matrix unterscheidet sich je nach Literatur. Häufig wird $\mathbf{J}_e = \mathbf{X}_{e,\xi}^T$ direkt als Transponierte der Ableitung der physikalischen nach den natürlichen Koordinaten definiert, s. z.B. ZIENKIEWICZ ET AL. [155]. Rein mathematisch ist die allgemeine Definition $J_{ik} = \partial f_i / \partial x_k$ üblich.

⁵Die vorangegangene Anmerkung zur Definition von $\mathbf{J}_e$ hat keinen Einfluss auf die Berechnung der Determinante da $\det \mathbf{J}_e = \det \mathbf{J}_e^T$.

Üblicherweise wird im Rahmen der Numerik die GAUSS-Integration verwendet. Details hierzu sind in WRIGGERS [152] zu finden.

2.3.2 Diskretisierung der Grundgleichungen

Mit den Ansätzen für die Geometrie und die Feldgrößen lässt sich die linearisierte schwache Form aus Gleichung (2.42) vom kontinuierlichen zum diskreten Modell überführen. Da die Formulierung praktischerweise in rein symmetrischen Tensoren erfolgt ist, kann hier der Übergang von der Tensor- zur Matrizennotation erfolgen. Tensoren 2. Stufe reduzieren sich dabei zu Vektoren und Tensoren 4. Stufe zu Matrizen. Die Vektornotationen für den GREEN-LAGRANGEN Verzerungstensor lautet hier

$$\mathbf{E}^v = [E_{11}, E_{22}, E_{33}, 2E_{12}, 2E_{13}, 2E_{23}]^T \quad (2.55)$$

und dazu passend für den zweiten PIOLA-KIRCHOFF'schen Spannungstensor

$$\mathbf{S}^v = [S_{11}, S_{22}, S_{33}, S_{12}, S_{13}, S_{23}]^T. \quad (2.56)$$

Die virtuellen sowie linearisierten Größen $\delta \mathbf{E}^v$ und $\Delta \mathbf{E}^v$ lassen sich sinngemäß darstellen. Mit den Ansätzen für die virtuellen Größen

$$\delta \mathbf{u}_e = \sum_{I=1}^{nel} N_I(\boldsymbol{\xi}) \delta \mathbf{u}_I, \quad \delta \mathbf{E}_e^v = \sum_{I=1}^{nel} \mathbf{B}_I(\mathbf{u}(\boldsymbol{\xi})) \delta \mathbf{u}_I \quad (2.57)$$

folgt die diskrete Form der virtuellen inneren Arbeit

$$\delta \Pi_i = \bigcup_{e=1}^{numel} \sum_{I=1}^{nel} \delta \mathbf{u}_I^T \int_{\tilde{\mathcal{B}}_e} \mathbf{B}_I^T \mathbf{S}_e^v dV_e = \bigcup_{e=1}^{numel} \delta \mathbf{u}_e^T \mathbf{F}_e(\mathbf{u}) \quad (2.58)$$

mit dem auf Elementebene definierten Vektor der inneren Kräfte $\mathbf{F}_e$. $\mathbf{B}_I^T$ wird häufig als Knotenoperatormatrix bezeichnet und kann z.B. WRIGGERS [152] entnommen werden. Die Integration wird, wie bei allen folgenden Integralen und wie in Gleichung (2.54) beschrieben, numerisch auf dem Referenzelement durch Auswertung an Integrationspunkten durchgeführt. Für den Anteil der virtuellen äußeren Arbeiten unter Vernachlässigung der Oberflächenlasten erhält man

$$\delta \Pi_{\text{ext}} = \bigcup_{e=1}^{numel} \sum_{I=1}^{nel} \delta \mathbf{u}_I^T \int_{\tilde{\mathcal{B}}_e} N_I^T \mathbf{B} dV_e = \bigcup_{e=1}^{numel} \delta \mathbf{u}_e^T \mathbf{P}_e^B, \quad (2.59)$$

mit dem aus den Körperlasten resultierenden Anteil des Elementlastvektors $\mathbf{P}_e^B$. Für die Berechnung des Elementlastvektors aus den Oberflächenlasten $\mathbf{P}_e^T$, s. WRIGGERS [152], sind ein- oder zweidimensionale Ansätze erforderlich. In Summe

ergibt sich der Elementlastvektor zu $\mathbf{P}_e = \mathbf{P}_e^B + \mathbf{P}_e^T$. Der Zusammenbau der Anteile (2.58) und (2.59) liefert die diskrete Form der virtuellen Arbeit

$$\delta\Pi(\mathbf{u}, \delta\mathbf{u}) = \bigcup_{e=1}^{numel} \delta\mathbf{u}_e^T [\mathbf{F}_e - \mathbf{P}_e] = \bigcup_{e=1}^{numel} \delta\mathbf{u}_e^T \mathbf{R}_e(\mathbf{u}). \quad (2.60)$$

$\mathbf{R}_e(\mathbf{u})$ ist hierbei der Residualvektor auf Elementebene. Mit dem Ansatz für die linearisierten Verzerrungen

$$\Delta\mathbf{E}_e^v = \sum_{I=1}^{nel} \mathbf{B}_I(\mathbf{u}(\boldsymbol{\xi})) \Delta\mathbf{u}_I \quad (2.61)$$

kann in analoger Vorgehensweise der tangentielle Anteil diskretisiert werden. Eine gesonderte Diskussion muss für den Anteil der linearisierten virtuellen Verzerrungen geführt werden, wobei auf eine ausführliche Herleitung an dieser Stelle verzichtet wird. Das Einsetzen der Ansätze führt jedoch auf die allgemeine Darstellung des Arbeitsanteils auf Elementebene

$$\Delta\delta\mathbf{E}_{e,IK}^v \mathbf{S}_e^v = \delta\mathbf{u}_I^T \mathbf{G}_{IK} \Delta\mathbf{u}_K \quad (2.62)$$

mit

$$\mathbf{G}_{IK} = N_{I,\mathbf{x}}^T \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} N_{K,\mathbf{x}} \mathbf{I}, \quad \mathbf{I} \in \mathbb{R}^{3 \times 3}. \quad (2.63)$$

Damit folgt

$$\Delta\delta\Pi_i = \bigcup_{e=1}^{numel} \sum_{I=1}^{nel} \sum_{K=1}^{nel} \delta\mathbf{u}_I^T \int_{\mathcal{B}_e} [\mathbf{B}_I^T \mathbf{C}_T \mathbf{B}_K + \mathbf{G}_{IK}] \Delta\mathbf{u}_K \, dV_e \quad (2.64)$$

$$= \bigcup_{e=1}^{numel} \delta\mathbf{u}_e^T \mathbf{K}_{Te}(\mathbf{u}) \Delta\mathbf{u}_e. \quad (2.65)$$

$\mathbf{K}_{Te}$ bezeichnet dabei die tangentielle Elementsteifigkeitsmatrix und $\mathbf{C}_T$ die Materialtangente. Aus den Gleichungen (2.60) und (2.65) folgt die diskrete linearisierte schwache Form

$$\bigcup_{e=1}^{numel} \delta\mathbf{u}_e^T [\mathbf{R}_e(\mathbf{u}) + \mathbf{K}_{Te}(\mathbf{u}) \Delta\mathbf{u}_e] = 0 \quad (2.66)$$

bzw. nach dem Zusammenbau aller Elemente und Berücksichtigung der Randbedingungen auf globaler Ebene

$$\delta\mathbf{u}^T [\mathbf{R}(\mathbf{u}) + \mathbf{K}_T(\mathbf{u}) \Delta\mathbf{u}] = 0. \quad (2.67)$$

Da die Testfunktion $\delta\mathbf{u}$ beliebig, jedoch für nichttriviale Lösungen $\neq 0$ sein muss, kann der Klammerterm nach den gesuchten Verschiebungssinkrementen aufgelöst

werden. Es gilt dabei $\mathbf{u} = \bar{\mathbf{u}}$, d.h., das Residuum und die tangential Steifigkeitsmatrix wird an einer festen Entwicklungsstelle ausgewertet. Im Rahmen des NEWTON-RAPHSON-Verfahrens erfolgt die Nullstellensuche durch eine iterative Berechnung der Verschiebungsinkremente

$$\Delta \mathbf{u}_{n+1} = -\mathbf{K}_T(\mathbf{u}_n)^{-1} \mathbf{R}(\mathbf{u}_n). \quad (2.68)$$

Dabei beziehen sich die Indizes n und $n+1$ auf den vorhergehenden und aktuellen Iterationsschritt. Häufig wird zusätzlich eine Inkrementierung des Lastniveaus der Form $\mathbf{P}^i = \lambda_i \mathbf{P}^0$ vorgenommen, bei der ausgehend von einem Anfangszustand $\mathbf{P}^0$ mittels des Laststeigerungsfaktors λ_i in Schrittweiten $i = 1, \dots, n$ der gewünschte Lastzustand $\mathbf{P}^n$ erreicht wird. Weitere Lösungsstrategien sind u.a. das modifizierte NEWTON-Verfahren oder das QUASI-NEWTON-Verfahren, bei welchen nicht mit der konsistenten Tangente $\mathbf{K}_T(\mathbf{u}_n)$ gerechnet wird. Details hierzu sind ebenfalls in WRIGGERS [152] zu finden. Letztlich müssen bei Stabilitätsproblemen gesonderte Methoden angewendet werden, z.B. Verschiebungssteuerungen oder das Bogenlängenverfahren. Ein Überblick zur Behandlung von Stabilitätsproblemen im Rahmen der FEM ist in WAGNER und WRIGGERS [142] gegeben.

Die in diesem Kapitel hergeleiteten Finite-Elemente Beziehungen sind unabhängig von der Wahl der Ansatzfunktionen. Im Gegensatz zu der Darstellung in Abbildung 2.5, werden im Rahmen dieser Arbeit in der Regel keine 8-Knoten Elemente verwendet, da diese besonders bei biegedominanten Problemen zu steife Lösungen liefern was auch als *locking* bezeichnet wird. Abhilfe schaffen dabei 27-Knoten Elemente mit quadratischen Ansätzen. Der Vollständigkeit halber sind noch weitere Alternativen zu erwähnen, wie beispielsweise die Verwendung von Solid-Schalenelementen (vgl. KLINKEL ET AL. [70] oder EAS-Elementen (vgl. SIMO ET AL [123]).

2.4 FEM für Schalenstrukturen

Bei schlanken, biegedominanten Problem hat sich gegenüber der Verwendung von Kontinuumsselementen (auch Volumenelemente oder „Bricks“) der Einsatz von Strukturelementen bewährt. Dabei werden die ursprünglich dreidimensionalen kontinuumsmechanischen Zusammenhänge auf eindimensionale (Balken) oder zweidimensionale (Schalen) Modelle reduziert. Bei schlanken Strukturen verringert sich durch diese Reduktion im Allgemeinen zusätzlich der Rechenaufwand, da weniger Freiheitsgrade zum Erreichen der gewünschten Approximationsgüten benötigt werden. Im Folgenden werden die wesentlichen Aspekte der FEM für das in dieser Arbeit verwendete Schalenelement hergeleitet. Details zu der Schalenformulierung sind in WAGNER und GRUTTMANN [140], GRUTTMANN und WAGNER [44] und WAGNER und GRUTTMANN [141] zu finden. Alle Herleitungen sind

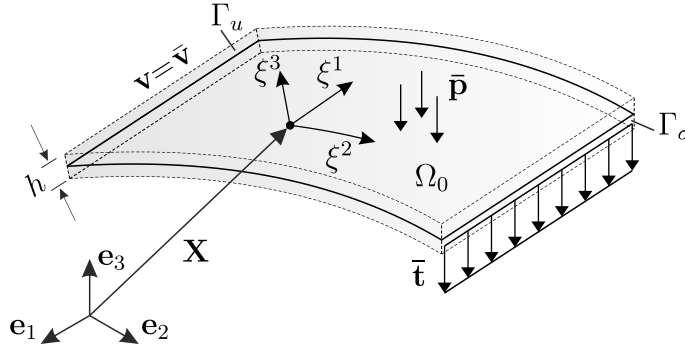


Abbildung 2.6: Darstellung des Schalenkontinuums in der Ausgangskonfiguration

den zuvor genannten Veröffentlichungen entnommen. Die aktuelle Version des Elements ist auf Basis einer Unterintegration implementiert (vgl. GRUTTMANN und WAGNER [47]) worauf hier nicht im Detail eingegangen wird.

2.4.1 Kinematik und Variationsformulierung

Zur Herleitung des Randwertproblems wird das dreidimensionale Kontinuum $\mathcal{B}_0$ auf ein zweidimensionales Kontinuum mit Referenzfläche Ω_0 und Schalendicke h in der Ausgangskonfiguration reduziert, vgl. Abbildung 2.6. Für den Rand gilt $\Gamma = \Gamma_\sigma \cup \Gamma_u$ und $\Gamma_\sigma \cap \Gamma_u = \emptyset$. Die Schale wird durch Flächenlasten $\bar{\mathbf{p}}$ und Streckenlasten $\bar{\mathbf{t}}$ auf Γ_σ belastet und Verschiebungsrandbedingungen werden auf Γ_u vorgegeben. Zur Beschreibung der Schalenebene werden lokale konvекtive Koordinaten ξ^i mit $i = 1, 2$ eingeführt, während $i = 3$ die Dickenrichtung mit $-h/2 \leq \xi^3 \leq h/2$ beschreibt. Im Folgenden gelten die Konvention $i, j = 1, 2, 3$ sowie $\alpha, \beta = 1, 2$ um die kartesischen Koordinaten von Flächenkoordinaten zu unterscheiden. Für den Ortsvektor Φ eines Punktes $P \in \mathcal{B}_0$ folgt

$$\Phi(\xi^1, \xi^2, \xi^3) = \mathbf{X} + \xi^3 \mathbf{D}(\xi^1, \xi^2), \quad |\mathbf{D}(\xi^1, \xi^2)| = 1. \quad (2.69)$$

$\mathbf{X}$ ist dabei der Ortsvektor eines Punktes auf der Schalenmittelfläche Ω_0 und $\mathbf{D}(\xi^1, \xi^2)$ der senkrecht auf dieser stehende Direktor. Entsprechend beschreibt φ den Ortsvektor eines Punktes der Momentankonfiguration

$$\varphi(\xi^1, \xi^2, \xi^3, t) = \mathbf{x} + \xi^3 \mathbf{d}(\xi^1, \xi^2), \quad |\mathbf{d}(\xi^1, \xi^2)| = 1 \quad (2.70)$$

mit dem Vektor zur Schalenmittelfläche $\mathbf{x}$ der Momentankonfiguration Ω_t und dem verformten Direktor $\mathbf{d}(\xi^1, \xi^2)$. Durch die Einschränkung in Gleichung (2.70) werden Dickenverzerrungen a priori ausgeschlossen. Allerdings steht der verformte Direktor nicht mehr senkrecht auf der Schalenmittelfläche, womit Schubzerrungen berücksichtigt werden. Mit der Definition (2.10) folgen die zu den konvекtiven

Koordinaten kovarianten Komponenten des GREEN-LAGRANGE Verzerrungstensors

$$E_{ij} = \frac{1}{2}(\boldsymbol{\varphi}_{,i} \cdot \boldsymbol{\varphi}_{,j} - \boldsymbol{\Phi}_{,i} \cdot \boldsymbol{\Phi}_{,j}). \quad (2.71)$$

Das Einsetzen der Gleichungen (2.69) und (2.70) in (2.71) liefert die Schalenverzerrungen

$$\begin{aligned} \varepsilon_{\alpha\beta} &= \frac{1}{2}(\mathbf{x}_{,\alpha} \cdot \mathbf{x}_{,\beta} - \mathbf{X}_{,\alpha} \cdot \mathbf{X}_{,\beta}) \\ \kappa_{\alpha\beta} &= \frac{1}{2}(\mathbf{x}_{,\alpha} \cdot \mathbf{d}_{,\beta} + \mathbf{x}_{,\beta} \cdot \mathbf{d}_{,\alpha} - \mathbf{X}_{,\alpha} \cdot \mathbf{D}_{,\beta} + \mathbf{X}_{,\beta} \cdot \mathbf{D}_{,\alpha}) \\ \gamma_{\alpha} &= \mathbf{x}_{,\alpha} \cdot \mathbf{d} - \mathbf{X}_{,\alpha} \cdot \mathbf{D}, \end{aligned} \quad (2.72)$$

mit den Membranverzerrungen $\varepsilon_{\alpha\beta}$, den Krümmungen $\kappa_{\alpha\beta}$ und den Schubverzerrungen γ_{α} . Diese lassen sich vektoriell zusammenfassen

$$\boldsymbol{\varepsilon}_G(\mathbf{v}) = [\varepsilon_{11}, \varepsilon_{22}, 2\varepsilon_{12}, \kappa_{11}, \kappa_{22}, 2\kappa_{12}, \gamma_1, \gamma_2]^T. \quad (2.73)$$

Der Vektor der Schalenfreiheitsgrade $\mathbf{v} = [\mathbf{u}, \boldsymbol{\omega}]^T$ unterteilt sich in Anteile aus der Verschiebung der Schalenmittelfläche $\mathbf{u} = \mathbf{x} - \mathbf{X}$ und der Freiheitsgrade $\boldsymbol{\omega}$ zur Beschreibung der Rotation des Direktors. Für Details wird auf WAGNER und GRUTTMANN [140] verwiesen. Es sei jedoch an dieser Stelle vorweggreifend erwähnt, dass bei koplanaren Knoten von angrenzenden Schalenelementen der Vektor $\mathbf{v}$ drei Verschiebungs- und zwei lokale Verdrehfreiheitsgrade beinhaltet. Im Allgemeinen müssen jedoch alle sechs Freiheitsgrade vorgehalten werden. Weiterhin ermöglicht die hier verwendete Parametrisierung des verformten Direktors die Beschreibung finiter Rotationen. Der den Verzerrungen zugeordnete zweite PIOLA-KIRCHOFFSCHE Vektor der Schalenschnittgrößen lautet

$$\boldsymbol{\sigma} = [n^{11}, n^{22}, n^{12}, m^{11}, m^{22}, m^{12}, q^1, q^2]^T. \quad (2.74)$$

Formal ist dieser bezüglich einer kontravarianten Basis definiert. Bei der verwendeten Formulierung wird jedoch ein lokales kartesisches Koordinatensystem eingeführt, wodurch die ko- und kontravarianten Basen identisch sind (vgl. WRIGGERS [152]). Dadurch müssen die ko- und kontravarianten Ableitungen der Tangentenvektoren im Rahmen der Numerik nicht berechnet werden. Im Weiteren wird dahingehend auf den hochgestellten Index bei den Schalenschnittgrößen verzichtet.

In der folgenden Herleitung wird eine Formulierung auf Basis des HU-WASHIZU Variationsprinzips verwendet (vgl. [143]). Dabei treten die Schalenfreiheitsgrade $\mathbf{v}$, die Verzerrungen $\boldsymbol{\varepsilon}$ und die Streckenschnittgrößen $\boldsymbol{\sigma}$ jeweils als unabhängige Variablen auf. Diese werden im Vektor

$$\boldsymbol{\theta} = [\mathbf{v}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}]^T \quad \text{und} \quad \delta\boldsymbol{\theta} = [\delta\mathbf{v}, \delta\boldsymbol{\varepsilon}, \delta\boldsymbol{\sigma}]^T \quad (2.75)$$

zusammengefasst. Damit lässt sich die schwache Form schreiben als

$$\begin{aligned} \delta\pi(\boldsymbol{\theta}, \delta\boldsymbol{\theta}) = & \int_{\Omega_0} \left[\delta\boldsymbol{\varepsilon}^T(\boldsymbol{\sigma}_S(\boldsymbol{\varepsilon}) - \boldsymbol{\sigma}) + \delta\boldsymbol{\sigma}^T(\boldsymbol{\varepsilon}_G(\mathbf{v}) - \boldsymbol{\varepsilon}) + \delta\boldsymbol{\varepsilon}_G^T \boldsymbol{\sigma} \right] dA \\ & - \int_{\Omega_0} \delta\mathbf{u}^T \bar{\mathbf{p}} dA - \int_{\Gamma_\sigma} \delta\mathbf{u}^T \bar{\mathbf{t}} ds = \mathbf{0}. \end{aligned} \quad (2.76)$$

Gemischte Variationsprinzipien reduzieren einerseits numerische Versteifungseffekte (*locking*) und ermöglichen andererseits eine bessere Approximation der Spannungsgrößen. Wie zuvor beschrieben, lässt sich Gleichung (2.76) unter Anwendung der verallgemeinerten Richtungsableitung (2.35) prinzipiell aus einem Potential $\pi(\boldsymbol{\theta})$ ableiten. Dies setzt allerdings die Existenz einer Energiefunktion⁶ Ψ voraus, aus der sich die konstitutive Beziehung ableiten lässt. Dies ist teilweise auch für inelastische Materialien möglich, vgl. SIMO ET. AL [124] für elastoplastisches und HOLZAPFEL ET. AL [56] für viskoelastisches Materialverhalten. Gleichung (2.76) ist allgemein unabhängig von der Existenz eines Potentials und beinhaltet direkt die EULER-LAGRANGE Gleichungen in den Integranden, die für beliebige Testfunktion $\delta\mathbf{v}$, $\delta\boldsymbol{\varepsilon}$ und $\delta\boldsymbol{\sigma}$ erfüllt werden müssen.

$$\delta\pi(\boldsymbol{\theta}, \delta\boldsymbol{\theta})_{\delta\boldsymbol{\varepsilon}} = \int_{\Omega_0} \delta\boldsymbol{\varepsilon}^T(\boldsymbol{\sigma}_S(\boldsymbol{\varepsilon}) - \boldsymbol{\sigma}) dA \quad (2.77)$$

$$\delta\pi(\boldsymbol{\theta}, \delta\boldsymbol{\theta})_{\delta\boldsymbol{\sigma}} = \int_{\Omega_0} \delta\boldsymbol{\sigma}^T(\boldsymbol{\varepsilon}_G(\mathbf{v}) - \boldsymbol{\varepsilon}) dA \quad (2.78)$$

$$\delta\pi(\boldsymbol{\theta}, \delta\boldsymbol{\theta})_{\delta\mathbf{u}} = \int_{\Omega_0} \delta\boldsymbol{\varepsilon}_G^T \boldsymbol{\sigma} dA - \int_{\Omega_0} \delta\mathbf{u}^T \bar{\mathbf{p}} dA - \int_{\Gamma_\sigma} \delta\mathbf{u}^T \bar{\mathbf{t}} ds. \quad (2.79)$$

Die Gleichung (2.78) fordert schwach die Kinematik, Gleichung (2.79) das Gleichgewicht und Gleichung (2.77) das Stoffgesetz ein, wobei für letzteres lediglich ein konstitutiver Zusammenhang zwischen den Streckenschnittgrößen $\boldsymbol{\sigma}_S$ und den Verzerrungen $\boldsymbol{\varepsilon}$ notwendig ist. Die Linearisierung

$$L[\delta\pi(\boldsymbol{\theta}, \delta\boldsymbol{\theta})]_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}} = \delta\pi(\bar{\boldsymbol{\theta}}, \delta\boldsymbol{\theta}) + \Delta\delta\pi(\bar{\boldsymbol{\theta}}, \delta\boldsymbol{\theta}) \quad (2.80)$$

führt auf ein Gleichungssystem in den inkrementellen Größen $\Delta\mathbf{v}$, $\Delta\boldsymbol{\varepsilon}$ und $\Delta\boldsymbol{\sigma}$, welches die Basis für die Diskretisierung im Rahmen der FEM stellt.

2.4.2 Finite-Elemente Formulierung

Mit den zuvor erläuterten Beziehungen werden in der weiteren Darstellung die wesentlichen Aspekte der FEM diskutiert. Grundsätzlich müssen sowohl die Geometrie der Referenz- und Momentankonfiguration als auch die gesuchten unabhängigen Feldgrößen mit geeigneten Ansätzen approximiert werden. Dafür werden

⁶Die Konstruktion von Energiefunktionen wird in Kapitel 3.2 ausführlich behandelt

4-Knoten Schalenelemente mit den bilinearen Ansatzfunktionen

$$N_I(\xi, \eta) = \frac{1}{4}(1 + \xi_I \xi)(1 + \eta_I \eta), \quad [\xi, \eta] \in [-1, 1] \quad (2.81)$$

eingeführt. ξ_I und η_I entsprechen dabei den Knotenkoordinaten des zweidimensionalen Referenzelements des zugrundeliegenden isoparametrischen Konzepts (vgl. Kapitel 2.3.1, Gleichungen (2.47) und (2.49)). Auf Basis der bekannten physikalischen Knotenkoordinaten $\mathbf{X}_I$ und der Knotendirektoren $\mathbf{D}_I$ lassen sich diese innerhalb des Elements approximieren

$$\mathbf{X}^h = \sum_{I=1}^4 N_I \mathbf{X}_I, \quad \mathbf{D}^h = \sum_{I=1}^4 N_I \mathbf{D}_I. \quad (2.82)$$

Für die Konstruktion einer lokalen kartesischen Basis⁷ gibt es grundsätzlich verschiedene Möglichkeiten. Ein Ansatz zur Konstruktion wäre über die Tangentenvektoren $\mathbf{X}_{,\alpha}$ und $\mathbf{X}_{,\beta}$ möglich, der hier im Weiteren nicht diskutiert wird (s. hierzu z.B. WRIGGERS [152]). Alternativ wäre die Verwendung der sogenannten Lamina-Basis möglich, bei der die Basisvektoren $\mathbf{t}_\alpha$ so konstruiert werden, dass sie in der Tangentialebene der Mittelfläche liegen und möglichst parallel zu den isoparametrischen Koordinaten ξ und η verlaufen (vgl. HUGHES [60]). Im Folgenden wird ein Ansatz mittels der Elementdiagonalen $\mathbf{d}_1$ und $\mathbf{d}_2$ verwendet (vgl. Abbildung 2.7).

$$\begin{aligned} \mathbf{d}_1 &= \mathbf{X}_3 - \mathbf{X}_1, & \hat{\mathbf{d}}_1 &= \frac{\mathbf{d}_1}{|\mathbf{d}_1|} \\ \mathbf{d}_2 &= \mathbf{X}_2 - \mathbf{X}_4, & \hat{\mathbf{d}}_2 &= \frac{\mathbf{d}_2}{|\mathbf{d}_2|} \end{aligned} \quad (2.83)$$

Damit lassen sich die kartesischen Basisvektoren $\mathbf{t}_i$ wie folgt bestimmen

$$\mathbf{t}_1 = \frac{\hat{\mathbf{d}}_1 + \hat{\mathbf{d}}_2}{|\hat{\mathbf{d}}_1 + \hat{\mathbf{d}}_2|}, \quad \mathbf{t}_2 = \frac{\hat{\mathbf{d}}_1 - \hat{\mathbf{d}}_2}{|\hat{\mathbf{d}}_1 - \hat{\mathbf{d}}_2|}, \quad \mathbf{t}_3 = \mathbf{t}_1 \times \mathbf{t}_2. \quad (2.84)$$

Dabei sei angemerkt, dass für beliebige nicht-planere Schalenelemente der Ursprung der kartesischen Basis nicht zwingend auf einem Kreuzungspunkt der Diagonalen liegen muss. Damit kann die Jacobi-Matrix berechnet werden, welche für die Transformation der Ableitungsbeziehungen zwischen den natürlichen und physikalischen Koordinaten benötigt wird

$$\begin{bmatrix} N_{I,1} \\ N_{I,2} \end{bmatrix} = \mathbf{J}^{-1} \begin{bmatrix} N_{I,\xi} \\ N_{I,\eta} \end{bmatrix}, \quad \mathbf{J} = \begin{bmatrix} \mathbf{X}_{,\xi}^h \cdot \mathbf{t}_1 & \mathbf{X}_{,\xi}^h \cdot \mathbf{t}_2 \\ \mathbf{X}_{,\eta}^h \cdot \mathbf{t}_1 & \mathbf{X}_{,\eta}^h \cdot \mathbf{t}_2 \end{bmatrix}. \quad (2.85)$$

⁷Die globale kartesische Basis kann bei Schalenelementen nicht verwendet werden, da die Abbildung im Rahmen des isoparametrischen Konzepts nicht singularitätenfrei ist (s. WRIGGERS [152]).

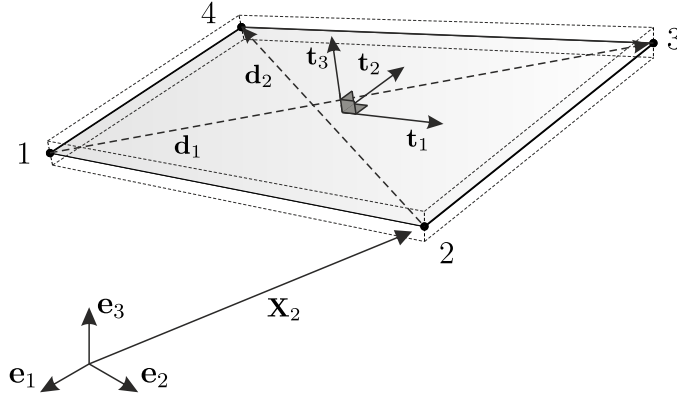


Abbildung 2.7: Schalenelement mit lokalem kartesischem Koordinatensystem

Die Berechnung der Einträge von $\mathbf{J}$ kann WAGNER und GRUTTMANN [140] entnommen werden. Damit folgt für die Ableitungen der geometrischen Größen in der Referenzkonfiguration

$$\mathbf{X}_{,\alpha}^h = \sum_{I=1}^4 N_{I,\alpha} \mathbf{X}_I, \quad \mathbf{D}_{,\alpha}^h = \sum_{I=1}^4 N_{I,\alpha} \mathbf{D}_I. \quad (2.86)$$

Analog gilt für die Approximation der Größen der Momentankonfiguration

$$\begin{aligned} \mathbf{x}^h &= \sum_{I=1}^4 N_I \mathbf{x}_I, & \mathbf{d}^h &= \sum_{I=1}^4 N_I \mathbf{d}_I, \\ \mathbf{x}_{,\alpha}^h &= \sum_{I=1}^4 N_{I,\alpha} \mathbf{x}_I, & \mathbf{d}_{,\alpha}^h &= \sum_{I=1}^4 N_{I,\alpha} \mathbf{d}_I, \end{aligned} \quad (2.87)$$

wobei sich die Knotengrößen der Momentankonfiguration $\mathbf{x}_I$ und $\mathbf{d}_I$ aus den Verschiebungsbeziehungen $\mathbf{x}_I = \mathbf{X}_I + \mathbf{u}_I$ und $\mathbf{d}_I = \mathbf{R}_I \mathbf{D}_I$ ergeben. $\mathbf{R}_I$ ist hierbei eine Rotationsmatrix die über die Freiheitsgrade $\boldsymbol{\omega}_I$ beschrieben wird. Damit stehen alle notwendigen Größen zur Verfügung, die unter Einsetzen in Gleichung (2.72) zur Approximation der geometrischen Verzerrungen $\boldsymbol{\varepsilon}_G^h$ benötigt werden. Dabei erfolgt zusätzlich eine Modifikation der Schubverzerrungsanteile γ_α , um numerischen Versteifungseffekten entgegenzuwirken. Diese Modifikation basiert auf dem Ansatz nach BATHE und DVORKIN [12], wonach die Schubverzerrungen auf Basis spezieller Kollokationspunkte interpoliert werden und nicht mittels der Elementknoten. Mit Ansätzen für die virtuellen Größen

$$\delta \mathbf{x}_{,\alpha}^h = \sum_{I=1}^4 N_{I,\alpha} \delta \mathbf{u}_I, \quad \delta \mathbf{d}_{,\alpha}^h = \sum_{I=1}^4 N_{I,\alpha} \delta \mathbf{d}_I, \quad (2.88)$$

sowie den speziellen Ansätzen für die virtuellen Schubverzerrungen $\delta \gamma_\alpha^h$ nach BATHE und DVORKIN folgt die allgemeine Form der virtuellen Verzerrungen

$$\delta \boldsymbol{\varepsilon}_G^h = \sum_{I=1}^4 \mathbf{B}_I \delta \mathbf{v}_I \quad (2.89)$$

Die Einträge der Knotenoperatormatrix $\mathbf{B}_I$ können WAGNER und GRUTTMANN [140] entnommen werden, worauf ebenfalls für die Herleitung und Spezifikation der virtuellen linearisierten Verzerrungen $\Delta\delta\boldsymbol{\varepsilon}_G^h$ verwiesen wird. Es verbleiben die unabhängigen Größen $\boldsymbol{\sigma}$ und $\boldsymbol{\varepsilon}$. Für die Streckenschnittgrößen wird dabei auf Elementebene ein diskontinuierlicher Ansatz mit 14 unabhängigen Parametern gewählt

$$\boldsymbol{\sigma}^h = \mathbf{N}_\sigma \hat{\boldsymbol{\sigma}}, \quad \delta\boldsymbol{\sigma}^h = \mathbf{N}_\sigma \delta\hat{\boldsymbol{\sigma}}, \quad (2.90)$$

die in dem Vektor $\hat{\boldsymbol{\sigma}}$ enthalten sind. In der Matrix $\mathbf{N}_\sigma$ sind sowohl Ansätze zur Interpolation der konstanten als auch der fluktuierenden Anteile der Streckenschnittgrößen enthalten. Details hierzu sind z.B. in WAGNER und GRUTTMANN [141] zu finden. Analog werden für die Verzerrungen Ansätze der Form

$$\boldsymbol{\varepsilon}^h = [\mathbf{N}_\varepsilon^1, \mathbf{N}_\varepsilon^2] \begin{bmatrix} \hat{\boldsymbol{\varepsilon}}_1 \\ \hat{\boldsymbol{\varepsilon}}_2 \end{bmatrix}, \quad \delta\boldsymbol{\varepsilon}^h = [\mathbf{N}_\varepsilon^1, \mathbf{N}_\varepsilon^2] \begin{bmatrix} \delta\hat{\boldsymbol{\varepsilon}}_1 \\ \delta\hat{\boldsymbol{\varepsilon}}_2 \end{bmatrix} \quad (2.91)$$

gewählt. Die Verzerrungsanteile $\hat{\boldsymbol{\varepsilon}}_1$ und $\delta\hat{\boldsymbol{\varepsilon}}_1$ werden dabei in gleicher Weise interpoliert wie die Spannungen und enthalten damit 14 unabhängige Parameter. Für die Verzerrungsanteile $\hat{\boldsymbol{\varepsilon}}_2$ und $\delta\hat{\boldsymbol{\varepsilon}}_2$ können für die Membran- und Biegeanteile jeweils eine unterschiedliche Anzahl von Parametern $n_{mb} = n_m + n_b$ gewählt werden wobei gilt

$$n_m \in [0, 2, 4, 7, 9, 11] \quad \text{und} \quad n_b \in [0, 2, 4, 6]. \quad (2.92)$$

Die genaue Besetzung der Matrizen $\mathbf{N}_\varepsilon^1$ und $\mathbf{N}_\varepsilon^2$ ist in WAGNER und GRUTTMANN [140] zu finden. Formal können abschließend die Ansätze für die linearisierten Größen zusammengefasst werden:

$$\Delta\boldsymbol{\theta}^h = \mathbf{N}_\theta \Delta\hat{\boldsymbol{\theta}} \quad \text{mit} \quad \Delta\hat{\boldsymbol{\theta}} = [\Delta\mathbf{v}, \Delta\hat{\boldsymbol{\varepsilon}}, \Delta\hat{\boldsymbol{\sigma}}]^T. \quad (2.93)$$

Diese werden anschließend in Gleichung (2.80) eingesetzt. Durch die Wahl der diskontinuierlichen Ansätze können die linearisierten Größen $\Delta\hat{\boldsymbol{\varepsilon}}$ und $\Delta\hat{\boldsymbol{\sigma}}$ auf Elementebene auskondensiert werden, wodurch global das Gleichungssystem rein in den Verschiebungsfreiheitsgraden zu lösen ist. Für die hier dargestellte Formulierung werden auf Elementebene 2×2 Integrationspunkte verwendet.

Für das beschriebene Schalenelement sind zum Stand der Arbeit drei mögliche Varianten im Forschungscode FEAP [128] implementiert, mit denen die konstitutive Beziehung berücksichtigt werden kann. Zunächst ist es möglich, geschichtetes linear elastisches Verhalten über eine Dickenintegration

$$\boldsymbol{\sigma}_S = \mathbf{D}\boldsymbol{\varepsilon}, \quad \mathbf{D} = \int_{-h/2}^{h/2} \mathbf{A}^T \mathbf{C} \mathbf{A} \, d\xi^3, \quad \mathbf{A} = \begin{bmatrix} \mathbf{I}_3 & \xi^3 \mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_2 \end{bmatrix} \quad (2.94)$$

zu berücksichtigen. Dabei ist $\mathbf{C}$ die Elastizitätsmatrix und $\mathbf{I}_n$ bezeichnet Einheitsmatrizen vom Rang n . Alternativ ist eine Kopplung am Integrationspunkt zu einem Mikroproblem möglich, über welches die konstitutiven Größen berechnet werden können. Dieses Vorgehen wird in Kapitel 5 diskutiert. Zuletzt kann der konstitutive Zusammenhang mit einem künstlichen neuronalen Netz abgebildet werden. Dies stellt den Kern der vorliegenden Arbeit dar und wird in Kapitel 7 erläutert. Für beide zuvor genannten Varianten soll das Materialverhalten ratenabhängige Effekte abbilden können. Dazu werden im folgenden Kapitel die Grundlagen viskoelastischer Materialien erläutert.

3 Viskoelastizität

Die in Kapitel 2 hergeleiteten mechanischen und numerischen Zusammenhänge wurden materialunabhängig formuliert. Für die vollständige Lösung des Randwertproblems ist jedoch die Spezifikation einer konstitutiven Gleichung erforderlich, welche die gewählten Spannungsmaße eindeutig mit geeigneten Verzerrungsmaßen sowie gegebenenfalls weiteren internen Zustandsvariablen verknüpft. Die Materialmodellierung stellt damit den abschließenden Bestandteil der Feldgleichungen dar und bildet zugleich einen zentralen Schwerpunkt aktueller Forschung. Während Kinematik und Gleichgewichtsbedingungen eine allgemein akzeptierte physikalische Basis zur Beschreibung eines materiellen Körpers liefern, eröffnet die Festlegung des Materialmodells erheblichen Gestaltungs- und Interpretationsspielraum. Diese Wahl richtet sich in der Regel nach den für die jeweilige technische Anwendung dominierenden Werkstoffeigenschaften.

Im weiteren Verlauf der Arbeit wird der Fokus auf die Beschreibung viskoelastischer Materialien gelegt. Diese zeichnen sich dadurch aus, dass sie sowohl elastisches als auch viskoses Verhalten zeigen und somit Eigenschaften von Festkörpern und Fluiden vereinen. Beobachten lässt sich das Verhalten z.B. bei Beton (vgl. ARGYRIS ET. AL. [8]) oder bei Polymeren (vgl. FERRY [30]). Konkret hängt der Spannungszustand in diesen Materialien nicht nur von der Dehnung des aktuellen Zeitpunkts, sondern auch von der Dehnungsgeschichte ab.

Die Viskoelastizität lässt sich in die Theorie der *einfachen Stoffe* nach TRUESDELL und NOLL einordnen, die in [134] begründet und in NOLL [101] weiterentwickelt wurde. Auf eine detaillierte Darstellung der mathematisch anspruchsvollen Grundlagen wird an dieser Stelle verzichtet. In Anlehnung an die Ausführungen in KRAWIETZ [72] lässt sich die Kernaussage der Theorie jedoch wie folgt zusammenfassen: Für jeden materiellen Punkt $\mathbf{X}$ existiert ein Ausgabefunktional, das für eine gegebene Eingangsgröße und eine Prozessvariable eine eindeutige Ausgangsgröße bestimmt. In rein mechanischen Systemen lassen sich diese Größen konkret als Dehnung (Eingangsgröße), Zeit (Prozessvariable) und Spannung (Ausgangsgröße) interpretieren. Zudem lassen sich aus der Theorie die grundlegenden Prinzipien des Materialverhaltens ableiten (vgl. KRAWIETZ [72]):

- **Determinismus** Die Spannungen am Punkt lassen sich eindeutig durch die Ausgangssituation des Körpers und den Prozess der Verzerrungen am Punkt $\mathbf{X}$ bestimmen.
- **Lokale Wirkung** Das Verhalten des Restkörpers ist ohne Einfluss auf die - infinitesimale Umgebung - des Punktes $\mathbf{X}$. Das Materialverhalten kann lokal betrachtet werden.

- **Materielle Objektivität** Das Materialverhalten am Punkt $\mathbf{X}$ ist unabhängig vom Beobachter und damit invariant gegenüber Beobachterwechsel.

Diese abstrakte Struktur wird sich in den dargestellten Modellansätzen dieses Kapitels wiederfinden.

Zur Beschreibung des konstitutiven Zusammenhangs werden nachfolgend eindimensionale rheologische Bausteine eingeführt, um die wesentlichen Eigenschaften viskoelastischer Materialien mit einfachen Modellen abzubilden und zu diskutieren. Anschließend werden diese in den thermodynamischen Kontext eingeordnet, der eine verallgemeinerte Formulierung auf Basis der freien Energie ermöglicht. Des Weiteren lassen sich an dieser Stelle Bedingungen für physikalische Konsistenz von Materialgesetze diskutieren. Darauf aufbauend folgt die mehrdimensionale Erweiterung des verwendeten Modells. Letztlich werden die analytischen Beziehungen in geeignete Formulierungen für deren Verwendung in der Numerik überführt. Die analytischen Ausführungen stützen sich im Wesentlichen auf CHRISTENSEN [21], LUBLINER [87] und MALVERN [88], die thermodynamischen und numerischen Herleitungen auf SIMO [122] und HOLZAPFEL [55]. Weitere verwendete Literatur wird an den entsprechenden Stellen ergänzend erwähnt.

3.1 Eindimensionale rheologische Modelle

Alle Herleitungen basieren zunächst auf der Theorie der infinitesimalen Verformungen, weshalb zwischen verschiedenen Spannungs- und Dehnungsmaßen nicht unterschieden werden muss. Spannungen werden daher mit σ und Dehnungen mit ε bezeichnet. Ausgangspunkt zur Beschreibung der viskoelastischen Spannungs-Dehnungs-Beziehung sind dabei eine Feder und ein Dämpfer als rheologische Grundbausteine: Die Feder (F), mit Elastizitätsmodul E , repräsentiert das elastische Verhalten eines Festkörpers mit dem Teilstoffgesetz

$$\sigma_F = E \varepsilon. \quad (3.1)$$

Der Dämpfer (D) repräsentiert das viskose Verhalten eines Fluids das durch die Viskosität η und das Teilstoffgesetz

$$\sigma_D = \eta \dot{\varepsilon} \quad (3.2)$$

beschrieben wird. Im einfachsten Fall werden diese in einer Parallelschaltung angeordnet, s. Abbildung 3.1, was auf das KELVIN-VOIGT-Modell führt. Die äußere aufgebrachte Spannung σ steht durch gedankliches Freischneiden im Gleichgewicht mit der Spannung in der Feder σ_F und des Dämpfers σ_D . Weiterhin muss die Dehnung in beiden Elementen gleich groß sein, wodurch sich mit den Teilstoffgesetzen (3.1) und (3.2) durch Superposition die konstitutive Beziehung

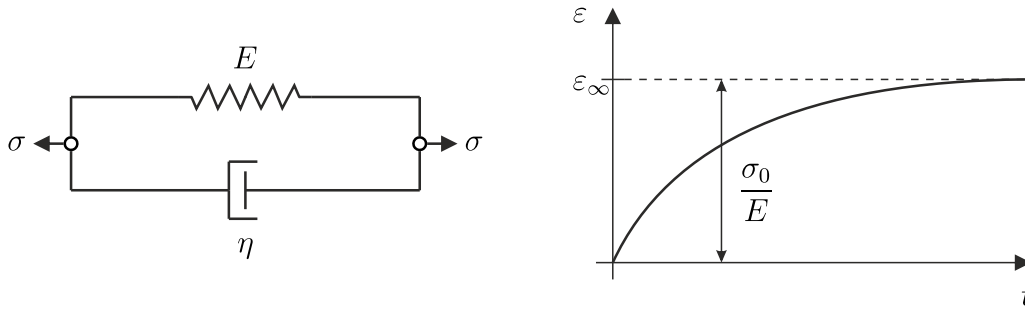


Abbildung 3.1: KELVIN-VOIGT-Modell und dazugehöriges Kriechverhalten

$$\sigma = \sigma_E + \sigma_D = E \varepsilon + \eta \dot{\varepsilon} \quad (3.3)$$

herleiten lässt. Unter Einführung der Retardationszeit

$$\hat{\tau} := \frac{\eta}{E} \quad (3.4)$$

lässt sich die Gleichung (3.3) umschreiben zu

$$\dot{\varepsilon} + \frac{\varepsilon}{\hat{\tau}} = \frac{\sigma}{\eta}, \quad (3.5)$$

was eine Differentialgleichung 1. Ordnung darstellt. Unter Vorgabe einer konstanten Spannung $\sigma(t) = \sigma_0$ lässt sich diese lösen und die Entwicklung der Verzerrung über die Zeit bestimmen

$$\varepsilon(t) = J(t) \sigma_0 \quad \text{mit} \quad J(t) = \frac{1}{E} (1 - e^{-t/\hat{\tau}}) \quad \text{und} \quad t \in [0, T]. \quad (3.6)$$

Hierbei ist $J(t)$ die Kriechfunktion. Das KELVIN-VOIGT-Modell liefert keine initiale elastische Antwort für $t \rightarrow 0$. Für $t \rightarrow \infty$ wird asymptotisch die Gleichgewichtsverzerrung $\varepsilon_\infty = \sigma_0/E$ erreicht, vgl. Abbildung 3.1. Zudem kann das Modell keine Spannungsrelaxation bei konstanter Dehnung $\varepsilon(t) = \varepsilon_0$ abbilden, da dehnungsbedingte Diskontinuitäten zu unendlich großen Spannungen führen.

Eine entsprechende Reihenschaltung der Feder und des Dämpfers wird als MAXWELL-Modell bezeichnet, was in Abbildung 3.2 dargestellt ist. In diesem Fall muss die Spannung in beiden Elementen gleich sein. Die Dehnung des Gesamtmodells setzt sich additiv aus dem Anteil der Feder ε_F und des Dämpfers ε_D zusammen. Die Addition und Ableitung nach der Zeit führen mit

$$\dot{\varepsilon}_F = \frac{\dot{\sigma}}{E}, \quad \dot{\varepsilon}_D = \frac{\sigma}{\eta} \quad (3.7)$$

auf die differentielle Spannungs-Dehnungsbeziehung des MAXWELL-Modells

$$\dot{\sigma} + \frac{1}{\tau} \sigma = \dot{\varepsilon} E. \quad (3.8)$$

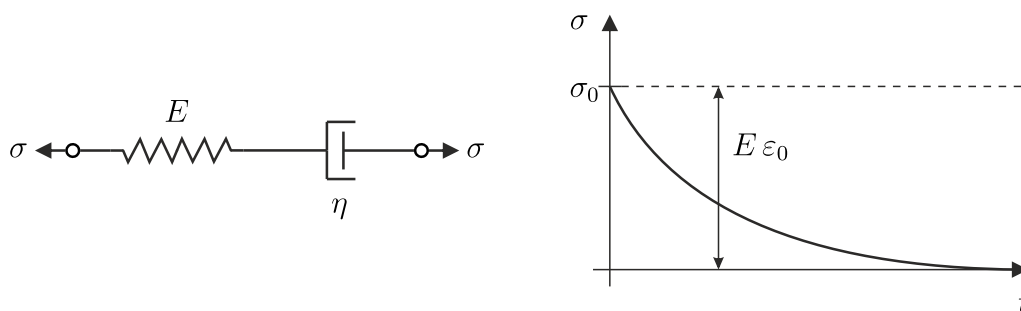


Abbildung 3.2: MAXWELL-Modell und dazugehöriges Relaxationsverhalten

Dabei wurde mit

$$\tau := \frac{\eta}{E} \quad (3.9)$$

die Relaxationszeit eingeführt⁸. Gleichung (3.8) ist für die Beschreibung festkörperartigen Verhaltens ungeeignet. Zwar lässt sich das Kriechverhalten bei konstanter Spannung σ_0 abbilden, jedoch konvergiert die Dehnung nicht gegen ein Gleichgewichtsniveau, was zu einer unbegrenzten (linearen) Zunahme führt. Bei konstanter Dehnung ε_0 hingegen zeigt das Modell ein typisches Relaxationsverhalten

$$\sigma(t) = E(t) \varepsilon_0 \quad \text{mit} \quad E(t) = E e^{-t/\tau} \quad \text{und} \quad t \in [0, T]. \quad (3.10)$$

$E(t)$ wird als Relaxationsfunktion bezeichnet. Das Modell liefert für einen instantanen Dehnungssprung ε_0 zum Zeitpunkt $t = 0$ eine Spannungsantwort $\sigma_0 = E \varepsilon_0$ und die Grenzwertbetrachtung liefert $\lim_{t \rightarrow \infty} \sigma(t) = 0$, s. Abbildung 3.2. Der Wert 0 wird zwar nur asymptotisch erreicht, allerdings reduziert sich die Spannung für $t \approx 5\tau$ auf $< 0.01 \sigma_0$.

Beide zuvor eingeführten rheologischen Anordnungen liefern somit keine zufriedenstellende Beschreibung viskoelastischer Effekte von Festkörpern. Entsprechend naheliegend ist die Kombination beider Modelle, was im Weiteren als linearer Standardkörper⁹ bezeichnet wird und in Abbildung 3.3 dargestellt ist. Gedankliches Freischneiden am rechten Teil des Modells durch die Feder E_∞ und den Dämpfer η führt mit der bekannten Dehnung in der Feder ε auf die Gesamtspannung

$$\sigma = E_\infty \varepsilon + \sigma_D. \quad (3.11)$$

Die Dehnung in der Feder E_1 beträgt $(\varepsilon - \varepsilon_D)$. Da aus Gleichgewichtsgründen die Spannung in jener Feder und dem Dämpfer gleich sein müssen, folgt der

⁸Formal entsprechen Retardations- und Relaxationszeit derselben Zeitkonstante. Die unterschiedliche Terminologie ergibt sich allein aus den jeweiligen physikalischen Interpretationen im *Kelvin-Voigt*- und *MAXWELL*-Modell

⁹Hier sind viele Bezeichnung verbreitet, häufig zu finden sind auch *POYNTING-THOMPSON*-Modell, *Drei-Parameter-Modell* oder *MAXWELL-ZENER*-Modell.

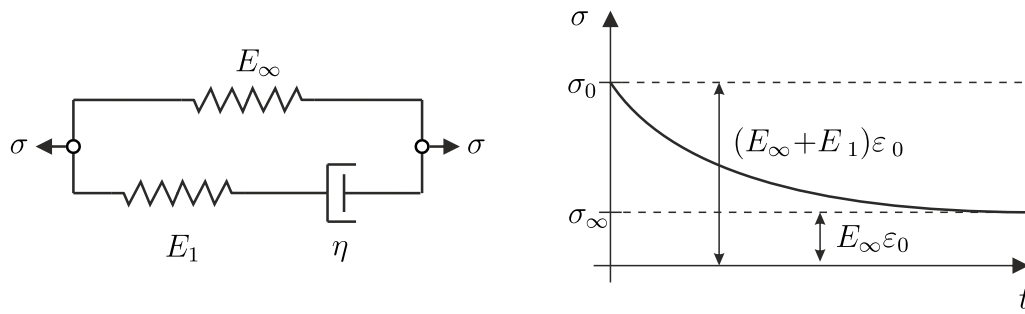


Abbildung 3.3: Linearer Standardkörper und dazugehöriges Relaxationsverhalten

konstitutive Zusammenhang

$$\sigma = (E_\infty + E_1) \varepsilon - E_1 \varepsilon_D. \quad (3.12)$$

Aus dem lokalen Gleichgewicht zwischen Feder und Dämpfer kann die differentielle Beziehung für die Entwicklung der viskosen Dehnung

$$\dot{\varepsilon}_D + \frac{1}{\tau} \varepsilon_D = \frac{1}{\tau} \varepsilon \quad \text{mit} \quad \tau := \frac{\eta}{E_1} \quad (3.13)$$

abgeleitet werden. Die Beziehung (3.13) wird auch als Evolutionsgleichung bezeichnet, da sie die zeitliche Entwicklung innerer Variablen, hier der viskosen Dehnungen, über die Zeit beschreibt. Alternativ kann die konstitutive Beziehung durch Einführung eines Integrationsfaktors $e^{t/\tau}$ hergeleitet werden (vgl. SIMO [122]). Mit diesem lässt sich Gleichung (3.13) auch schreiben als

$$\frac{d}{dt} [e^{t/\tau} \cdot \varepsilon_D] = \frac{1}{\tau} e^{t/\tau} \varepsilon. \quad (3.14)$$

Die anschließende Integration liefert

$$\varepsilon_D(t) = e^{-t/\tau} \cdot \varepsilon_D(-\infty) + \frac{1}{\tau} + \int_{-\infty}^t e^{(-t-s)/\tau} \varepsilon(s) ds \quad (3.15)$$

wobei der erste Summand der rechten Seite mit der Anfangsbedingung

$$\lim_{t \rightarrow -\infty} \varepsilon(t)_D = 0 \quad (3.16)$$

entfällt. Die partielle Integration von (3.15), unter Berücksichtigung der Bedingung $\varepsilon(t) = 0$ für $t \rightarrow -\infty$ und Einsetzen der daraus erhaltenen Beziehungen für $\varepsilon_D(t)$ in (3.12) führt auf das Faltungsintegral

$$\sigma(t) = \int_{-\infty}^t E(t-s) \dot{\varepsilon}(s) ds. \quad (3.17)$$

Gleichung (3.17) ist dahingehend vorteilhaft, da für beliebige Dehnungsverläufe $\varepsilon(s)$ keine explizite Lösung der differentiellen Beziehungen notwendig ist. Sie bildet zudem die Grundlage der numerischen Formulierung. Erforderlich ist dafür lediglich die Relaxationsfunktion

$$E(t) = E_\infty + E_1 e^{-t/\tau}, \quad (3.18)$$

die sich aus den Anteilen der Feder und des MAXWELL-Elements zusammensetzt. Das Relaxationsverhalten ist exemplarisch in Abbildung 3.3 zu sehen. Mit der Definition des initialen Elastizitätsmoduls

$$E_0 := E_\infty + E_1 \quad (3.19)$$

führt eine konstante Dehnung ε_0 zum Zeitpunkt $t = 0$ auf eine Spannungsantwort $\sigma_0 = E_0 \varepsilon_0$ welche aus Gleichgewichtsgründen für $t \rightarrow \infty$ auf den Wert $\sigma_\infty = E_\infty \varepsilon_0$ asymptotisch abklingt. Entsprechend sind durch die Anordnung der Feder, die Dehnungen im Falle eines Kriechversuchs begrenzt, womit beide viskoelastischen Effekte zufriedenstellend abgebildet werden.

Um zusätzlich die Flexibilität des Modells zu erhöhen, kann durch Hinzufügen weiterer Feder-Dämpfer-Elemente die Anzahl der freien Materialparameter des Modells erhöht werden. Dies wird als verallgemeinertes MAXWELL-Modell bezeichnet, vgl. Abbildung 3.4. Analog zu den zuvor gemachten Überlegungen folgt die Relaxationsfunktion

$$E(t) = E_\infty + \sum_{i=1}^M E_i e^{-t/\tau_i} \quad (3.20)$$

und weiterhin die Definitionen

$$\tau_i = \frac{\eta_i}{E_i}, \quad E_0 := E_\infty + \sum_{i=1}^M E_i \quad (3.21)$$

mit der Anzahl von MAXWELL-Elementen M . Die dazugehörige verallgemeinerten Evolutionsgleichungen lauten

$$\dot{\varepsilon}_{D_i} + \frac{1}{\tau_i} \varepsilon_{D_i} = \frac{1}{\tau_i} \varepsilon \quad \text{mit} \quad i = 1, \dots, M. \quad (3.22)$$

Das Modell bietet nicht nur ausreichend Flexibilität, sondern beinhaltet aufgrund der mathematischen Struktur der Relaxationsfunktion praktische Vorteile bei der algorithmischen Formulierung, da die Geschichte des Materials durch Zustandsvariablen ersetzt werden kann.

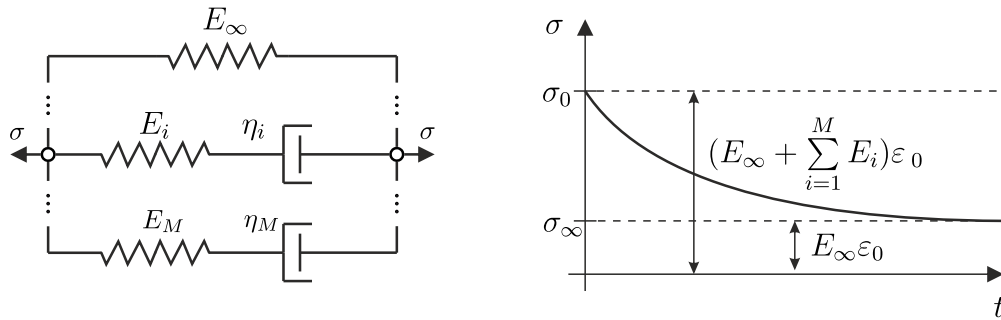


Abbildung 3.4: Verallgemeinertes MAXWELL-Modell und dazugehöriges Relaxationsverhalten

3.2 Thermodynamische Grundlagen und Einordnung der Modelle

Die bisherigen Modelle basieren rein auf differentiellen Gleichgewichtsbeziehungen, ohne weitere thermodynamische Überlegungen mit einzubeziehen. Die Thermodynamik liefert allerdings die theoretischen Rahmenbedingungen, auf dem eine verallgemeinerte Formulierung aufgebaut werden kann. Einhergehend lassen sich dadurch grundsätzliche Prinzipien der Materialmodellierung diskutieren, die im Laufe der Arbeit – insbesondere bei der Diskussion der verwendeten viskoelastischen Materialmodelle auf Basis künstlicher neuronaler Netze – noch benötigt werden. Zudem werden damit die bislang eindimensionalen Zusammenhänge auf mehrdimensionale Beziehungen erweitert. Die Formulierung erfolgt in den objektiven Größen $\mathbf{E}$ und $\mathbf{S}$ – die hier in der tensoriellen Notation verwendet werden – um dem Prinzip der materiellen Objektivität zu genügen.

3.2.1 Energiefunktionen mit inneren Variablen

Für die Herleitung der konstitutiven Beziehungen aus dem thermodynamischen Kontext wird die Existenz einer freien Energiefunktion der Form

$$\Psi = \Psi(\mathbf{E}, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n) \quad (3.23)$$

vorausgesetzt, die für isotherme Prozesse mit homogener Temperaturverteilung zulässig ist. Für Materialien, die durch Nicht-Gleichgewichtspfade¹⁰ charakterisiert sind, ist neben den externen Variablen – hier der Verzerrungstensor $\mathbf{E}$ – ein zusätzlicher Satz n an inneren Variablen $\boldsymbol{\xi}_i$, mit $i = 1, \dots, n$, notwendig, um die

¹⁰Ein System befindet sich im Gleichgewicht, wenn es ohne Änderung der äußeren Lasten keine Zustandsänderung erfährt. Kriechen und Relaxation sind Beispiele für Prozesse, in denen das System Nicht-Gleichgewichtszustände durchläuft (vgl. LUBLINER [87]).

Energiefunktion zu beschreiben. Innere Variablen lassen sich nicht direkt messen, häufig kann ihnen jedoch eine physikalische Größe zugeordnet werden. Für eine ausführliche Diskussion zur Bedeutung von inneren Variablen im Rahmen der Thermodynamik wird auf LUBLINER [86, 87] verwiesen. Sie beschreiben die Vorgeschichte des Materials, wobei hier davon ausgegangen wird, dass nicht die Kenntnis der gesamten Vorgeschichte, sondern nur der Zustand der Variablen zum aktuellen Zeitpunkt notwendig ist. Entsprechend ist ein funktionaler Zusammenhang ϱ_i notwendig, der die Entwicklung der inneren Variablen über die Zeit beschreibt:

$$\dot{\xi}_i(t) = \varrho_i(\mathbf{E}, \xi_1, \dots, \xi_n). \quad (3.24)$$

Eine Anforderung an die o.g. Gleichung, die ebenfalls als Evolutionsgleichung identifiziert werden kann, ist, dass im Falle des Gleichgewichts die Änderung der inneren Variablen zu Null werden

$$\varrho_i(\mathbf{E}, \xi_1, \dots, \xi_n) = \mathbf{0}. \quad (3.25)$$

Für rein elastische Materialien entfallen die inneren Variablen, da jeder Zustand des Materials einem Gleichgewichtspunkt entspricht und alle Prozesse reversibel sind¹¹. Dies motiviert eine additive Zerlegung der Energie in einen freien elastischen Anteil $\mathcal{W}(\mathbf{E})$ und in die dissipativen Anteile $\Gamma_i(\mathbf{E}, \xi_i)$, womit für Gleichung (3.23) folgt¹²

$$\Psi(\mathbf{E}, \xi_1, \dots, \xi_n) = \mathcal{W}(\mathbf{E}) + \sum_{i=1}^n \Gamma_i(\mathbf{E}, \xi_i). \quad (3.26)$$

Häufig wird bei viskoelastischen Materialien beobachtet, dass sich nur die deviatorischen Anteile viskos verhalten (vgl. FERRY [30]). Diese Beobachtung motiviert die Zerlegung des GREEN-LAGRANGE-Verzerrungsmaßes

$$\mathbf{E} = \frac{1}{3}\Theta\mathbf{I}_3 + \bar{\mathbf{E}} \quad (3.27)$$

in die Volumendehnungen Θ und die deviatorischen Verzerrungsanteile $\bar{\mathbf{E}}$:

$$\Theta = \text{tr}[\mathbf{E}], \quad \bar{\mathbf{E}} = \mathbf{E} - \frac{1}{3}\Theta\mathbf{I}_3 \quad \text{und} \quad \text{tr}[\bar{\mathbf{E}}] = E_{ii}. \quad (3.28)$$

Entsprechend wird die Trennung der freien Energie (3.26) wie folgt vorgenommen:

$$\Psi(\mathbf{E}, \xi_1, \dots, \xi_n) = \mathcal{W}_{\text{vol}}(\Theta) + \mathcal{W}_{\text{dev}}(\bar{\mathbf{E}}) + \sum_{i=1}^n \Gamma_i(\bar{\mathbf{E}}, \xi_i). \quad (3.29)$$

¹¹Im Falle ratenunabhängiger Plastizität können irreversible Prozesse auch in Abwesenheit von Nicht-Gleichgewichtszuständen auftreten, s. LUBLINER [87].

¹²Die Definition der elastischen Verzerrungsenergiedichtefunktion in Abhängigkeit von $\mathbf{E}$ ist im Falle finiter Elastizität, auch Hyperelastizität genannt, eher unüblich. Gängig ist die Definition in Abhängigkeit des rechten CAUCHY-GREEN-Verzerrungstensors $\mathbf{C}$, vgl. TRUESDELL und NOLL [134] oder MALVERN [88]. Da im Rahmen der Arbeit nur lineare viskoelastische Materialgesetze behandelt werden, wird diesem Umstand im Weiteren keine Aufmerksamkeit gewidmet.

An dieser Stelle lässt sich die Normalisierungsbedingung¹³ einführen. Unter der Annahme einer spannungsfreien Referenzkonfiguration folgt

$$\mathcal{W}_{\text{vol}}(0) = 0, \quad \mathcal{W}_{\text{dev}}(\mathbf{0}) = 0, \quad \sum_{i=1}^n \Gamma_i(\mathbf{0}, \mathbf{0}) = 0. \quad (3.30)$$

Nun kann der zweite Hauptsatz der Thermodynamik herangezogen werden¹⁴, um die konstitutiven Beziehungen aus den Energiefunktionen abzuleiten. Unter der Beachtung der rein mechanischen Größen lautet die CLAUSIUS-PLANK Ungleichung

$$\mathbf{S} : \dot{\mathbf{E}} - \dot{\Psi}(\mathbf{E}, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n) \geq 0. \quad (3.31)$$

Die Ungleichung fordert, dass die Differenz zwischen Spannungsleistung und Änderungsrate der freien Energie nicht negativ ist. Im rein elastischen Fall gilt die Gleichheit, sodass die gesamte Spannungsleistung in Form gespeicherter Energie wiedergewonnen werden kann. Bei dissipativen Materialmodellen geht ein Teil der Spannungsleistung irreversibel verloren. Diese Bedingung stellt eine zentrale Restriktion für die Konstruktion konsistenter Materialmodelle dar. Unter Anwendung der Kettenregel folgt

$$\dot{\Psi}(\mathbf{E}, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n) = \frac{\partial \Psi(\mathbf{E}, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n)}{\partial \mathbf{E}} : \dot{\mathbf{E}} + \sum_{i=1}^n \frac{\partial \Psi(\mathbf{E}, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n)}{\partial \boldsymbol{\xi}_i} : \dot{\boldsymbol{\xi}}_i \quad (3.32)$$

und durch Einsetzen in die Ungleichung (3.31)

$$\left(\mathbf{S} - \frac{\partial \Psi(\mathbf{E}, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n)}{\partial \mathbf{E}} \right) : \dot{\mathbf{E}} - \sum_{i=1}^n \frac{\partial \Psi(\mathbf{E}, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n)}{\partial \boldsymbol{\xi}_i} : \dot{\boldsymbol{\xi}}_i \geq 0. \quad (3.33)$$

Um die Ungleichung für beliebige $\dot{\mathbf{E}}$ zu erfüllen, folgt die Beziehung für den zweiten PIOLA-KIRSCHOFFschen Spannungstensor

$$\mathbf{S} = \frac{\partial \Psi(\mathbf{E}, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n)}{\partial \mathbf{E}} = \mathbf{S}_{\text{vol}} + \mathbf{S}_{\text{dev}} \quad (3.34)$$

mit

$$\mathbf{S}_{\text{vol}} = \frac{\partial \mathcal{W}_{\text{vol}}(\Theta)}{\partial \mathbf{E}}, \quad \mathbf{S}_{\text{dev}} = \frac{\partial \mathcal{W}_{\text{dev}}(\bar{\mathbf{E}})}{\partial \mathbf{E}} + \sum_{i=1}^n \frac{\partial \Gamma_i(\bar{\mathbf{E}}, \boldsymbol{\xi}_i)}{\partial \mathbf{E}}. \quad (3.35)$$

Folglich kann die Definition der Nicht-Gleichgewichtsspannungen

$$\mathbf{Q}_i = \frac{\partial \Gamma_i(\bar{\mathbf{E}}, \boldsymbol{\xi}_i)}{\partial \mathbf{E}} \quad (3.36)$$

¹³In Abhängigkeit der gewählten Variablen zur Beschreibung von Ψ muss die Normalisierungsbedingung angepasst werden. Bei der Verwendung von $\mathbf{C}$ gilt diese für $\mathbf{C} = \mathbf{1}$ bzw. $\text{tr}[\mathbf{C}] = 1$.

¹⁴Formal werden der erste und zweite Hauptsatz der Thermodynamik bereits benötigt, um die Form der freien Energie abzuleiten. Für die entsprechenden Herleitungen wird auf bspw. LUBLINER [87] verwiesen.

erfolgen, womit sich die deviatorischen Spannungsanteile wie folgt ausdrücken lassen

$$\mathbf{S}_{\text{dev}} = \frac{\partial \mathcal{W}_{\text{dev}}(\bar{\mathbf{E}})}{\partial \mathbf{E}} + \sum_{i=1}^n \mathbf{Q}_i \quad (3.37)$$

Ferner lässt sich aus Gleichung (3.33) die Bedingung an die Dissipation

$$\mathcal{D} := - \sum_{i=1}^n \frac{\partial \Psi(\mathbf{E}, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n)}{\partial \boldsymbol{\xi}_i} : \dot{\boldsymbol{\xi}}_i \geq 0 \quad (3.38)$$

ableiten, die wiederum durch eine geeignete Wahl der Evolutionsgleichung für $\dot{\boldsymbol{\xi}}$, vgl. Gleichung (3.24), erfüllt werden muss. Da im Falle des Gleichgewichts keine Dissipation vorliegt, lassen sich die treibenden thermodynamischen Kräfte $\mathbf{Q}_i$ auch aus der internen konstitutiven Beziehung

$$\mathbf{Q}_i = - \frac{\partial \Psi(\mathbf{E}, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n)}{\partial \boldsymbol{\xi}_i} = - \frac{\partial \Gamma_i(\bar{\mathbf{E}}, \boldsymbol{\xi}_i)}{\partial \boldsymbol{\xi}_i} \quad (3.39)$$

bestimmen. Weiterhin lässt sich daraus schließen, dass ohne Dissipation im Gleichgewichtszustand $\mathbf{Q}_i = \mathbf{0}$ gelten muss. Das Material verhält sich rein elastisch.

Damit wurde das thermodynamische Fundament zusammengefasst. Notwendig ist zum einen die Wahl geeigneter Funktionen für die Beschreibung der freien Energie und der Dissipation sowie zum anderen die Konstruktion von Evolutionsgleichungen, die der Dissipationsungleichung genügen.

3.2.2 Eindimensionales verallgemeinertes Maxwell-Modell

Auf Grundlage der vorangegangenen Überlegungen wird im Folgenden das rheologisch motivierte verallgemeinerte MAXWELL-Modell in den thermodynamischen Rahmen eingeordnet. Der Begriff der Einordnung wird dabei bewusst gewählt: Die Modellbeziehungen lassen sich nicht strikt aus den thermodynamischen Prinzipien herleiten; vielmehr soll gezeigt werden, dass die gewählten Annahmen mit den im vorhergehenden Kapitel formulierten Bedingungen konsistent sind¹⁵. Vergleichbare Einordnungen sind u.a. in HOLZAPFEL [55] und Simo [122] zu finden. Im Eindimensionalen reduziert sich die freie Energie aus Gleichung (3.26) zu

$$\Psi(\varepsilon, \boldsymbol{\xi}) = W(\varepsilon) + \sum_{i=1}^n \Gamma_i(\varepsilon, \xi_i), \quad (3.40)$$

mit den Komponenten $(\xi_1, \dots, \xi_n)$ des Vektors der inneren Variablen $\boldsymbol{\xi}$. Physikalisch entspricht diese freie Energie der in den Federn des rheologischen Modells gespeicherten Energie, während die inneren Variablen die zeitabhängigen, viskosen

¹⁵Der Vollständigkeit halber sei auf die Ausführungen in KRAWIETZ [72] hingewiesen, in denen eine Einordnung in die Theorie der *einfachen Stoffe* erfolgt.

Anteile repräsentieren. Unter dieser Voraussetzung motiviert sich die Darstellung der Gleichung (3.40) als

$$\Psi(\varepsilon, \boldsymbol{\xi}) = \frac{1}{2} E_{\infty} \varepsilon^2 + \frac{1}{2} \sum_{i=1}^M E_i (\varepsilon - \xi_i)^2 \quad (3.41)$$

worin sich die Anteile

$$W(\varepsilon) = \frac{1}{2} E_{\infty} \varepsilon^2, \quad \Gamma_i(\varepsilon, \xi_i) = \frac{1}{2} E_i (\varepsilon - \xi_i)^2 \quad (3.42)$$

identifizieren lassen. Weiter lassen sich die inneren Variablen den viskosen Dehnungen des Dämpfers zuordnen $\xi_i := \varepsilon_{Di}$. Damit entspricht die Anzahl der inneren Variablen n der Anzahl der MAXWELL-Elemente M . Die Anteile erfüllen jeweils die Normalisierungsbedingung

$$W(0) = 0, \quad \Gamma_i(0, 0) = 0. \quad (3.43)$$

Die Spannungen lassen sich mit Gleichung (3.34) berechnen

$$\sigma = \frac{\partial \Psi(\varepsilon, \boldsymbol{\xi})}{\partial \varepsilon} = E_{\infty} \varepsilon + \sum_{i=1}^n E_i (\varepsilon - \xi_i) = E_{\infty} \varepsilon + \sum_{i=1}^n Q_i, \quad (3.44)$$

was für $n = 1$ wieder auf die Beziehung (3.12) führt. Die Dissipation des Modells lässt sich angeben mit

$$\mathcal{D}(\varepsilon, \boldsymbol{\xi}, \dot{\boldsymbol{\xi}}) = - \sum_{i=1}^n \frac{\partial \Psi(\varepsilon, \boldsymbol{\xi})}{\partial \xi_i} \dot{\xi}_i = \sum_{i=1}^n E_i (\varepsilon - \xi_i) \dot{\xi}_i = \sum_{i=1}^n Q_i \dot{\xi}_i \geq 0, \quad (3.45)$$

wobei hier erkennbar ist, dass sich die Nicht-Gleichgewichtsspannungen aus der internen konstitutiven Beziehung ableiten lassen und gerade den Spannungen im Dämpfer entsprechen. Einsetzen des Stoffgesetzes $Q_i = \eta_i \dot{\xi}_i$ in die Ungleichung (3.45) führt zu

$$\mathcal{D}(\varepsilon, \boldsymbol{\xi}, \dot{\boldsymbol{\xi}}) = \sum_{i=1}^n \eta_i (\dot{\xi}_i)^2 \geq 0, \quad (3.46)$$

was für $\eta_i > 0$ stets erfüllt und somit auch die Evolutionsgleichung (3.22) im thermodynamischen Sinne verträglich ist. Häufig werden die Evolutionsgleichungen auch in den Nicht-Gleichgewichtsspannungen dargestellt. Mit den Beziehungen

$$\dot{Q}_i = E_i (\dot{\varepsilon} - \dot{\xi}_i), \quad \dot{\xi}_i = \frac{Q_i}{\eta_i} \quad (3.47)$$

folgt

$$\dot{Q}_i + \frac{Q_i}{\tau_i} = E_i \dot{\varepsilon}. \quad (3.48)$$

Hierbei sei angemerkt, dass in der Literatur unterschiedliche Darstellungen der Evolutionsgleichung zu finden sind. Gleichung (3.48) folgt der Notation in HOLZAPFEL [55]. In SIMO [122] werden die Nicht-Gleichgewichtsspannungen mit

$$q_i := E_i \xi_i \quad (3.49)$$

definiert, was zu einer Modifikation der rechten Seite von Gleichung (3.48) führt.

3.2.3 Mehrdimensionale Erweiterung

Abschließend sollen die wesentlichen Aspekte des mehrdimensionalen Modells erläutert werden. In Anlehnung an SIMO [121] werden für die freien elastischen Anteile der Energiegleichung (3.29) folgende Darstellungen

$$\mathcal{W}_{\text{vol}}(\Theta) = \frac{1}{2} K_{\infty} \Theta^2 \quad \mathcal{W}_{\text{dev}}(\bar{\mathbf{E}}) = G_{\infty} \bar{\mathbf{E}} : \bar{\mathbf{E}}, \quad (3.50)$$

mit dem Kompressionsmodul K_{∞} und dem Schubmodul der Einzelfeder des verallgemeinerten MAXWELL-Modells G_{∞} , gewählt, wobei isotropes Materialverhalten vorausgesetzt wird. Die explizite Definition eines Potentials zur Beschreibung der dissipativen Anteile ist nicht zwingend erforderlich; mögliche Konstruktionen sind jedoch in SIMO [121, 122] zu finden. Eine zulässige Erweiterung der eindimensionalen Beziehung (3.42) ist sinngemäß

$$\Gamma_i(\bar{\mathbf{E}}, \boldsymbol{\xi}_i) = \frac{1}{2} G_i (\bar{\mathbf{E}} - \boldsymbol{\xi}_i)^2, \quad (3.51)$$

mit den Schubmoduln der Einzelfedern G_i . Ebenfalls auf Basis der eindimensionalen Überlegungen (3.48) lässt sich die mehrdimensionale Form der Evolutionsgleichungen nach HOLZAPFEL [55] angeben:

$$\dot{\mathbf{Q}}_i + \frac{\mathbf{Q}_i}{\tau_i} = G_i \dot{\bar{\mathbf{E}}} = \dot{\mathbf{S}}_{i,\text{dev}} \quad \text{mit } i = 1, \dots, n. \quad (3.52)$$

Die konstitutiven Beziehungen lassen sich aus Gleichung (3.35) ableiten. Der volumetrische Spannungsanteil ist damit unabhängig von der Zeit und berechnet sich zu

$$\mathbf{S}_{\text{vol}} = \frac{\partial \mathcal{W}_{\text{vol}}(\Theta)}{\partial \Theta} = K_{\infty} \Theta \mathbf{1}. \quad (3.53)$$

Die deviatorischen Spannungsanteile sind wiederum von der Zeit abhängig

$$\mathbf{S}_{\text{dev}} = 2G_{\infty} \bar{\mathbf{E}} + \sum_{i=1}^n \mathbf{Q}_i. \quad (3.54)$$

Die Integration der Evolutionsgleichungen (3.52) - analog zu dem Vorgehen in Kapitel 3.1, s. Gleichungen (3.13)-(3.17) - und anschließendes Einsetzen in (3.54) führt auf das Faltungsintegral für den Spannungsdeviator

$$\mathbf{S}_{\text{dev}} = \int_{-\infty}^t G(t-s) \dot{\bar{\mathbf{E}}}(s) ds \quad (3.55)$$

wobei die Relaxationsfunktion mit

$$G(t) = 2G_{\infty} + \sum_{i=1}^n G_i e^{-t/\tau_i} \quad (3.56)$$

definiert ist. Damit wurden die notwendigen mehrdimensionalen Zusammenhänge des verallgemeinerten MAXWELL-Modells eingeordnet, sodass im folgenden Kapitel die algorithmische Umsetzung diskutiert werden kann.

3.3 Algorithmische Umsetzung

Im Folgenden werden die wesentlichen Schritte zur Integration der Evolutionsgleichungen dargelegt, um unter Einbau der Randbedingungen für die Numerik geeignete Vorschriften zu erhalten. Diese werden nicht nur im Rahmen der FEM benötigt. Sie werden ebenfalls genutzt, um am Materialpunkt synthetische Daten zu generieren, die im späteren Verlauf der Arbeit für den Aufbau eines Materialmodells auf Basis eines künstlichen neuronalen Netzes benötigt werden. Die Herleitungen orientieren sich an den Notationen die in SIMO [122] und RUST [116] zu finden sind. Alle bereitgestellten Algorithmen wurden in MATLAB [131] implementiert.

3.3.1 Eindimensionale Implementierung

Die Herleitungen erfolgen zunächst für den eindimensionalen Fall auf Basis der in Kapitel 3.1 hergeleiteten Zusammenhänge. Ausgangspunkt für die Numerik ist das Faltungsintegral¹⁶ (3.17), das hier aus Gründen der Vollständigkeit nochmals angegeben wird.

$$\sigma(t) = \int_{-\infty}^t E(t-s) \dot{\varepsilon}(s) ds \quad \text{mit} \quad E(t) = E_{\infty} + \sum_{i=1}^n E_i e^{-t/\tau_i} \quad (3.57)$$

Sowohl die Relaxationsfunktion $E(t)$ als auch der Dehnungsverlauf $\varepsilon(s)$ werden als bekannt vorausgesetzt. Betrachtet wird die Zeit im Intervall $t \in [T_0, T]$ mit und $T > T_0$. Im Folgenden wird davon ausgegangen, dass alle Prozesse bei $T_0 = 0$ beginnen. Mit den gegebenen Beziehungen lässt sich der Ausdruck (3.57) zu

$$\sigma(t) = \int_0^t E_{\infty} \dot{\varepsilon}(s) ds + \sum_{i=1}^n \int_0^t E_i e^{-(t-s)/\tau_i} \dot{\varepsilon}(s) ds \quad (3.58)$$

$$= E_{\infty} \varepsilon(t) + \underbrace{\sum_{i=1}^n \int_0^t E_i e^{-(t-s)/\tau_i} \dot{\varepsilon}(s) ds}_{h^i(t)} \quad (3.59)$$

umformen. Der wesentliche Vorteil der Darstellung der Relaxationsfunktion liegt in der Möglichkeit, die Identität

$$e^{(t_{n+1}+\Delta t_{n+1})/\tau_i} = e^{t_{n+1}/\tau_i} e^{\Delta t_{n+1}/\tau_i} \quad (3.60)$$

mit der Definition

$$t = t_{n+1} = t_n + \Delta t_{n+1} \quad (3.61)$$

¹⁶Alternativ können die Evolutionsgleichungen auch mit geeigneten Zeitintegrationsverfahren gelöst werden. Eine Diskussion hierzu ist in WRIGGERS [152] zu finden.

für die Berechnung des Terms $h^i(t)$ auszunutzen. Es folgt mit $h_{n+1}^i := h^i(t)$

$$h_{n+1}^i = \int_0^{t_n + \Delta t_{n+1}} E_i e^{-(t_n + \Delta t_{n+1} - s)/\tau_i} \dot{\varepsilon}(s) ds \quad (3.62)$$

$$= e^{-\Delta t_{n+1}/\tau_i} \left[\underbrace{\int_0^{t_n} E_i e^{-(t_n - s)/\tau_i} \dot{\varepsilon}(s) ds}_{h_n^i} + \underbrace{\int_{t_n}^{t_{n+1}} E_i e^{-(t_{n+1} - s)/\tau_i} \dot{\varepsilon}(s) ds}_{\Delta h_{n+1}^i} \right] \quad (3.63)$$

$$= e^{-\Delta t_{n+1}/\tau_i} [h_i(t_n) + \Delta h_{n+1}^i] . \quad (3.64)$$

Gleichung (3.64) ist eine Rekursionsformel, da für die Bestimmung der viskosen Spannungsanteile h_{n+1}^i lediglich die Kenntnis deren Zustandes im letzten Zeitschritt und deren Zuwachses Δh_{n+1}^i im aktuellen Zeitinkrement Δt_{n+1} notwendig ist. Für die Berechnung des Zuwachses wird im Rahmen der Numerik die Dehngeschwindigkeit über

$$\dot{\varepsilon} \approx \frac{\Delta \varepsilon_{n+1}}{\Delta t_{n+1}}, \quad \Delta \varepsilon_{n+1} = \varepsilon_{n+1} - \varepsilon_n, \quad (3.65)$$

unter der Annahme, dass $\dot{\varepsilon}$ innerhalb eines Zeitschrittes konstant ist (vgl. SIMO [122]), approximiert. Damit verbleibt

$$\Delta h_{n+1}^i = E_i \frac{\Delta \varepsilon_{n+1}}{\Delta t_{n+1}} \int_{t_n}^{t_{n+1}} e^{-(t_{n+1} - s)/\tau_i} ds . \quad (3.66)$$

Der Integralausdruck in Gleichung 3.66 kann entweder mittels numerischer Integration oder nach TAYLOR ET AL. [129] analytisch gelöst werden¹⁷. Im Falle der analytischen Integration wird der folgende Term erhalten

$$\Delta h_{n+1}^i = E_i \frac{\Delta \varepsilon_{n+1}}{\Delta t_{n+1}} \tau_i \left(1 - e^{-\Delta t_{n+1}/\tau_i} \right) . \quad (3.67)$$

Mit den hergeleiteten Beziehungen folgt für die Spannung die Berechnungsvorschrift

$$\sigma_{n+1} = E_\infty(\varepsilon_n + \Delta \varepsilon_{n+1}) + \sum_{i=1}^n e^{-\Delta t_{n+1}/\tau_i} h_n^i + \sum_{i=1}^n E_i \frac{\Delta \varepsilon_{n+1}}{\Delta t_{n+1}} \tau_i \left(1 - e^{-\Delta t_{n+1}/\tau_i} \right) . \quad (3.68)$$

¹⁷Die analytische Lösung basiert auf der Annahme der Temperaturunabhängigkeit der viskoelastischen Materialparameter, für Details s. TAYLOR ET AL. [129] oder RUST [116]. Im Falle von Temperaturabhängigkeit muss das Integral numerisch gelöst werden.

Initialisierung:	$\varepsilon_n = 0, h_n^i = 0$
Gegeben	$\varepsilon_n, \varepsilon_{n+1}, t_n, h_n^i$
Zeitschleife	for $t_{n+1} = 0, \dots, T$
	$\Delta t_{n+1} = t_{n+1} - t_n$
	$\Delta \varepsilon_{n+1} = \varepsilon_{n+1} - \varepsilon_n$
	$\sigma_0 = 0, C_0 = 0$
Schleife über Maxwell-Elemente	for $i = 1, \dots, n$
	$\delta_i := e^{-\Delta t_{n+1}/\tau_i}$
	$\Delta h_{n+1}^i = E_i \frac{\Delta \varepsilon_{n+1}}{\Delta t_{n+1}} \tau_i (1 - \delta_i)$
	$h_{n+1}^i = \delta_i h_n^i + \Delta h_{n+1}^i$
	$\sigma_i = \sigma_{i-1} + h_{n+1}^i$
	$C_i = C_{i-1} + \frac{E_i}{\Delta t_{n+1}} \tau_i (1 - \delta_i)$
	end for
	$\sigma_{n+1} = E_\infty \varepsilon_{n+1} + \sigma_i$
	$C_{n+1} = E_\infty + C_i$
	end for
Update Variablen	$\varepsilon_n = \varepsilon_{n+1}$
	$h_n^i = h_{n+1}^i$

Algorithmus 3.1: Eindimensionale Berechnungsvorschrift für lineare Viskoelastizität

Letztlich wird im Rahmen der FEM noch die Tangente C für die Linearisierung der Materialgleichung benötigt. Mit der Kettenregel folgt

$$C_{n+1} = \frac{\partial \sigma_{n+1}}{\partial \varepsilon_{n+1}} = \frac{\partial \sigma_{n+1}}{\partial \Delta \varepsilon_{n+1}} \frac{\partial \Delta \varepsilon_{n+1}}{\partial \varepsilon_{n+1}} = \frac{\partial \sigma_{n+1}}{\partial \Delta \varepsilon_{n+1}} \frac{\partial (\varepsilon_{n+1} - \varepsilon_n)}{\partial \varepsilon_{n+1}} = \frac{\partial \sigma_{n+1}}{\partial \Delta \varepsilon_{n+1}}. \quad (3.69)$$

Entsprechende Differenzierung der Gleichung (3.68) liefert

$$C_{n+1} = E_\infty + \sum_{i=1}^n E_i \frac{\tau_i}{\Delta t_{n+1}} \left(1 - e^{-\Delta t_{n+1}/\tau_i}\right), \quad (3.70)$$

womit alle notwendigen Formeln zur Implementierung des eindimensionalen Stoffgesetzes hergeleitet wurden. Die Berechnungsvorschrift ist zusammenfassend in Algorithmus 3.1 gegeben. Abschließend bleibt anzumerken, dass die Tangente unabhängig von dem aktuellen Verzerrungszustand ist und lediglich von der aktuellen Zeitschrittweite abhängt. Damit ist das Stoffgesetz linear in den Verzerrungen definiert und dahingehend ist keine lokale Newton-Iteration am Materialpunkt erforderlich. Dieser Aspekt wird im späteren Verlauf der Arbeit nochmals aufgegriffen.

3.3.2 Mehrdimensionale Implementierung

Nach den bereits zuvor vorgestellten eindimensionalen Gleichungen werden nachfolgend die wesentlichen Unterscheidungen zu der dreidimensionalen Darstellung bereitgestellt. Während die eindimensionalen Zusammenhänge in den linearen Größen ε und σ formuliert wurden, erfolgt konsistent mit Kapitel 3.2.3 die Formulierung in $\mathbf{E}$ und $\mathbf{S}$, bzw. hier direkt in Vektor-Matrix Notation in $\mathbf{E}^v$ und $\mathbf{S}^v$. Die volumetrischen $\mathbf{E}_{\text{vol}}^v$ und deviatorischen Verzerrungsanteile $\mathbf{E}_{\text{dev}}^v$ können über die Beziehungen

$$\mathbf{E}_{\text{vol}}^v = \frac{1}{3} \underbrace{\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{\mathbf{P}^{\text{vol}}} \mathbf{E}^v \quad \text{und} \quad \mathbf{E}_{\text{dev}}^v = \underbrace{\begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & 0 & 0 & 0 \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{P}^{\text{dev}}} \mathbf{E}^v \quad (3.71)$$

bestimmt werden (vgl. ZIENKIEWICZ ET AL.) [155]. Unter der Annahme, dass sich nur die deviatorischen Verzerrungsanteile viskos verhalten, wird deren Rate

$$\dot{\mathbf{E}}_{\text{dev}}^v \approx \frac{\Delta \mathbf{E}_{\text{dev}_{n+1}}^v}{\Delta t_{n+1}} \quad (3.72)$$

benötigt. Mit der sinngemäßen mehrdimensionalen Erweiterung von Gleichung (3.67)

$$\Delta \mathbf{h}_{n+1}^i = G_i \frac{\Delta \mathbf{E}_{\text{dev}_{n+1}}^v}{\Delta t_{n+1}} \tau_i \left(1 - e^{-\Delta t_{n+1}/\tau_i}\right) \in \mathbb{R}^{6 \times i} \quad (3.73)$$

folgt die Aktualisierung der Nicht-Gleichgewichtsspannungen

$$\mathbf{h}_{n+1}^i = e^{-\Delta t_{n+1}/\tau_i} \left[\mathbf{h}_n^i + \Delta \mathbf{h}_{n+1}^i \right] \quad (3.74)$$

und entsprechend die konstitutive Beziehung

$$\mathbf{S}_{n+1}^v = K_\infty \mathbf{E}_{\text{vol}_{n+1}}^v + 2G_\infty \mathbf{E}_{\text{dev}_{n+1}}^v + \sum_{i=1}^n \mathbf{h}_{n+1}^i. \quad (3.75)$$

Die Materialtangente berechnet sich mittels

$$\mathbf{C}_T = K_\infty \mathbf{P}^{\text{vol}} + \left[2G_\infty + \sum_{i=1}^n \frac{G_i}{\Delta t_{n+1}} \tau_i \left(1 - e^{-\Delta t_{n+1}/\tau_i}\right) \right] \mathbf{P}^{\text{dev}}. \quad (3.76)$$

Der Algorithmus 3.1 kann sinngemäß mit den gegebenen Anpassungen übernommen werden. Bei isotropen Materialverhalten erfolgt die Angabe der Materialparameter häufig mittels des initialen Elastizitätsmoduls E_0 und der initialen

Querkontraktion ν_0 . Entsprechend berechnen sich

$$K_0 = \frac{E}{3(1 - 2\nu_0)} \quad \text{und} \quad G_0 = \frac{E_0}{2(1 - \nu_0)}. \quad (3.77)$$

Mit den Definitionen

$$\mu_i := \frac{G_i}{G_0} \quad \text{und} \quad \mu_0 := 1 - \sum_{i=1}^n \mu_i \quad (3.78)$$

erhält man die Gleichgewichtsmoduln

$$K_\infty = K_0 \mu_0 \quad \text{und} \quad G_\infty = G_0 \mu_0. \quad (3.79)$$

Kommentar zum ebenen Spannungszustand

Bei ebenen Problemen wird häufig auf eine zweidimensionale Modellierung zurückgegriffen. Im ebenen Spannungszustand wird dabei die Annahme getroffen, dass $S_{33} = 0$ gilt, jedoch $E_{33} \neq 0$. Im rein elastischen Fall lässt sich E_{33} dabei über die Beziehung

$$E_{33} = \frac{-\nu}{1 - \nu}(E_{11} + E_{22}) \quad (3.80)$$

bestimmen. Im Falle viskoelastischen Materialverhaltens ist die Querkontraktion ν von der Zeit abhängig, womit die Berechnung der deviatorischen und volumetrischen Anteile nicht ohne weiteres möglich ist. Dahingehend muss der volumetrische Anteil θ als Geschichtsvariable mitgeführt und pro Zeitschritt entsprechend aktualisiert werden

$$\theta_{n+1} = \theta_n + \Delta\theta_{n+1}. \quad (3.81)$$

Weitere Details und die Implementierung sind in RUST [116] zu finden.

3.4 Zeitabhängige viskoelastische Materialversuche

Das letzte Unterkapitel widmet sich einer kurzen Beschreibung der klassischen Versuche zur Bestimmung der viskoelastischen Materialparameter. Diese werden im späteren Verlauf der Arbeit nochmals aufgegriffen, um synthetische Daten für den Aufbau von Materialmodellen mit künstlichen neuronalen Netzen zu erzeugen. Ausführlichere Beschreibungen hierzu sind in CHRISTENSEN [21] und FERRY [30] speziell für Viskoelastizität und im erweiterten Kontext z.B. in LEMAITRE ET AL. [82] zu finden.

Grundsätzlich unterliegen Materialversuche der Voraussetzung, dass sich an einem Materialpunkt eindeutig der Spannungs-Dehnungszustand identifizieren lässt. Dies wiederum ist am einfachsten umsetzbar, wenn diese homogen sind, sprich

$$\mathbf{S} \neq \mathbf{S}(\mathbf{X}), \quad \mathbf{E} \neq \mathbf{E}(\mathbf{X}), \quad (3.82)$$

wobei $\mathbf{S}$ und $\mathbf{E}$ nicht die gemessenen Größen darstellen, sondern beispielsweise aus gemessenen Verschiebungen und der technischen Spannung $\mathbf{P}$ berechnet werden. Rein experimentell lässt sich dies am einfachsten mit uniaxialen Versuchsaufbauten durchführen. Bi- und Triaxialversuche sind grundsätzlich möglich. Zum Erreichen homogener - oder zumindest bereichsweise homogener - Felder sind jedoch deutlich aufwändigere Aufbauten notwendig. Für isotrope Materialien gilt im uniaxialen Fall exemplarisch (vgl. LEMAITRE ET AL. [82]):

$$\mathbf{S} = \begin{bmatrix} S_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{und} \quad \mathbf{E} = \begin{bmatrix} E_{11} & 0 & 0 \\ 0 & E_{22}^* & 0 \\ 0 & 0 & E_{33}^* \end{bmatrix}. \quad (3.83)$$

E_{11} ist hierbei die unabhängige Variable, während sich die Komponenten E_{22}^* und E_{33}^* aus den Querkontraktionseigenschaften des Materials ergeben. Die unabhängige Variable kann grundsätzlich jede beliebige Komponente des Verzerrungstensors sein.

Für viskoelastische Materialien motiviert die Darstellung (3.17) die wohl naheliegendste Art des uniaxialen Versuchs, den Relaxationsversuch. Dafür wird ein Verzerrungszustand

$$E_{11}(t) = E_{11}^0, \quad t > 0 \quad (3.84)$$

vorgegeben, womit sich die Relaxationsfunktion durch

$$E(t) = \frac{S_{11}(t)}{E_{11}^0} \quad (3.85)$$

bestimmen lässt. Die freien Materialparameter der Relaxationsfunktion können dann durch eine Schätzung ermittelt werden, s. z.B. KUNTSCHKE [76]. Entsprechend kann durch Vorgabe einer konstanten Spannung

$$S_{11}(t) = S_{11}^0, \quad t > 0 \quad (3.86)$$

durch Messung der Dehnung über die Zeit die Kriechfunktion

$$J(t) = \frac{E_{11}(t)}{S_{11}^0} \quad (3.87)$$

bestimmt werden. Neben Kriech- und Relaxationsversuchen besteht die Möglichkeit, mittels zyklischer Versuche die Materialparameter zu ermitteln. Dafür wird eine harmonische Dehnung der Form

$$E_{11}(t) = E_{11}^0 \sin(\omega t) \quad (3.88)$$

vorgegeben. Aus der Spannungsantwort des Materials lassen sich die frequenzabhängigen Anteile der Relaxationsfunktion – der Speichermodul $E'(\omega)$ und der Verlustmodul $E''(\omega)$ – ableiten. Auch hier wird für weitere Ausführung auf die bereits genannte Literatur verwiesen.

4 Einführung in die Mikromechanik

In Kapitel 2.1 wurde der Begriff des *Partikels* eingeführt, der sich als Materialpunkt des Körpers $\mathcal{B}$ identifizieren lässt. Tatsächlich handelt es sich dabei um ein konzeptuelles Volumenelement, das auf der betrachteten Skala als punktförmig erscheint. Zwangsläufig ist die Definition des Materialpunkts damit skalenabhängig, da angenommen wird, dass innerhalb dieses Punktes keine weiteren relevanten Wechselwirkungen mehr auftreten, die auf der gewählten Betrachtungsskala berücksichtigt werden müssen. Rein theoretisch müsste solch ein Punkt auf der atomaren Ebene angesiedelt sein. Die Modellierung eines Körpers mit all seinen atomaren Punkten wäre jedoch in der Praxis nicht sinnvoll, zum einen hinsichtlich des Rechenaufwands und zum anderen haben atomare Effekte keine Relevanz für Ingenieur Anwendungen. Stattdessen ist ein Materialpunkt im mechanischen Sinne als eine Repräsentativregion des Körpers zu verstehen, in der die physikalischen Wechselwirkungen auf kleineren Skalen bereits durch eine effektive konstitutive Beziehung abgebildet werden.

In bestimmten Fällen müssen die Wechselwirkungen zwischen größeren und kleineren Skalen des Körpers explizit berücksichtigt werden, da die effektive konstitutive Beziehung noch nicht bekannt ist. Das Konzept der Mehrskalenmodellierung führt daher neben der Makroskala – also der Skala des materiellen Körpers – eine kleinere, sogenannte Mikroskala ein. Auf der Makroskala wird die Strukturebene beschrieben, also die Geometrie des Bauteils mit den zugehörigen Randbedingungen. Die Mikroskala erfasst hingegen die Inhomogenitäten des Materials, wie etwa Poren, Einschlüsse, Risse oder unterschiedliche Materialphasen. Die Aufgabe der Mikromechanik besteht darin, eine Brücke zwischen beiden Skalen zu schlagen. Dazu gehören die Festlegung geeigneter Randbedingungen für eine konsistente Kopplung der Randwertprobleme, die Definition von Volumenmittelwertsätzen – der *Homogenisierung* – sowie die Herleitung makroskopisch wirksamer Materialgesetze.

Die Anfänge der Mikromechanik lassen sich bis ins 19. Jahrhundert zurückverfolgen. Im Bereich der Festkörpermechanik sind dabei die Arbeiten von VOIGT [139] und REUSS [109] zu nennen, die sich mit der Abschätzung effektiven Steifigkeiten, bzw. Nachgiebigkeiten, von Vielkristallen auf Basis der Theorie von Einkristallen auseinander gesetzt haben. Ihre Ansätze – basierend auf der Annahme gleichmäßiger Dehnung (VOIGT) bzw. Spannung (REUSS) – bilden die äußeren Schranken effektiver Materialparameter. Einen bedeutenden Fortschritt stellte in der zweiten Hälfte des 20. Jahrhunderts der Beitrag von HILL [54] dar, der die Äquivalenz der Formänderungsenergie auf Mikro- und Makroskala postulierte – ein Konzept, das als Grundlage für viele Homogenisierungsverfahren dient und auch im weiteren Verlauf dieser Arbeit Anwendung findet. In jüngerer Zeit hat sich der Fokus

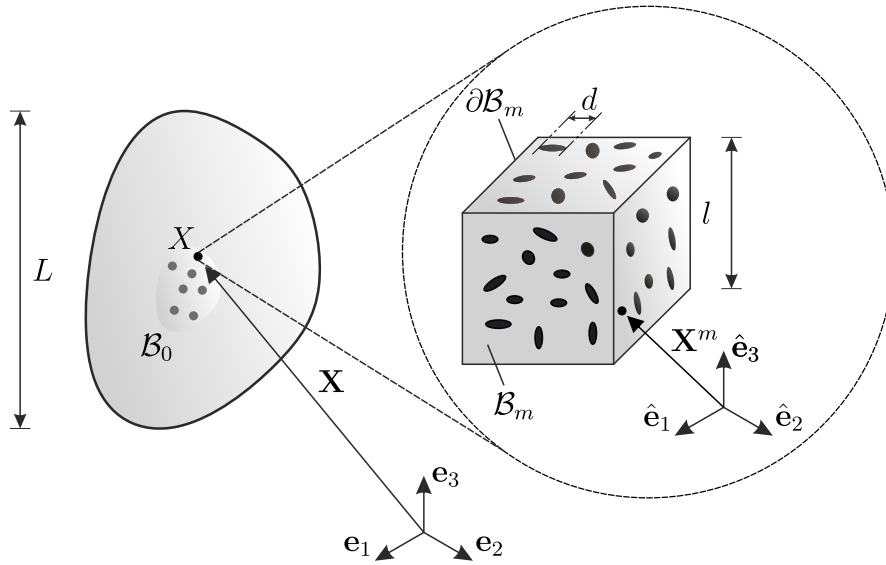


Abbildung 4.1: Makroskopischer Körper und repräsentatives Volumenelement

zunehmend auf die numerische Behandlung von Mehrskalenproblemen sowie die Modellierung inelastischen Materialverhaltens verlagert, etwa im Rahmen sogenannter FE^2 -Methoden.

Im Folgenden werden die grundlegenden Konzepte der Mikromechanik im Rahmen der für diese Arbeit relevanten Fragestellungen dargestellt. In diesem Kapitel wird dabei von einem dreidimensionalen Kontinuum auf beiden Skalen ausgegangen. In Kapitel 5 werden die Beziehungen auf Basis der nachfolgenden Ausführungen für die Kopplung von Schalen und Kontinua diskutiert. Ausgangspunkt ist die Definition einer geeigneten Einheitszelle, um die wesentlichen Merkmale der Mikroskala abzubilden. Darauf aufbauend wird die Mittelung mikroskopischer Deformations- und Spannungsfelder diskutiert, um abschließend eine konsistente Kopplung zur Makroskala ermöglichen. Als begleitende Literatur sei auf ABOUDI [3], TORQUATO [132], oder auch GROSS und SEELIG [43] verwiesen. Die numerische Behandlung von Mehrskalenproblemen wird ausführlich bei ZHODI und WRIGGERS [156] behandelt.

4.1 Das repräsentative Volumenelement als Einheitszelle

Die Verwendung des Mehrskalenkonzepts erfordert eine entsprechende Definition der verwendeten Skalen. Die Makroskala ist in der Regel einfach zu identifizieren und beschreibt beispielsweise im bisher diskutierten Kontext den Körper in seiner Referenzkonfiguration $\mathcal{B}_0$. Diesem wird eine Länge L zugeordnet, die dessen makroskopische Skala charakterisiert, s. Abbildung 4.1. Im allgemeinen Kontext wird davon ausgegangen, dass die Mikrostruktur des Körpers aus rein zufällig angeord-

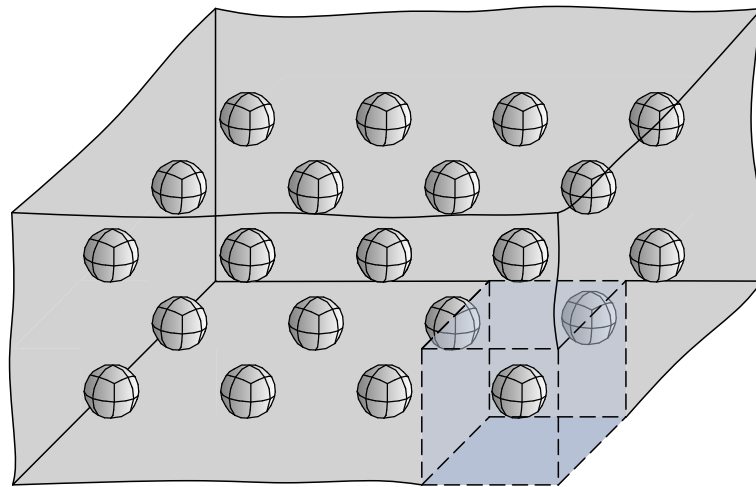


Abbildung 4.2: Ausschnitt einer periodischen Mikrostruktur und Darstellung eines möglichen RVEs

neten Inhomogenitäten besteht womit die Definition der Einheitszelle dem Konzept des *repräsentativen Volumenelements* (RVE) folgt. Die geometrische Skalenseparation erfolgt dann unter der Annahme, dass für ein RVE mit Abmessungen l und darin enthalten Inhomogenitäten mit charakteristischem Durchmesser d die Bedingung der Skalenseparation (vgl. HASHIN [51])

$$d \ll l \ll L \quad (4.1)$$

erfüllt ist. Dieses Konzept basiert auf der Tatsache, dass durch die geometrische Trennung der Skalen zwangsläufig genug Inhomogenitäten innerhalb eines RVEs vorhanden sein müssen, um die Bedingung der statistischen Repräsentativität¹⁸ zu erfüllen. Statistisch repräsentativ bedeutet hier, dass die globalen Eigenschaften, wie beispielsweise die Volumenanteile der unterschiedlichen Materialphasen des RVEs, unabhängig von der materiellen Koordinate $\mathbf{X}$ der Makroskala sind (vgl. HASHIN [51]). Streng genommen ist diese Eigenschaft nur für eine mathematisch unendliche Anzahl von eingeschlossenen Inhomogenitäten erfüllt (vgl. STROEVEN ET AL. [126]).

In dieser Arbeit werden ausschließlich periodische Mikrostrukturen behandelt. Beispiele hierfür sind Faserverbundlamine, *Honeycomb*-Strukturen oder, wie später auch gezeigt wird, Bauteile mit periodisch auftretenden sphärischen Einschlüssen, wie in Abbildung 4.2 exemplarisch dargestellt ist. Diese stellen Sonderfälle dar, bei denen das RVE als die kleinstmögliche sich wiederholende Einheitszelle definiert wird.¹⁹ Die Gesamtstruktur setzt sich dann aus einer periodischen

¹⁸TORQUATO [132] definiert den Begriff „statistischen Repräsentativität“ mathematische auf Basis einer n -Punkt Wahrscheinlichkeitsdichtefunktion.

¹⁹Ausgenommen sind Probleme, bei denen Instabilitäten auf der Mikrostruktur auftreten. Das

Anordnung dieser zusammen. Dies ist exemplarisch in Abbildung 4.2 dargestellt. Angemerkt sei dazu, dass für die dargestellte Struktur viele Definitionen des RVEs zulässig sind, da dreidimensionale ideal periodische Mikrostrukturen translatorisch invariant bezüglich aller ihrer drei lokalen Basisvektoren sind (vgl. MICHEL ET AL. [93]). Im Rahmen der Numerik erfolgt die Wahl des RVEs dahingehend meist in Abhängigkeit von lokalen Symmetrieebenen, was unter anderem die Aufbringung der Randbedingungen vereinfachen kann. Es sei jedoch angemerkt, dass die weiteren Methoden unabhängig von zufälligen oder periodischen Mikrostrukturen sind und lediglich voraussetzen, dass ein geeignetes RVE identifiziert wurde.

4.2 Dreidimensionaler Mikro-Makro Übergang

Die grundlegendes Konzept der Mikromechanik ist die *Homogenisierung* der Mikrofelder. *Homogenisierung* bedeutet dabei, dass die Makrogrößen aus der Mittelung der Mikrofelder erhalten werden. Im Folgenden wird dafür die Definition des Mittelwerts $\langle \bullet \rangle$ einer Mikrogröße $(\bullet)$ eingeführt

$$\langle \bullet \rangle = \frac{1}{V} \int_{\mathcal{B}^m} (\bullet) dV \quad \text{mit} \quad V = \int_{\mathcal{B}^m} dV, \quad (4.2)$$

die durch den hochgestellten Index m gekennzeichnet wird. $\mathcal{B}^m$ bezeichnet damit das RVE in seiner Referenzkonfiguration, vgl. auch Abbildung 4.1. Entsprechend lässt sich so der gemittelte Deformationsgradient

$$\langle \mathbf{F}^m \rangle = \frac{1}{V} \int_{\mathcal{B}^m} \mathbf{F}^m(\mathbf{X}^m, t) dV = \frac{1}{V} \int_{\partial \mathcal{B}^m} \mathbf{x}^m(\mathbf{X}^m, t) \otimes \mathbf{N}^m dA \quad (4.3)$$

sowie der gemittelte erste PIOLA-KIRCHOFF Spannungstensor

$$\langle \mathbf{P}^m \rangle^T = \frac{1}{V} \int_{\mathcal{B}^m} \mathbf{P}^{mT}(\mathbf{X}^m, t) dV = \frac{1}{V} \int_{\partial \mathcal{B}^m} \mathbf{T}^m(\mathbf{X}^m, t) \otimes \mathbf{X}^m dA. \quad (4.4)$$

berechnen. Sowohl in Gleichung (4.3) als auch in (4.4) wurde der GAUSSsche Integralsatz angewendet, was die Berechnung der effektiven Größen allein über die Auswertung der dargestellten Integrale auf dem Rand ermöglicht.

Die grundlegende Fragestellung die sich im Rahmen der *Homogenisierung* stellt ist, in welcher Art und Weise eine makroskopische Deformation $\mathbf{F}$ aufgebracht werden muss, um daraus auf dem RVE die homogenisierten Spannungen $\langle \mathbf{P}^m \rangle$ zu ermitteln, unter der Voraussetzung, dass Gleichung (4.3) erfüllt ist. Hier lässt sich ohne weitere Annahmen zeigen, dass zum einen eine vorgeschriebene Deformation

RVE muss dann so gewählt werden, dass die maßgebenden periodischen Beulformen abgebildet werden (vgl. z. B. MIEHE ET AL. [96]).

mittels eines konstanten Deformationsgradienten $\mathbf{F} = \mathbf{F}^0$ auf dem Rand $\partial\mathcal{B}^m$ des RVEs die Gleichung (4.3) erfüllt. Mit den vorgegebenen Verschiebungen

$$\mathbf{x}^m = \mathbf{F}^0 \mathbf{X}^m \quad \text{auf} \quad \partial\mathcal{B}^m \quad (4.5)$$

folgt durch Einsetzen in (4.3)₂ gerade

$$\langle \mathbf{F}^m \rangle = \frac{1}{V} \mathbf{F}^0 \int_{\partial\mathcal{B}^m} \mathbf{X}^m \otimes \mathbf{N}^m dA = \mathbf{F}^0 \quad \text{mit} \quad \int_{\partial\mathcal{B}^m} \mathbf{X}^m \otimes \mathbf{N}^m dA = V. \quad (4.6)$$

Dies wird auch als *Satz über den mittleren Deformationsgradienten* bezeichnet²⁰. Zum anderen folgt durch Vorgabe einer konstanten Randspannung $\mathbf{P}^{0T}$

$$\mathbf{T}^m = \mathbf{P}^{0T} \mathbf{N}^m \quad \text{auf} \quad \partial\mathcal{B}^m \quad (4.7)$$

und Einsetzen in Gleichung (4.4)₂ der *Satz über die mittlere Spannung*

$$\langle \mathbf{P}^m \rangle^T = \frac{1}{V} \mathbf{P}^{0T} \int_{\partial\mathcal{B}^m} \mathbf{N}^m \otimes \mathbf{X}^m dA = \mathbf{P}^{0T}. \quad (4.8)$$

Eine ausführliche Herleitung ist unter anderem in SAEB ET AL. [117] zu finden. Die beiden Sätze motivieren damit die Äquivalenzen

$$\mathbf{F} \equiv \langle \mathbf{F}^m \rangle \quad \text{und} \quad \mathbf{P}^T \equiv \langle \mathbf{P}^m \rangle^T. \quad (4.9)$$

sodass die mikroskopischen Mittelwerte von $\mathbf{F}^m$ und $\mathbf{P}^{mT}$ gerade den makroskopischen Größen $\mathbf{F}$ und $\mathbf{P}^T$ entsprechen.

Grundsätzlich kann die *Homogenisierung* in beliebigen Spannungs- und Verzerrungsgrößen erfolgen. Während im Kontext finiter Deformation die Darstellung in $\langle \mathbf{F}^m \rangle$ und $\langle \mathbf{P}^m \rangle$ geläufig ist, findet sich bei geometrisch und materiell linearen Problemen die Formulierung in den Ingenieurverzerrungen $\langle \boldsymbol{\varepsilon}^m \rangle$ und den CAUCHY-Spannungen $\langle \boldsymbol{\sigma}^m \rangle$. Hier lassen sich die Beziehungen (4.3) und (4.4) sowie die darauffolgenden Sätze (4.6) und (4.8) in gleicher Weise anwenden. Aus den homogenisierten Größen können über die in den Kapiteln 2.1.1 und 2.1.2 hergeleiteten kontinuumsmechanischen Beziehungen weitere makroskopische Größen berechnen. Für die makroskopischen Verzerrungen und Spannungen gilt beispielsweise

$$\mathbf{E} = \frac{1}{2} \left(\langle \mathbf{F}^m \rangle^T \langle \mathbf{F}^m \rangle - \mathbf{1} \right) \quad \text{und} \quad \mathbf{S} = \langle \mathbf{P}^m \rangle^T \langle \mathbf{F}^m \rangle^{-1}. \quad (4.10)$$

Zu beachten ist dabei, dass die Aussage (4.9) nicht sinngemäß übertragbar ist, d.h.

$$\mathbf{E} \neq \langle \mathbf{E}^m \rangle \quad \text{und} \quad \mathbf{S} \neq \langle \mathbf{S}^m \rangle. \quad (4.11)$$

²⁰Das lineare Pendant ist der *Satz über die mittlere Verzerrung* mit $\boldsymbol{\varepsilon} = \langle \boldsymbol{\varepsilon}^m \rangle$.

Dies ist der Nichtlinearität der beiden Größen geschuldet. So lässt sich beispielsweise leicht für das GREEN-LAGRANGE Verzerrungsmaß erkennen

$$\langle \mathbf{E}^m \rangle = \frac{1}{2V} \int_{\mathcal{B}^m} (\mathbf{F}^{mT}(\mathbf{X}, t) \mathbf{F}^m(\mathbf{X}, t) - \mathbf{1}) dV \neq \frac{1}{2} (\langle \mathbf{F}^m \rangle^T \langle \mathbf{F}^m \rangle - \mathbf{1}), \quad (4.12)$$

wie beispielsweise in NEMAT-NASSER [99] ausführlich beschrieben. Im Falle finiter Deformation ist dahingehend eine Formulierung rein in $\langle \mathbf{E} \rangle$ und $\langle \mathbf{S} \rangle$ eher unüblich, da zusätzliche Überlegungen über die Definition der Randbedingungen in eben diesen notwendig sind. Einen Sonderfall stellen Probleme mit großen Verschiebungen aber kleinen Verzerrungen dar, was beispielsweise bei Schalentragwerken relevant wird. Dieser Fall wird in Kapitel 5 diskutiert.

Für den konsistenten Einsatz der Homogenisierungsschemas im Rahmen der FEM ist letztlich das makroskopische Spannungsinkrement

$$\Delta \langle \mathbf{P}^m \rangle^T = \frac{\partial \langle \mathbf{P}^m \rangle^T}{\partial \langle \mathbf{F}^m \rangle} \Delta \langle \mathbf{F}^m \rangle \quad (4.13)$$

zu berechnen. Da im Allgemeinen keine analytische Beziehung für den makroskopischen Tangenten-Modul

$$\mathbb{A} := \frac{\partial \langle \mathbf{P}^m \rangle^T}{\partial \langle \mathbf{F}^m \rangle} \quad (4.14)$$

vorliegt, wird dieser in der Regel numerisch berechnet. Eine ausführliche Diskussion hierzu überschreitet den Rahmen dieses Kapitels, als weiterführende Literatur wird auf MIEHE [94] sowie TEMIZER und WRIGGERS [130] verwiesen. Auch auf analytische Herleitungen des Elastizitätsgesetzes

$$\langle \boldsymbol{\sigma}^m \rangle = \mathbb{C}^* \langle \boldsymbol{\varepsilon}^m \rangle \quad (4.15)$$

mit dem effektiven Elastizitätstensors $\mathbb{C}^*$ wird nicht weiter eingegangen, welche z.B. in GROSS [43] zu finden sind²¹. Die im Rahmen der Arbeit benötigte konsistente makroskopische Tangente bei der Kopplung von Schalen und Kontinua wird in Kapitel 5.3 hergeleitet.

4.3 Herleitung mikroskopischer Randbedingungen

Zur Lösung des Randwertproblems auf der Mikroebene und Berechnung der makroskopischen Größen sind geeignete Randbedingung zu formulieren. Im vorherigen Kapitel konnte bereits mit den Gleichungen (4.6) und (4.8) über die Volumennittelwertssätze die Zulässigkeit von uniformen Deformationen und konstanten

²¹Sowohl der makroskopische Tangentenmodul als auch der effektive Elastizitätstensor sind im Allgemeinen nicht deren mikroskopische Volumennittelwerte, d.h. $\mathbb{A} \neq \langle \mathbb{A}^m \rangle$ und $\mathbb{C}^* \neq \langle \mathbb{C}^m \rangle$. Letzteres kann jedoch als oberer Grenzwert für die effektiven Eigenschaften des RVEs identifiziert werden und wird auch als VOIGT-Schranke bezeichnet.

Spannungen als Randwerte gezeigt werden. Im Weiteren wird ein energetischer Ansatz nach HILL [54] eingeführt, welcher die Identifikation weiterer zulässiger Bedingungen auf dem Rand ermöglicht. Das entsprechende Lemma postuliert die Äquivalenzen der virtuellen Arbeiten auf der Mikro- und Makroskala²². Im Falle finiter Deformationen und im allgemein thermodynamischen Kontext wird die sogenannte HILL-MANDEL-Bedingung mit geeigneten arbeitskonjugierten Größen ausgedrückt, z.B.:

$$\langle \mathbf{P}^{mT} : \delta \mathbf{F}^m \rangle - \mathbf{P}^T : \delta \mathbf{F} \stackrel{!}{=} 0. \quad (4.16)$$

Durch entsprechende Umformung, wie sie z.B. in SAEB [117] zu finden ist, lässt sich die Gleichung (4.16) als ein Integral über den Rand des RVEs schreiben

$$\langle \mathbf{P}^{mT} : \delta \mathbf{F}^m \rangle - \mathbf{P}^T : \delta \mathbf{F} = \int_{\partial \mathcal{B}_0} (\delta \mathbf{x}^m - \delta \mathbf{F} \mathbf{X}^m) (\mathbf{T}^m - \mathbf{P}^T \mathbf{N}^m) dA. \quad (4.17)$$

Um die HILL-MANDEL-Bedingung (4.16) zu erfüllen, muss das Randintegral von Gleichung (4.17) zu Null werden, woraus sich die folgenden zulässigen Vorgaben ableiten lassen:

- Konstante Verschiebungen für jeden Punkt des RVEs

$$\mathbf{x}^m = \mathbf{F} \mathbf{X}^m \quad \text{in} \quad \mathcal{B}_0 \quad (4.18)$$

Dies entspricht gerade der VOIGT-Schranke, bei der davon ausgegangen wird, dass alle materiellen Punkte des RVEs die gleichen Deformationen erfahren. Die Vorgabe dieser Bedingung liefert damit einen oberen Grenzwert der effektiven Größen, da bei beliebiger Anordnung von Inhomogenitäten grundsätzlich weichere Phasen größeren und steifere Phasen kleineren Verformungen ausgesetzt sind. Die effektive Steifigkeit wird damit insgesamt überschätzt.

- Konstante Verschiebungen auf dem Rand des RVEs

$$\mathbf{x}^m = \mathbf{F} \mathbf{X}^m \quad \text{auf} \quad \partial \mathcal{B}^m \quad (4.19)$$

Dies entspricht der bereits eingeführten Gleichung (4.5).

- Periodische Verschiebungen auf dem Rand des RVEs

$$\mathbf{x}^{m+} - \mathbf{x}^{m-} = \mathbf{F} \mathbf{X}^{m+} - \mathbf{F} \mathbf{X}^{m-} \quad \text{auf} \quad \partial \mathcal{B}^m \quad (4.20)$$

²²Die Veröffentlichung von HILL postuliert korrekterweise zunächst die Gleichheit der Energien. Diese Forderung kann allerdings nur im Falle linearer Elastizität erfüllt werden. Die Formulierung in den virtuellen Arbeiten ist gültig für beliebiges Materialverhalten und kann alternativ auch in eine Darstellung in den Gleichheiten der Spannungsleistungen überführt werden.

Dabei wurde die Konvention $(...)^+$ und $(...)^-$ zur Bezeichnung zweier gegenüberliegenden symmetrischer Ränder eingeführt. Da nicht zwangsläufig gewährleistet ist, dass bei beliebigen Mikrostrukturen die Ränder des RVEs symmetrisch sind, eignen sich periodische Randbedingungen folglich vor allem für periodische Mikrostrukturen, was jedoch keine zwingende Voraussetzung darstellt.

- Konstante Spannungen auf dem Rand des RVEs

$$\mathbf{T}^m = \mathbf{P}^T \mathbf{N}^m \quad \text{auf} \quad \partial\mathcal{B}^m \quad (4.21)$$

Dies entspricht gerade wieder der Gleichung (4.7).

- Konstante Spannungen in jedem Punkt des RVEs

$$\mathbf{T}^m = \mathbf{P}^T \mathbf{N}^m \quad \text{in} \quad \mathcal{B}^m \quad (4.22)$$

Diese Vorgabe führt auf die REUSS-Schranke und stellt den unteren Grenzwert für die effektiven Größen dar. Ohne Anforderungen an die Anordnung von Inhomogenitäten in der Mikrostruktur unterschätzt diese Annahme grundsätzlich die effektiven Steifigkeit, da in der Realität die Spannungen in steiferen größer als in weicheren Phasen sind.

Die Randbedingungen (4.18)-(4.22) können dahingehend genutzt werden, um das Randwertproblem auf dem RVE zu lösen und die benötigten makroskopischen Größen zu berechnen. Eine zusammenfassende Übersicht über die Implementierung der Bedingungen im Rahmen der Numerik und deren Einfluss auf die Berechnung der makroskopischen Größen ist in, z. B., SAEB ET AL. [117] zu finden.

Alle Herleitung sind bislang im allgemeinen Kontext der Kopplung von Kontinua auf der Mikro- und Makroskala vorgenommen worden. Diese stellen die Basis für den eigentlichen Kern der Arbeit, die Kopplung von Schalen und Kontinua dar. Durch geeignete Modifikation der HILL-MANDEL-Bedingung können auch hier die entsprechenden Randbedingungen abgeleitet werden, womit der Einsatz des Konzepts für die Berechnung der effektiven Größen für Schalentragerwerke ermöglicht wird.

5 Mikro-Makro Übergang für Schalen

Die mikromechanischen Konzepte aus Kapitel 4 basieren auf der Annahme, dass sowohl auf der Makro- als auch auf der Mikroskala dreidimensionale Kontinua vorliegen. Handelt es sich auf der Makroskala jedoch um schlanke, biegedominante Flächentragwerke, bietet es sich erneut an, zur Modellierung des makroskopischen Verhaltens Schalenelemente zu verwenden. Für einige Arten von Inhomogenitäten ist dabei nicht direkt ein Mehrskalenmodell notwendig. Insbesondere bei geschichteten Materialien und rein linearem Materialverhalten stellt eine Dickenintegration einen effizienten Berechnungsansatz dar, der sich durch numerische Integration auch auf inelastisches Materialverhalten erweitern lässt. Für die Fälle, in denen dieses Vorgehen nicht angewendet werden kann, bietet sich erneut die Modellierung auf zwei Skalen an, bei welcher die Mikroskala weiterhin durch ein dreidimensionales Kontinuum beschrieben wird. Dies ermöglicht die Modellierung der effektiven Eigenschaften schlanker Tragwerke mit komplexer Mikrostruktur.

Die folgenden Ziele lassen im Rahmen des Kapitels identifizieren:

1. Das Mehrskalenkonzept für Schalen wird konsistent hergeleitet unter Verwendung einer modifizierten Form der HILL-MANDEL-Bedingung. Dadurch ergibt sich eine Alternative zu einem vollständig dreidimensionalen Modell zur Modellierung dünnwandiger Strukturen mit inhomogener Mikrostruktur.
2. Das Homogenisierungsschema kann alternativ zur Erzeugung von synthetischen Spannungs-Verzerrungs-Daten verwendet werden. Die zur Definition der Randbedingungen erforderlichen makroskopischen Größen werden dabei explizit vorgegeben, beispielsweise über uniaxiale Verzerrungszustände (vgl. Kapitel 3.4). Daraus lassen sich die zugehörigen homogenisierten Spannungen ermitteln. Diese synthetischen Materialversuche dienen der Erzeugung eines Datensatzes aus Verzerrungs-Spannungs-Paaren, der in Kapitel 7 genutzt wird, um mithilfe eines künstlichen neuronalen Netzes ein effektives Materialmodell zu entwickeln.

Dazu wird zunächst das mikroskopische Homogenisierungsproblem unter den Annahmen der Schalenkinematik definiert. Aufbauend auf der modifizierten HILL-MANDEL-Bedingung werden geeignete Randbedingungen für das RVE hergeleitet und die dabei speziell im Zusammenhang mit der Kopplung von Schalen und Kontinua auftretenden Problematiken diskutiert. Abschließend wird die numerische Umsetzung zur Berechnung der homogenisierten Schalenschnittgrößen sowie der effektiven Materialmatrix bereitgestellt. Da die Tensor-Notation im weiteren Verlauf der Arbeit nur noch lokal benötigt wird, erfolgt die Darstellung aller Größen

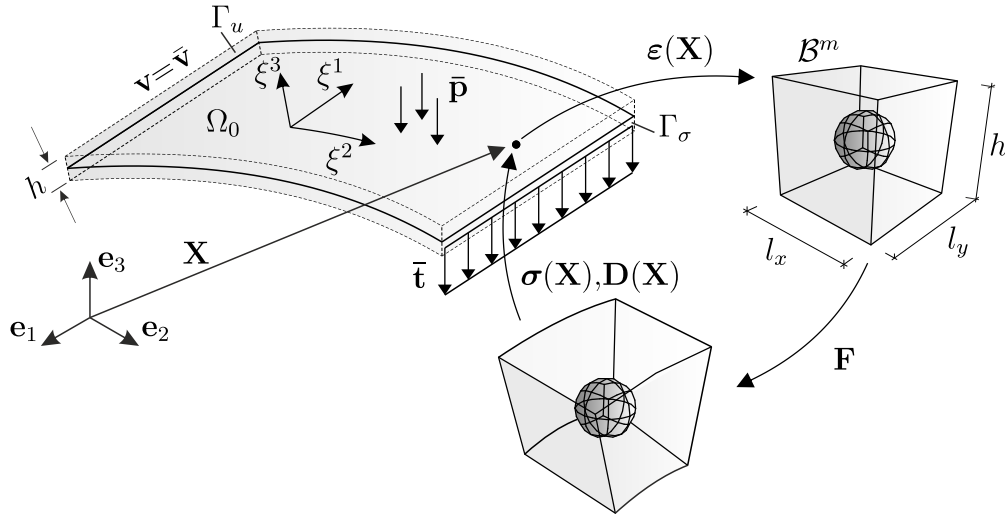


Abbildung 5.1: Makroskopisches Randwertproblem der Schale und schematische Darstellung der Homogenisierung der Mikroskala abgebildet durch ein RVE.

grundsätzlich in Vektor-Matrix Notation, d. h.,

$$\mathbf{E} := \mathbf{E}^v \quad \text{und} \quad \mathbf{S} := \mathbf{S}^v. \quad (5.1)$$

Auf die explizite Kennzeichnung als Mikrogröße mit dem hochgestellten Index m wird dabei im Weiteren verzichtet.

5.1 Mikroskopisches Homogenisierungsproblem

Im Folgenden wir das bislang unvollständig gestellte Randwertproblem der Schale aus Kapitel 2.4.1 aufgegriffen. Die entsprechende Lösung erfordert die Definition eines konstitutiven Zusammenhangs. In Anlehnung an die Ausführungen in GRUTTMANN und WAGNER [45], wird dem makroskopischen Materialpunkt, der später im Rahmen der Numerik dem Integrationspunkt entspricht, ein RVE mit der bereits in Kapitel 4.2 eingeführten Referenzkonfiguration $\mathcal{B}^m$ zugeordnet (vgl. Abbildung 5.1). Auf dem RVE mit Länge l_x , Breite l_y und Höhe h wird ein lokales kartesisches Koordinatensystem definiert. Das Gebiet $\mathcal{B}^m$ ist über die Bedingungen

$$-\frac{l_x}{2} \leq x \leq \frac{l_x}{2}, \quad -\frac{l_y}{2} \leq y \leq \frac{l_y}{2} \quad \text{und} \quad h^- \leq z \leq h^+ \quad (5.2)$$

begrenzt. Hierbei werden die x - und y -Koordinate, entsprechend der Referenzfläche der Schale, als Flächenkoordinaten gewählt, während z die Koordinate in Dickenrichtung beschreibt. Wird die Mittelfläche der Schale als Referenzfläche gewählt, so gilt:

$$h^- = -\frac{h}{2} \quad \text{und} \quad h^+ = +\frac{h}{2}. \quad (5.3)$$

Bei Schalen-RVEs entspricht deren Höhe stets derjenigen der Schale. Die Länge l_x und Breite l_y ergeben sich im Falle periodischen Mikrostrukturen durch die Wiederholung der Einheitszelle; bei Laminaten wiederum sind diese Größen beliebig, da unendlich viele zulässige Periodenlängen existieren²³.

Auf der Mikroskala wird das Materialverhalten der einzelnen Phasen, z.B. durch eine dreidimensionale viskoelastische konstitutive Beziehung beschrieben (vgl. Kapitel 3.2.3). Während das makroskopische Randwertproblem durch das Fehlen einer Materialgleichung unvollständig gestellt ist, müssen auf der Mikroskala geeignete Randbedingungen zur Berechnung des Verschiebungsfeldes $\mathbf{u}^m$ formuliert werden. Die Kopplung zwischen Mikro- und Makroskala wird durch Anwendung der HILL-MANDEL-Bedingung hergestellt. Dabei wird angenommen, dass die Schale großen Verformungen, aber nur kleinen Verzerrungen unterliegt. Makroskopisch können somit große Verformungen korrekt beschrieben werden, ohne dass die Randbedingungen über den makroskopischen Deformationsgradienten vorgegeben werden müssen. Auf dem RVE selbst wird von kleinen Verformungen ausgegangen, sodass $\mathbf{P}^T \approx \mathbf{F}\mathbf{S}$ gilt, womit eine Formulierung in den zweiten PIOLA-KIRCHOFF Spannungen möglich ist.

Mit diesen Annahmen lauten die Beziehungen zwischen den makroskopischen Schalen- und gemittelten mikroskopischen Kontinuumsgrößen demnach

$$\mathbf{A}\boldsymbol{\varepsilon} = \langle \mathbf{E} \rangle \quad \text{und} \quad \boldsymbol{\sigma} = \int_{-h/2}^{+h/2} \mathbf{A}^T \langle \mathbf{S} \rangle d\xi^3 \quad (5.4)$$

mit der bereits in Gleichung (2.94) in kompakter Darstellung eingeführten Transformationsmatrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & \xi^3 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \xi^3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \xi^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (5.5)$$

Dabei ist zu beachten, dass aufgrund der Annahme eines inextensiblen Direktors $\langle E_{33} \rangle$ gleich Null gilt. Gleichung (5.4)₁ wird im weiteren Verlauf noch für die Formulierung der Randbedingungen auf dem RVE benötigt.

Für die Herleitung der Randbedingungen muss die eingeführte HILL-MANDEL-Bedingung (4.16) entsprechend für die Kopplung von Schalen und Kontinua modifiziert werden. Zunächst ergibt sich unter der Annahme, dass Körper- bzw. Flächenlasten sowie Randlasten ausschließlich auf der Makroskala angesetzt werden

²³Die Problematik der Längenabhängigkeit bei Schalen RVEs wird in Kapitel 5.2 gesondert behandelt.

für die innere virtuelle Arbeit auf dem RVE:

$$\delta \Pi^m(\mathbf{u}^m, \delta \mathbf{u}^m) = \int_{\mathcal{B}^m} \delta \mathbf{E}^T \mathbf{S} \, dV. \quad (5.6)$$

Damit folgt über die Gleichheit der makro- und mikroskopischen virtuellen Arbeit

$$\frac{1}{h} \delta \boldsymbol{\varepsilon}^T \boldsymbol{\sigma} - \frac{1}{V_0} \int_{\mathcal{B}^m} \delta \mathbf{E}^T \mathbf{S} \, dV = 0. \quad (5.7)$$

Dabei bezeichnet $V_0 = l_x l_y h$ das unverformte Referenzvolumen des RVEs. Damit lässt sich die Bedingung (5.7) auch wie folgt schreiben:

$$\delta \boldsymbol{\varepsilon}^T \boldsymbol{\sigma} - \frac{1}{A_0} \int_{\mathcal{B}^m} \delta \mathbf{E}^T \mathbf{S} \, dV = 0. \quad (5.8)$$

Entsprechend werden die homogenisierten Größen durch Mittelung über die Referenzfläche des RVEs $A_0 = l_x l_y$ berechnet, s. hierzu auch CONEN ET AL. [24]. Aus Gleichung (5.8) lassen sich analog zum Vorgehen in Kapitel 4.3 zulässige Randbedingungen zum Lösen des mikroskopischen Randwertproblems ableiten.

5.2 Rand- und Nebenbedingungen für Schalen-RVEs

Die Annahme kleiner Verzerrungen makroskopisch, sowie kleiner Verformungen mikroskopisch erlaubt im Weiteren die Formulieren der Randbedingungen in den GREEN-LAGRANGE-Verzerrungen und den zweiten PIOLA-KIRCHHOFF Spannungen. Da im Rahmen der Arbeit ausschließlich periodische Randbedingungen verwendet werden, werden nur diese im Weiteren erläutert. Eine Formulierung von Gleichung (4.20) in den makroskopischen Verzerrungen führt auf

$$\mathbf{u}^+ = \mathbf{u}^- + \langle \mathbf{E} \rangle \Delta \mathbf{X}^m, \quad \Delta \mathbf{X}^m = \mathbf{X}^{m+} - \mathbf{X}^{m-}. \quad (5.9)$$

Dabei ist $\langle \mathbf{E} \rangle$ der mittlere GREEN-LAGRANGE-Verzerrungstensor. Im Gegensatz zu den dreidimensional formulierten periodischen Randbedingungen (4.20) werden bei Schalen-RVEs mit den Konventionen $(\dots)^+$ und $(\dots)^-$ nur die zu den x - und y -Richtungen senkrecht stehenden seitlich Flächen des RVEs angesprochen. Dies gewährleistet Konformität zu der Makroskala, da auf den Ober- und Unterseiten Spannungsfreiheit vorliegen muss und in diese Richtung weiter keine Periodizität vorliegt. Ausschreiben der Relation (5.4)₁ mit (5.5) zwischen den

mittleren Kontinuums- und Schalenverzerrungen führt auf die Einzelterme

$$\begin{aligned}
\langle E_{11} \rangle &= \varepsilon_{11} + z\kappa_{11} \\
\langle E_{22} \rangle &= \varepsilon_{22} + z\kappa_{22} \\
\langle E_{33} \rangle &= 0 \\
\langle E_{12} \rangle &= \langle E_{21} \rangle = \varepsilon_{12} + z\kappa_{12} \\
2\langle E_{13} \rangle &= 2\langle E_{31} \rangle = \gamma_1 \\
2\langle E_{23} \rangle &= 2\langle E_{32} \rangle = \gamma_2.
\end{aligned} \tag{5.10}$$

Das Einsetzen der Beziehungen (5.10) in (5.9) führt auf eine Darstellung der Randbedingungen in den makroskopischen Schalenverzerrungen

$$\begin{bmatrix} u_x^+ \\ u_y^+ \\ u_z^+ \end{bmatrix} = \begin{bmatrix} u_x^- \\ u_y^- \\ u_z^- \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} + z\kappa_{11} & \varepsilon_{12} + z\kappa_{12} & \varepsilon_{13} \\ \varepsilon_{12} + z\kappa_{12} & \varepsilon_{22} + z\kappa_{22} & \varepsilon_{23} \\ \varepsilon_{13} & \varepsilon_{23} & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}. \tag{5.11}$$

Unter der Voraussetzung, dass $\Delta z = 0$ für Knoten auf gegenüberliegenden Flächen $x = \pm l_x/2$ und $y = \pm l_y/2$ gilt, folgt die modifizierte Darstellung

$$\begin{bmatrix} u_x^+ \\ u_y^+ \\ u_z^+ \end{bmatrix} = \begin{bmatrix} u_x^- \\ u_y^- \\ u_z^- \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} + z\kappa_{11} & \varepsilon_{12} + z\kappa_{12} & 0 \\ \varepsilon_{12} + z\kappa_{12} & \varepsilon_{22} + z\kappa_{22} & 0 \\ 2\varepsilon_{13} & 2\varepsilon_{23} & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}. \tag{5.12}$$

Die Gleichung (5.12) lässt sich unter Einführung der Matrix

$$\bar{\mathbf{A}} = \begin{bmatrix} \Delta x & 0 & \frac{1}{2}\Delta y & \Delta x z & 0 & \frac{1}{2}\Delta y z & 0 & 0 \\ 0 & \Delta y & \frac{1}{2}\Delta x & 0 & \Delta y z & \frac{1}{2}\Delta x z & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \Delta x & \Delta y \end{bmatrix}. \tag{5.13}$$

umschreiben zu

$$\mathbf{u}^+ = \mathbf{u}^- + \bar{\mathbf{A}}(\Delta x, \Delta y, z) \boldsymbol{\varepsilon}(\mathbf{X}). \tag{5.14}$$

Die Bedingung $\mathbf{u}^+ = \mathbf{u}^-$ aus Gleichung (5.14) wird über geeignete Zwangsbedingungen realisiert. Dabei ist zu beachten, dass die ebenen Verschiebungskomponenten symmetrisch und die Verschiebungskomponenten aus der Ebene heraus antimetrisch, gemäß Tabelle 5.1, gekoppelt werden. Für weitere Details wird auf GRUTTMANN und WAGNER [46] verwiesen.

Im Fall von Schalen-RVEs hat sich gezeigt, dass die alleinige Vorgabe von periodischen Randbedingungen nach Gleichung (5.14) keine zufriedenstellende Lösung darstellt. Dabei lassen sich zwei wesentliche Problematiken identifizieren:

- Das Aufbringen einer Schubdeformation γ_α kann aus Gleichgewichtsgründen zu einer reinen Starrkörperbewegung des RVEs führen. Folglich werden

keine Querschubspannungen auf dem RVE aktiviert und entsprechend keine Querschubsteifigkeit. Auch hier wird für eine ausführliche Diskussion auf KLARMANN ET AL. [67] verwiesen.

- Die effektive Schubsteifigkeit des RVEs weist eine Abhängigkeit der Länge auf. Dies ist vor allem bei Mikrostrukturen problematisch, bei denen grundsätzlich die Wahl der Abmessungen l_x und l_y aufgrund unendlicher Periodizität beliebig sein sollte. Beim Aufbringen einer reinen makroskopischen Schubdeformation γ_α werden für große l/h -Verhältnisse auf dem RVE überwiegend Biegesteifigkeiten aktiviert. Dies ist der Tatsache geschuldet, dass eine reine Schubverformung nicht nur eine konstante Querschub-, sondern zusätzlich eine lineare Momentenverteilung in die x - und y -Richtung auf dem RVE liefert. Im Extremfall führt dies zu einem Abfall der Querschubsteifigkeit auf null. Dies ist anschaulich für Balken RVEs in KLARMANN ET AL. [67] dargestellt.

Zur Vermeidung beider zuvor genannten Problematiken kann der in KLARMANN ET AL. [67] vorgeschlagene Ansatz gemäß GRUTTMANN und WAGNER [46] für Schalen RVEs erweitert werden. Die Kernidee besteht darin, die parasitär eingebrachten linearen Momentenverläufe durch Formulierung geeigneter Nebenbedingungen zu vermeiden. Exemplarisch soll das Prinzip an der Streckenschnittgröße m_x verdeutlicht werden. Mit

$$m_x = \int_{-h/2}^{+h/2} \sigma_x z \, dz \quad (5.15)$$

lautet die zugehörige Nebenbedingung

$$g := \int_{-l_x/2}^{l_x/2} \int_{-l_y/2}^{l_y/2} m_x \kappa_x \, dx \, dy \stackrel{!}{=} 0. \quad (5.16)$$

Knoten I auf Fläche	Bedingung ($\alpha = x, y$)
$x = \pm \ell_x/2$	$u_\alpha(\ell_x/2, y, z) = u_\alpha(-\ell_x/2, y, z)$ $u_z(\ell_x/2, y, z) = u_z(-\ell_x/2, -y, z)$
$y = \pm \ell_y/2$	$u_\alpha(x, \ell_y/2, z) = u_\alpha(x, -\ell_y/2, z)$ $u_z(x, \ell_y/2, z) = u_z(-x, -\ell_y/2, z)$

Tabelle 5.1: Kopplungsbedingungen auf dem RVE nach GRUTTMANN und WAGNER [46].

Für kleine Verformungen ist die Annahme einer linearen Krümmungsverteilung $\kappa_x = \lambda_1 x$ zulässig. λ_1 ist dabei ein LAGRANGE-Parameter, der als zusätzlicher Freiheitsgrad eingeführt wird. Da die Nebenbedingung (5.16) für lineare Momentenverteilung nicht erfüllt wird, werden diese damit ausgeschlossen. Analog lässt sich die Nebenbedingung für m_y formulieren. Im Falle von Kopplungen zwischen Membran- und Biegeverformungen müssen zusätzliche Forderungen an die Streckenschnittgrößen n_x und n_y gestellt werden. Damit werden nicht nur die Längenabhängigkeit der Querschubsteifigkeiten eliminiert, sondern gleichzeitig auch Starrkörperrotationen um die x - und y -Achse. Letztlich sind weitere Nebenbedingungen für die Vermeidung von Starrkörperverschiebungen in die z -Richtung notwendig. Im allgemeinen Fall müssen somit neun Nebenbedingungen berücksichtigt werden, die sich bei Entkopplung von Membran- und Biegeverformungen auf fünf reduzieren. Dies führt entsprechend auf die benötigte Anzahl von Lagrange-Parametern die im Vektor

$$\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_9]^T \quad (5.17)$$

zusammengefasst werden. Diese sind unabhängig von den x - und y -Koordinaten des RVEs, womit sie aus dem Integranden ausgeklammert werden können. Die Nebenbedingungen selbst werden im Vektor

$$\mathbf{g} = [g_1, \dots, g_9]^T \quad (5.18)$$

zusammengefasst.

Aus Gleichung (5.16) ist durch die arbeitskonforme Paarung der Charakter eines Energieterms zu erkennen, womit sich diese auch physikalisch interpretieren lassen. Entsprechend gehen die Variationen $\delta\boldsymbol{\lambda}$ und $\delta\mathbf{g}$ zusätzlich in die mikroskopische virtuelle Arbeit ein

$$\delta\Pi^m(\mathbf{u}^m, \delta\mathbf{u}^m) = \int_{\mathcal{B}^m} [\delta\mathbf{E}^T \mathbf{S} + \delta\boldsymbol{\lambda}^T \mathbf{g} + \boldsymbol{\lambda}^T \delta\mathbf{g}] \, dV. \quad (5.19)$$

Es lässt sich zeigen, dass die Erweiterung der mikroskopischen virtuellen Arbeit weiterhin konform mit der Hill-Mandel-Bedingung ist. Die numerische Behandlung der Nebenbedingung überschreitet den Rahmen dieses Kapitels und wird im Folgenden nicht weiter erläutert. Hierzu wird auf die bereits zuvor genannte Literatur verwiesen.

5.3 Numerische Homogenisierung

Mit der Definition der Randbedingungen auf dem RVE werden im Weiteren durch eine konsistente Kopplung zur Makroskala die homogenisierten Schalenschnittgrößen $\boldsymbol{\sigma}(\mathbf{X})$ und die effektive Materialmatrix $\mathbf{D}(\mathbf{X})$ aus Gleichung (2.94) berechnet.

Die Herleitung erfolgt auf Basis der Linearisierung der HILL-MANDEL-Bedingung, woraus die entsprechenden Größen direkt abgeleitet werden können. Die numerische Behandlung der Nebenbedingungen wird an dieser Stelle nicht erläutert, die folgenden Herleitungen sind allerdings unabhängig von dieser gültig und können um die Nebenbedingungen erweitert werden.

Auf der Makroskala wird die Schale entsprechend den Ausführungen aus Kapitel 2.4 diskretisiert. An jedem GAUSS-Punkt wird ein RVE eingeführt, welches mit dreidimensionalen Verschiebungselementen gemäß den Definitionen aus Kapitel 2.3 diskretisiert wird. Ausgangspunkt für das gekoppelte Problem ist die konsistente Linearisierung der HILL-MANDEL-Bedingung (5.8). Mit den Definitionen

$$g(\boldsymbol{\varepsilon}) := \delta \boldsymbol{\varepsilon}^T \boldsymbol{\sigma} \quad \text{und} \quad f(\mathbf{E}) := \frac{1}{A_0} \int_{\mathcal{B}^m} \delta \mathbf{E}^T \mathbf{S} \, dV \quad (5.20)$$

lässt sich diese auch schreiben als

$$G(\boldsymbol{\varepsilon}, \mathbf{E}) := g(\boldsymbol{\varepsilon}) - f(\mathbf{E}) = 0. \quad (5.21)$$

Für das totale Differential der Funktion (5.21) folgt

$$DG(\boldsymbol{\varepsilon}, \mathbf{E}) = \frac{\partial G}{\partial \boldsymbol{\varepsilon}} \Delta \boldsymbol{\varepsilon} + \frac{\partial G}{\partial \mathbf{E}} \Delta \mathbf{E} \quad (5.22)$$

mit den jeweiligen Anteilen

$$\frac{\partial G}{\partial \boldsymbol{\varepsilon}} \Delta \boldsymbol{\varepsilon} = \delta \boldsymbol{\varepsilon}^T \mathbf{D} \Delta \boldsymbol{\varepsilon} \quad \text{und} \quad \frac{\partial G}{\partial \mathbf{E}} \Delta \mathbf{E} = \frac{1}{A_0} \int_{\mathcal{B}^m} [\delta \mathbf{E}^T \mathbf{C}_T \Delta \mathbf{E} + \Delta \delta \mathbf{E}^T \mathbf{S}] \, dV. \quad (5.23)$$

Damit lautet die vollständige Linearisierung

$$\begin{aligned} L[G(\boldsymbol{\varepsilon}, \mathbf{E})] &= G(\boldsymbol{\varepsilon}, \mathbf{E}) + DG(\boldsymbol{\varepsilon}, \mathbf{E}) \\ &= \delta \boldsymbol{\varepsilon}^T (\mathbf{D} \Delta \boldsymbol{\varepsilon} + \boldsymbol{\sigma}) - \frac{1}{A_0} \int_{\mathcal{B}^m} [\delta \mathbf{E}^T \mathbf{C}_T \Delta \mathbf{E} + \Delta \delta \mathbf{E}^T \mathbf{S} + \delta \mathbf{E}^T \mathbf{S}] \, dV. \end{aligned} \quad (5.24)$$

Die Anordnung der Anteile in der Form

$$\delta \boldsymbol{\varepsilon}^T (\mathbf{D} \Delta \boldsymbol{\varepsilon} + \boldsymbol{\sigma}) = \frac{1}{A_0} \int_{\mathcal{B}^m} [\delta \mathbf{E}^T \mathbf{C}_T \Delta \mathbf{E} + \Delta \delta \mathbf{E}^T \mathbf{S} + \delta \mathbf{E}^T \mathbf{S}] \, dV. \quad (5.25)$$

dient der späteren Identifikation der makroskopischen Größen mittels Koeffizientenvergleichs. Für die linke Seite der Gleichung (5.25) ist keine Diskretisierung notwendig, da diese bereits den diskreten Größen am makroskopischen GAUSS-Punkt i der Schale, mit $i = 1, \dots, ngp$, entsprechen. Der mikroskopische Anteil der

rechten Seite ist allerdings noch kontinuierlich. Das RVE wird folglich in diskrete Elemente

$$\mathcal{B}^m \approx \mathcal{B}^h = \bigcup_{e=1}^{numel} \mathcal{B}^e \quad (5.26)$$

unterteilt. Durch das Einsetzen der Ansätze für die linearisierten Verzerrungen (2.61) und virtuellen linearisierten Verzerrungen (2.62) wird damit die diskrete Darstellung - vgl. Gleichung (2.66) - in den gesuchten Verschiebungsgrößen erhalten:

$$\delta \boldsymbol{\varepsilon}_i^T (\mathbf{D}_i \Delta \boldsymbol{\varepsilon}_i + \boldsymbol{\sigma}_i) = \bigcup_{e=1}^{numel} \delta \mathbf{u}_e^T [\mathbf{K}_{Te} \Delta \mathbf{u}_e + \mathbf{R}_e]. \quad (5.27)$$

Im nächsten Schritt werden die Randbedingungen des Mikroproblems berücksichtigt. Dafür wird der Verschiebungsvektor auf Elementebene in die Anteile aus den makroskopischen Verzerrungen am GAUSS-Punkt $\boldsymbol{\varepsilon}_i$ und den Vektor der allgemeinen Verschiebungsfreiheitsgrade aufgeteilt

$$\mathbf{u}_e = \mathbf{a}_e \mathbf{u} + \mathbf{A}_e \boldsymbol{\varepsilon}_i. \quad (5.28)$$

Dabei ist $\mathbf{a}_e$ eine allgemeine Assemblierungsmatrix und $\mathbf{A}_e$ ist in GRUTTMANN und WAGNER [46] spezifiziert. Letztere ist zuständig für die Einforderung der Rand- und Zwangsbedingungen, womit der Anteil $\mathbf{A}_e \boldsymbol{\varepsilon}_i$ nur bei Elementen auf dem Rand des RVEs auftritt. Die additive Zerlegung der Freiheitsgrade erfolgt analog für die Vektoren der virtuellen und linearisierten Verschiebungsgrößen

$$\delta \mathbf{u}_e = [\mathbf{a}_e, \mathbf{A}_e] \begin{bmatrix} \delta \mathbf{u} \\ \delta \boldsymbol{\varepsilon}_i \end{bmatrix} \quad \text{und} \quad \Delta \mathbf{u}_e = [\mathbf{a}_e, \mathbf{A}_e] \begin{bmatrix} \Delta \mathbf{u} \\ \Delta \boldsymbol{\varepsilon}_i \end{bmatrix}. \quad (5.29)$$

Das Einsetzen der Zusammenhänge (5.29) in die rechte Seite von Gleichung (5.27) liefert die Beziehung

$$\begin{aligned} \delta \boldsymbol{\varepsilon}_i^T (\mathbf{D}_i \Delta \boldsymbol{\varepsilon}_i + \boldsymbol{\sigma}_i) &= \frac{1}{A_0} \bigcup_{e=1}^{numel} \begin{bmatrix} \delta \mathbf{u} \\ \delta \boldsymbol{\varepsilon}_i \end{bmatrix}^T \left\{ \begin{bmatrix} \mathbf{a}_e^T \mathbf{K}_{Te} \mathbf{a}_e & \mathbf{a}_e^T \mathbf{K}_{Te} \mathbf{A}_e \\ \mathbf{A}_e^T \mathbf{K}_{Te} \mathbf{a}_e & \mathbf{A}_e^T \mathbf{K}_{Te} \mathbf{A}_e \end{bmatrix} \begin{bmatrix} \Delta \mathbf{u} \\ \Delta \boldsymbol{\varepsilon}_i \end{bmatrix} \right. \\ &\quad \left. + \begin{bmatrix} \mathbf{a}_e^T \mathbf{R}_e \\ \mathbf{A}_e^T \mathbf{R}_e \end{bmatrix} \right\} \\ &= \frac{1}{A_0} \begin{bmatrix} \delta \mathbf{u} \\ \delta \boldsymbol{\varepsilon}_i \end{bmatrix}^T \left\{ \begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{21} & \mathbf{K}_{22} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{u} \\ \Delta \boldsymbol{\varepsilon}_i \end{bmatrix} + \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{bmatrix} \right\} \end{aligned} \quad (5.30)$$

mit den modifizierten assemblierten Steifigkeitsmatrizen

$$\begin{aligned} \mathbf{K}_{11} &= \bigcup_{e=1}^{numel} \mathbf{a}_e^T \mathbf{k}_e \mathbf{a}_e, \quad \mathbf{K}_{12} = \bigcup_{e=1}^{numel} \mathbf{a}_e^T \mathbf{k}_e \mathbf{A}_e, \\ \mathbf{K}_{21} &= \mathbf{K}_{12}^T \quad \text{und} \quad \mathbf{K}_{22} = \bigcup_{e=1}^{numel} \mathbf{A}_e^T \mathbf{k}_e \mathbf{A}_e \end{aligned} \quad (5.31)$$

sowie den modifizierten assemblierten Residualvektoren

$$\mathbf{R}_1 = \bigcup_{e=1}^{numel} \mathbf{a}_e^T \mathbf{R}_e \quad \text{und} \quad \mathbf{R}_2 = \bigcup_{e=1}^{numel} \mathbf{A}_e^T \mathbf{R}_e. \quad (5.32)$$

Da die makroskopischen Verzerrungsinkremente $\Delta \boldsymbol{\varepsilon}_i$ bekannt sind, können die gesuchten mikroskopischen Verschiebungsinkremente $\Delta \mathbf{u}_i$ auf RVE-Ebene durch statische Kondensation aus dem Gleichungssystem eliminiert werden:

$$\Delta \mathbf{u}_i = -\mathbf{K}_{11}^{-1} (\mathbf{R}_1 + \mathbf{K}_{12} \Delta \boldsymbol{\varepsilon}_i). \quad (5.33)$$

Durch Einsetzen in Gleichung (5.30) folgt die Darstellung

$$\delta \boldsymbol{\varepsilon}^T (\mathbf{D}_i \Delta \boldsymbol{\varepsilon}_i + \boldsymbol{\sigma}_i) = \frac{1}{A_0} \left[(\mathbf{K}_{22} - \mathbf{K}_{21} \mathbf{K}_{11}^{-1} \mathbf{K}_{12}) \Delta \boldsymbol{\varepsilon}_i + \mathbf{R}_2 - \mathbf{K}_{21} \mathbf{K}_{11}^{-1} \mathbf{R}_1 \right]. \quad (5.34)$$

Aus dieser Umformung lassen sich mittels Koeffizientenvergleichs die homogenisierten Schalenschnittgrößen

$$\boldsymbol{\sigma}_i = \frac{1}{A_0} (\mathbf{R}_2 - \mathbf{K}_{21} \mathbf{K}_{11}^{-1} \mathbf{R}_1) \quad (5.35)$$

und die effektive Materialmatrix am GAUSS-Punkt

$$\mathbf{D}_i = \frac{1}{A_0} (\mathbf{K}_{22} - \mathbf{K}_{21} \mathbf{K}_{11}^{-1} \mathbf{K}_{12}) \quad (5.36)$$

extrahieren. Damit sind alle notwendigen Beziehungen zur Lösung des gekoppelten Randwertproblems hergeleitet worden. An jedem Integrationspunkt der Schale ist mit den vorgegebenen Verzerrungen eine FE-Berechnung des Mikroproblems durchzuführen, womit anschließend das übergeordnete Randwertproblem gelöst werden kann.

In Anlehnung an GRUTTMANN UND WAGNER [46] wird dabei zunächst von der diskreten linearisierten Form des gekoppelten Problems

$$L[\delta \pi(\boldsymbol{\theta}, \delta \boldsymbol{\theta})]_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}} = \bigcup_{e=1}^{numel} \left[\begin{array}{c} \delta \mathbf{v} \\ \delta \mathbf{u}_1 \\ \vdots \\ \delta \mathbf{u}_i \\ \vdots \\ \delta \mathbf{u}_{ngp} \end{array} \right]_e^T \left(\left[\begin{array}{c|cccc} \mathbf{k}_T & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{K}_1 & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{K}_i & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{K}_{ngp} \end{array} \right] \left[\begin{array}{c} \Delta \mathbf{v} \\ \Delta \mathbf{u}_1 \\ \vdots \\ \Delta \mathbf{u}_i \\ \vdots \\ \Delta \mathbf{u}_{ngp} \end{array} \right] + \left[\begin{array}{c} \mathbf{f} \\ \mathbf{F}_1 \\ \vdots \\ \mathbf{F}_i \\ \vdots \\ \mathbf{F}_{ngp} \end{array} \right] \right)_e. \quad (5.37)$$

ausgegangen. Mit den zuvor erfolgten Herleitungen lässt sich das Gleichungssystem so umschreiben, dass sich die direkte Abhängigkeit der Makroebene von den homogenisierten Schalenschnittgrößen σ_i und der effektive Materialmatrix $\mathbf{D}_i$ der Mikroebene und die Abhängigkeit der Mikroebene von den makroskopischen Schalenverzerrungen ε_i erkennen lässt:

$$L[\delta\pi(\boldsymbol{\theta}, \delta\boldsymbol{\theta})]_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}} = \bigcup_{e=1}^{numel} \left[\begin{array}{c} \delta\mathbf{v} \\ \delta\boldsymbol{\varepsilon}_1 \\ \vdots \\ \delta\boldsymbol{\varepsilon}_i \\ \vdots \\ \delta\boldsymbol{\varepsilon}_{ngp} \end{array} \right]^T_e \left\{ \left[\begin{array}{c|ccccc} \mathbf{k}_T(\mathbf{D}_i) & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{D}_1 & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D}_i & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{D}_{ngp} \end{array} \right] \left[\begin{array}{c} \Delta\mathbf{v} \\ \Delta\boldsymbol{\varepsilon}_1 \\ \vdots \\ \Delta\boldsymbol{\varepsilon}_i \\ \vdots \\ \Delta\boldsymbol{\varepsilon}_{ngp} \end{array} \right] + \left[\begin{array}{c} \mathbf{f}(\boldsymbol{\sigma}_i) \\ \boldsymbol{\sigma}_1 \\ \vdots \\ \boldsymbol{\sigma}_i \\ \vdots \\ \boldsymbol{\sigma}_{ngp} \end{array} \right] \right\}_e. \quad (5.38)$$

Bei geometrischer und oder materieller Nichtlinearität ist die entsprechende Gleichgewichtssiteration auf beiden Skalen notwendig. Diese Iteration wird bevorzugt simultan durchgeführt, d. h., es erfolgt eine gleichzeitige Aktualisierung der gesuchten Größen auf beiden Skalen. Dies ist aufgrund des reduzierten Berechnungsaufwands einer geschachtelten (engl.: *nested*) Iteration vorzuziehen, bei der das makroskopische Gleichungssystem erst nach Erreichen des Gleichgewichts auf der Mikroebene aktualisiert wird. Die unterschiedlichen Lösungsstrategien werden in GRUTTMANN und WAGNER [45] gegenübergestellt und weiterhin gezeigt, dass bei gleichzeitiger Aktualisierung die quadratische Konvergenz des Newton-Raphson-Verfahrens erhalten bleibt. Aus der Darstellung (5.38) ist zudem ersichtlich, dass die Mikroprobleme entkoppelt sind. Dahingehend werden diese im Rahmen der Implementierung parallel gelöst. Falls auf der Mikroskala geschichtsabhängiges Materialverhalten vorliegt, müssen weiterhin die entsprechenden Zustandsvariablen der Schalen-RVEs entsprechend abgespeichert werden, um für die Berechnung im Rahmen des nächsten Last- oder Zeitschritts zur Verfügung zu stehen.

Kommentar zur FE²-Methode für Schalen

Dieser Abschnitt dient der Einordnung des Potentials des Mehrskalenansatzes für Schalen im Hinblick auf den Berechnungsaufwand. Gleichzeitig soll damit der in den folgenden Kapiteln verfolgte Ansatz über ein Ersatzmodellierung mit einem künstlichen neuronalen Netz motiviert werden. Es werden drei Fälle bezüglich des Materialverhaltens der Mikrostruktur diskutiert, wobei die genannten Punkte für die Annahme geometrischer Linearität auf der Mikroebene, jedoch geometrischer Nichtlinearität auf der Makroebene, unter Annahme kleiner Verzerrungen, gelten.

- **Linear elastisches Materialverhalten**

Bei linear elastischen Strukturen mit komplexer Mikrostruktur kann der FE²-Ansatz besonders effizient genutzt werden. Das Mikroproblem muss nur einmalig gelöst werden; anschließend kann die effektive Materialmatrix gespeichert und im makroskopischen Modell verwendet werden. Die homogenisierten Spannungen lassen sich an jedem GAUSS-Punkt analog zu Gleichung (2.94) berechnen:

$$\sigma_i = \mathbf{D}_i \varepsilon_i \quad \text{und} \quad \mathbf{D}_i = \text{konst.}, \quad i = 1, \dots, ngp. \quad (5.39)$$

- **Linear inelastisches Materialverhalten**

Für Materialien mit geschichtsabhängigem Verhalten ist eine Lösung des Mikroproblems in jedem Last-, bzw., (Makro) Iterationsschritt erforderlich. Aufgrund der Linearität entfällt jedoch eine lokale Iteration auf Mikroebene. Der Rechenaufwand hängt hierbei von der Komplexität der Mikrostruktur und der Summe der makroskopischen GAUSS-Punkte ab. Wie in Kapitel 8 gezeigt wird, kann dieser Ansatz – je nach Struktur – dennoch effizienter sein als ein vollständig dreidimensionales Modell. Um die wiederholte Lösung des Mikroproblems zu vermeiden, bietet sich jedoch bereits in diesem Fall der Einsatz eines Ersatzmodells, z. B. auf Basis eines künstlichen neuronalen Netzes, an. Dieser Aspekt bildet einen zentralen Fokus der vorliegenden Arbeit, da auf der Mikroebene viskoelastisches Materialverhalten berücksichtigt wird.

- **Nichtlinear inelastisches Materialverhalten**

Zusätzlich zur Lösung des Mikroproblems in jedem Lastschritt ist hier auch eine lokale Gleichgewichtsiteration auf Mikroebene erforderlich. Ein typisches Beispiel ist elastoplastisches Materialverhalten. In solchen Fällen ist der Rechenaufwand besonders hoch, sodass der Einsatz eines Ersatzmodells auch hier vorteilhaft erscheint.

6 Künstliche neuronale Netze

Künstliche Intelligenz (KI) – ein Begriff, der seinesgleichen sucht, wenn man die rasante Entwicklung vergleicht, mit der dieser in kürzester Zeit Einzug in alle Bereiche unserer Gesellschaft erhalten hat. Spätestens seit der Einführung unterschiedlicher KI-Assistenten (Alexa, Google Gemini, ChatGPT, Microsoft Copilot,...) hat sich nicht nur der Begriff, sondern auch die Technologie fest in unserem Alltag etabliert. KI kann augenscheinlich alles – Übersetzen, Texte verfassen, Recherchieren, Bilder zeichnen, ganze Songs kreieren, Daten analysieren – womit viele Aufgaben vermehrt mit ihrer Unterstützung gelöst werden.

Der grundsätzliche Gedanke hinter *künstlicher Intelligenz* ist auf Basis der Funktionsweise des menschlichen Gehirns adaptive Algorithmen zu entwickeln, um Aufgaben zu erledigen, die für einen Computer vergleichsweise einfach sind. In vielen Bereichen ist die KI dem Menschen dabei schon überlegen und das nicht erst seit den Entwicklungen der letzten Jahre. Bereits im Jahr 1997 unterlag der damalige Schachweltmeister GARRI KASPAROV dem Schachprogramm *Deep Blue* von IBM (vgl. HSU [58]). Für die heutigen Programme wie *AlphaZero* und *Stockfish* ist der Mensch ein leichter Gegner und die komplexen Duelle der Programme gegeneinander nur schwer nachvollziehbar. Schach ist dabei nicht ohne Grund prädestiniert für den Einsatz von KI: Ein einfaches Spielfeld und Figuren mit eingeschränkten Bewegungsmöglichkeiten führen dazu, dass auch der Mensch sehr schnell die Formalismen versteht. Er braucht jedoch Jahre an Übung, um sie zu beherrschen und erfolgreich zu spielen, was für Schachprogramme innerhalb von wenigen Stunden möglich ist.

KI lässt sich zunächst allgemein in *wissensbasierte Systeme* und *maschinelles Lernen* unterteilen (vgl. GOODFELLOW ET AL. [42]). Bei wissensbasierten Ansätzen werden logische Zusammenhänge explizit vorgegeben, sodass die KI ausschließlich auf Basis dieser programmierten Regeln eigene Schlüsse ziehen kann. *Maschinelles Lernen* hingegen verfolgt den deutlich erfolgreicherer Ansatz: Ein Datensatz wird bereitgestellt, aus dem das System selbständig die relevanten Merkmale zur Entscheidungsfindung extrahiert. Die so gewonnenen Merkmale unterscheiden sich dabei häufig von jenen, die ein Mensch für relevant erachten würde. Um einer KI die nötige Flexibilität zu geben, solche Merkmale eigenständig zu erkennen, wird sie aus einer Vielzahl an einfachen Verarbeitungseinheiten²⁴ aufgebaut. Mit zunehmender Komplexität der zu erlernenden Aufgabe steigt die Anzahl der benötigten Einheiten, was auf den Begriff des *Deep Learnings* führt. Werden die Einheiten wiederum sequenziell in Schichten angeordnet, findet man sich bei

²⁴Der Begriff Verarbeitungseinheit ist formal bei der Datenverarbeitung von Computern für die Prozessoren (CPUs) oder Grafikprozessoren (GPUs) vorgesehen. Im weiteren Verlauf des Kapitels ist wird die Analogie zu diesen jedoch offensichtlich.

der Definition einer konkreten KI-Architektur wieder, dem *Multilayer-Perceptron* (MLP), welches auch im Rahmen dieser Arbeit Verwendung findet.

Historisch gesehen ist die Entwicklung von Algorithmen, die in ihrem Aufbau und ihrer Funktionsweise, dem menschlichen Gehirn ähneln, nicht erst auf die letzten Jahrzehnte zurückzuführen. KI bezeichnet dabei die gewünschte Charaktereigenschaft des Algorithmus – die abstrakte Fähigkeit Probleme zu lösen – womit sich der Begriff *künstliches neuronales Netz* (KNN) als Überbegriff verschiedener Architekturen etabliert hat. Die informationsverarbeitenden Einheiten entsprechen dabei den Neuronen im Gehirn und geben über gewichtete Verbindungen – analog der biologischen Synapsen – Datenströme an die Neuronen folgender Schichten weiter. Der erste Ansatz, diese Prozesse mit mathematischen Modellen zu beschreiben, geht auf die Arbeit von MCCULLOCH und PITTS [92] aus dem Jahr 1943 zurück, die darin erfolgreich zeigen konnten, dass logische und arithmetische Funktionen berechnet werden können. In Addition dazu konkretisiert die HEBB'sche Lernregel [52] das Zusammenspiel angrenzender Neurone dahingehend, dass der Informationsfluss zwischen ihnen proportional zur Aktivität ist. Dieser Grundsatz bildet weiterhin die Basis vieler maschineller Lernverfahren. Der Begriff des *Perceptrons* wurde 1958 von ROSENBLATT [112] geprägt und lässt sich bis heute noch als grundlegender Baustein vieler KI-Architekturen identifizieren (vgl. *Multi-Layer-Perceptron*). Weitere Arbeiten in dem Feld folgten, bevor mit der Erstausgabe (1969) von *Perceptrons* der Autoren MINSKY und PAPERT [97] erste Limitationen der Lösungsfähigkeit der Modelle aufgezeigt wurden²⁵. Das „Aufblühen“ in den 1980er-Jahren wird überwiegend mit der Einführung des HOPFIELD-Netzes [57] und der Anwendung des Gradientenverfahrens für den Lernprozess – auch bekannt als *Fehlerrückführung* oder *Backpropagation of Error* – durch RUMELHART ET AL. [115] assoziiert. Eine ausführlichere Darstellung weiterer historischer Meilensteine findet sich bei KRIESEL [73]. Die Vielzahl an Veröffentlichungen und deren Anwendung in unterschiedlichen Disziplinen lässt an dieser Stelle keine komprimierte Darstellung der maßgeblichen Fortschritte der letzten Jahrzehnte zu.

Entsprechend haben KNN auch Einzug in den Bereich der numerischen Mechanik erhalten (s. HERRMANN und KOLLMANNSSBERGER [53]). Als Beispiel hierfür wurde in Kapitel 2.3 bereits die Anwendung zur Lösung partieller Differentialgleichungen in Form von *Physics-Informed Neural Networks* genannt. Ohne Kapitel 7 vorwegzugreifen, ist eine weitere Anwendung nach den Ausführungen in Kapitel 3 naheliegend: die Materialmodellierung. Der entsprechende Überblick

²⁵In der Literatur wird häufig davon gesprochen, dass die Veröffentlichung maßgeblich für den Stillstand der wissenschaftlichen Fortschritte in den folgenden Jahren zurückzuführen ist. In der Neuauflage [97] widersprechen die Autoren im Prolog dieser Ansicht, da aus ihrer Sicht durch fehlende theoretische Grundlage der Stillstand davor bereits eingeleitet wurde.

der Entwicklungen hierzu wurde bereits in Kapitel 1 gegeben. Die theoretische Symbiose der zwei grundlegenden Voraussetzungen zur Konstruktion eines konstitutiven Zusammenhangs ist nicht von der Hand zu weisen: Die experimentellen Materialversuche liefern den Datensatz und das KNN das Modell. Ersteres stellt dabei, wie in Kapitel 3.4 beschrieben, oftmals die einschränkende Problematik dar. Entsprechend wünschenswert ist ein unbegrenzter Datensatz, was wiederum die *Homogenisierung* ins Spiel bringt. Numerisch lassen sich – wie in Kapitel 5.3 ausgeführt – theoretisch unendliche Datenmengen in Form von synthetischen Materialversuchen auf dem RVE erzeugen. Letztlich verbleibt nur die Konstruktion eines geeigneten KNNs zur Beschreibung des konstitutiven Zusammenhangs.

Für den Einsatz eines Materialmodells in der beschriebenen Form werden zunächst die Definitionen und Berechnungsvorschriften eines KNNs – spezifisch eines *Feedforward*-Netzes – eingeführt. Die Aufbereitung des Datensatzes, die Auswertung des Netzes sowie die Festlegung eines geeigneten Fehlermaßes zur Quantifizierung des Lernerfolgs stellen die Voraussetzung zur Anwendung der Optimierungsalgorithmen im Rahmen des Lernprozesses dar. Abschließend wird die Einbindung von Nebenbedingungen diskutiert, die einen wichtigen Faktor bei der Anwendung für die Materialmodellierung darstellen werden. Alle hergeleiteten Zusammenhänge wurden in einer eigenen Programmumgebung in MATLAB [131] implementiert und es wurden keine der bereits vorhandenen Funktionen des *Machine Learning und Deep Learning* Pakets verwendet, was zu einer äußerst flexiblen und eigenständigen Programmstruktur beitrug.

6.1 Das Feedforward-Netz

Ein *Feedforward*-Netz (FFN) ist ein vollverknüpftes KNN, dessen Informationsfluss ausschließlich unidirektional von der Eingabeschicht zur Ausgabeschicht verläuft. Liegen zwischen Ein- und Ausgabeschicht weitere *verdeckte Schichten*, kann das *Feedforward*-Netz auch als MLP bezeichnet werden. Im Rahmen der Arbeit wird dahingehend keine explizite Differenzierung zwischen den Begriffen FFN, KNN und MLP vorgenommen.

Mathematisch lässt sich ein KNN als die Abbildung

$$f(\mathbf{x}; \mathbf{w}) \mapsto \mathbf{z} \quad (6.1)$$

darstellen. Bei den unabhängigen Variablen wird dabei zwischen dem Eingangsvektor $\mathbf{x} \in \mathbb{R}^{n_i}$ und den freien Parametern des Netzes $\mathbf{w} \in \mathbb{R}^{n_w}$ – den Gewichten – unterschieden. n_i bezeichnet dabei die Eingangsdimension und n_w ist die Anzahl aller Gewichte des KNN und damit die Anzahl der Freiheitsgrade. Weiterhin wird $\mathcal{W} = \mathbb{R}^{n_w}$ als die Menge aller möglichen Gewichtskombinationen definiert. Neben

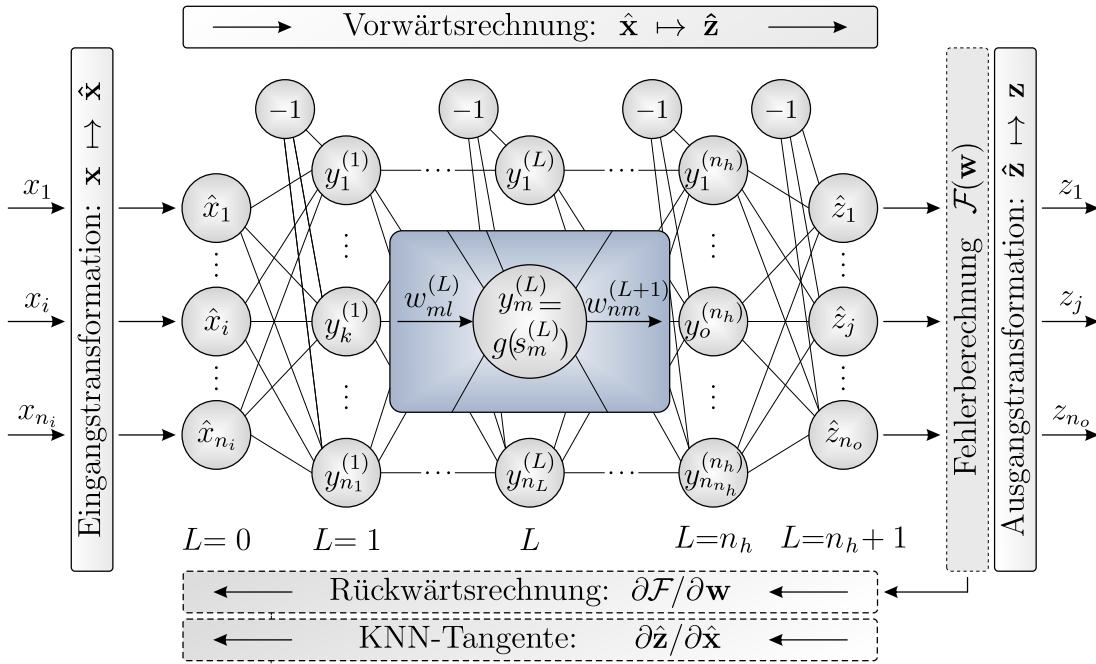


Abbildung 6.1: Darstellung eines *Feedforward*-Netzes und Detailansicht eines Neurons sowie der zugehörigen Berechnungsschritte. Die gestrichelt dargestellten Operationen sind optional und hier nur im Rahmen des KNN-Trainings oder für den Einsatz als Materialmodell notwendig.

diesen existieren noch sogenannte Hyperparameter; Variablen, deren Werte vom Anwender fest definiert und nicht im Verlauf des Lernprozesses angepasst werden. Der Vektor der Ausgangsgrößen wird mit $\mathbf{z} \in \mathbb{R}^{n_o}$, mit der Ausgangsdimension n_o , bezeichnet. In dieser Arbeit wird das Netz ausschließlich zur Funktionsapproximation genutzt²⁶. Ziel ist es damit, eine Referenz-Funktion $\tilde{f}(\mathbf{x})$ durch Bestimmung eines geeigneten Satzes an Gewichten $\tilde{\mathbf{w}}$ aus der Menge $\mathcal{W}$ zu approximieren, sodass gilt

$$\mathbf{z}(\mathbf{x}; \tilde{\mathbf{w}}) \approx \tilde{f}(\mathbf{x}). \quad (6.2)$$

In der Regel liegt keine kontinuierliche Beschreibung der Funktion $\tilde{f}(\mathbf{x})$ vor, sondern nur ein diskreter Datensatz

$$\mathcal{D} = \{(\mathbf{x}_k, \mathbf{t}_k) \mid \mathbf{x}_k \in \mathcal{X} \subset \mathbb{R}^{n_i}, \mathbf{t}_k = \tilde{f}(\mathbf{x}_k), k = 1, 2, \dots, P\}. \quad (6.3)$$

$\mathcal{X}$ ist dabei der den Eingangsdaten zugrundeliegende physikalische Raum und P die maximale Anzahl an Datenpaaren. Der Vektor $\mathbf{t}_k$ enthält die Zielgrößen, die als Referenzlösungen dienen. Dieser Datensatz stellt die Basis für den Lernprozess

²⁶Üblicherweise werden zwei Hauptgebiete für den Einsatz von KNN unterschieden: die Anwendung für Klassifizierungsprobleme (Mustererkennung) und für die Funktionsapproximation (Regression). Für Details, s. auch [73].

– im Weiteren auch Training genannt – dar. Im Folgenden werden die für das KNN-Training benötigten Beziehungen und Algorithmen erläutert. Als Grundlage dient Abbildung 6.1, an der sich der Aufbau dieses Unterkapitels orientiert. Die Diskussion folgt direkt der im Rahmen der Implementierung vorzunehmenden Schritte.

6.1.1 Transformation der KNN-Daten

Der erste entscheidende Schritt für den erfolgreichen Aufbau eines KNNs ist die Vorverarbeitung des Datensatzes. Dies ist dahingehend entscheidend, da sowohl die Eingangs- als auch die Ausgangsvariablen unterschiedliche Größenordnungen annehmen können, insbesondere wenn es sich um physikalische Problemstellungen handelt. Die Transformation der Daten auf einen normalisierten Bereich garantiert damit zum einen eine gleichmäßige Wichtung aller Ein- und Ausgangsvariablen, was das Training erheblich stabilisiert. Zum anderen erlaubt dies sowohl eine standardisierte Initialisierung der Gewichte als auch eine von den Trainingsdaten unabhängige Wahl der Hyperparameter des Netzes (vgl. SILVA ET AL. [27]).

Der physikalische Raum $\mathcal{X}$ wird damit über eine lineare Abbildungsvorschrift auf einen normalisierten Raum $\hat{\mathcal{X}}$ mit abgebildet²⁷, sodass für die normalisierten Eingangsgrößen $\hat{\mathbf{x}} \in [-1, 1]^{n_i}$ gilt. Dieselbe Transformation wird für die normalisierten Zielgrößen $\hat{\mathbf{t}} \in [-1, 1]^{n_o}$ vorgenommen. Komponentenweise lässt sich dies mit der Vorschrift

$$\hat{x}_i = 2 \left(\frac{x_i - x_{i,\min}}{x_{i,\max} - x_{i,\min}} \right) - 1 \quad \text{mit} \quad i = 1, \dots, n_i \quad (6.4)$$

und

$$\hat{t}_j = 2 \left(\frac{t_j - t_{j,\min}}{t_{j,\max} - t_{j,\min}} \right) - 1 \quad \text{mit} \quad j = 1, \dots, n_o \quad (6.5)$$

umsetzen. Die Minimal- und Maximalwerte der Eingangsvariablen $x_{i,\min}$ und $x_{i,\max}$ lassen sich in den Vektoren $\mathbf{x}_{\min}, \mathbf{x}_{\max} \in \mathbb{R}^{n_i}$ zusammenfassen:

$$\mathbf{x}_{\min} = [x_{1,\min}, \dots, x_{n_i,\min}]^T \quad \text{und} \quad \mathbf{x}_{\max} = [x_{1,\max}, \dots, x_{n_i,\max}]^T. \quad (6.6)$$

In gleicher Weise erfolgt die Einordnung für die Werte der Zielgrößentransformation in $\mathbf{t}_{\min}, \mathbf{t}_{\max} \in \mathbb{R}^{n_o}$:

$$\mathbf{t}_{\min} = [t_{1,\min}, \dots, t_{n_o,\min}]^T \quad \text{und} \quad \mathbf{t}_{\max} = [t_{1,\max}, \dots, t_{n_o,\max}]^T. \quad (6.7)$$

²⁷Das Vorgehen ähnelt dem des in Kapitel 2.3.1 erläuterten Einführens eines Referenzelements. Während in beiden Fällen dadurch gewisse Standardisierungen möglich gemacht werden, hat dieses Konzept aber im Rahmen der FEM – im Gegensatz zum KNN-Training – keinen Einfluss auf die Robustheit des Verfahrens.

Entsprechend lässt sich die komponentenweise Darstellung der Transformation auch vektoriell formulieren:

$$\hat{\mathbf{x}} = 2 [(\mathbf{x} - \mathbf{x}_{\min}) \oslash (\mathbf{x}_{\max} - \mathbf{x}_{\min})] - \mathbf{1}_{n_i} \quad (6.8)$$

und

$$\hat{\mathbf{t}} = 2 [(\mathbf{t} - \mathbf{t}_{\min}) \oslash (\mathbf{t}_{\max} - \mathbf{t}_{\min})] - \mathbf{1}_{n_o}. \quad (6.9)$$

Dabei ist $\mathbf{1}_n = [1, \dots, 1]^T \in \mathbb{R}^n$ ein n -dimensionaler Spaltenvektor und $\oslash$ bezeichnet eine elementweise durchzuführende Division. Die Gleichung (6.9) stellt die Basis für die Rücktransformation der normalisierten Ausgangsgrößen des KNN in den physikalischen Raum dar. Diese erfolgt über die Beziehung

$$\mathbf{z} = \frac{(\mathbf{t}_{\max} - \mathbf{t}_{\min}) \odot (\hat{\mathbf{z}} + \mathbf{1}_{n_o})}{2} + \mathbf{t}_{\min}, \quad (6.10)$$

wobei $\odot$ für eine elementweise Multiplikation verwendet wird. Weiterhin gilt die inverse Beziehung

$$\mathbf{x} = \frac{(\mathbf{x}_{\max} - \mathbf{x}_{\min}) \odot (\hat{\mathbf{x}} + \mathbf{1}_{n_i})}{2} + \mathbf{x}_{\min}. \quad (6.11)$$

Die Minimal- und Maximalwerte der Eingangs- und Zielgrößen sind damit essenzielle Transformationsparameter, ohne deren Kenntnis das KNN im Rahmen einer weiteren Verwendung – z. B. als Materialmodell – unbrauchbar ist.

6.1.2 Die Vorwärtsrechnung

Die Vorwärtsrechnung ist das „Nervensystem“ des KNNs und kann biologisch gesehen als Reaktion $\mathbf{z}$ auf einen Reiz $\mathbf{x}$ interpretiert werden. Mathematisch bedeutet dies die Abbildung

$$\hat{\mathbf{x}} \mapsto \hat{\mathbf{z}}, \quad (6.12)$$

die hier direkt in den normalisierten Größen erfolgt. Hierfür ist es notwendig einige Definitionen zur Beschreibung der KNN-Architektur (vgl. Abbildung 6.1) einzuführen. Graphisch ausgedrückt erfolgt die Vorwärtsrechnung des KNNs durch Weitergabe der Eingangsvariablen $\hat{\mathbf{x}}$ in der Schicht $L = 0$ über eine Anzahl $n_h \in \mathbb{N}$ Zwischenschichten an die Ausgangsschicht $L = n_h + 1$. Jeder Schicht wird eine Anzahl n_L Neuronen zugewiesen, die für die Verarbeitung der Information der vorherigen Schicht zuständig sind. Für die Ein- und Ausgangsschicht sind mit $n_{L=0} \equiv n_i$ sowie $n_{L=n_h+1} \equiv n_o$ die Anzahl der Neuronen bereits durch die zu approximierende Problemstellung festgelegt. Sowohl die Anzahl der Zwischenschichten als auch die Anzahl der darin enthaltenen Neurone muss gewählt

werden. Für die Bezeichnung einer KNN-Architektur wird damit die Konvention $n_i - n_{n_L} - n_{n_o}$ verwendet.

Die Ausgangsgrößen der einzelnen Neuronen einer Schicht L können im Vektor

$$\mathbf{y}^{(L)} = [-1, \bar{\mathbf{y}}^{(L)T}]^T = [-1, y_1^{(L)}, \dots, y_m^{(L)}, \dots, y_{n_L}^{(L)}]^T \in \mathbb{R}^{n_L+1} \quad (6.13)$$

zusammengefasst werden. Dabei wurde mit $y_0^{(L)} = -1$ das sogenannte *Bias*-Neuron eingeführt, s. Abbildung 6.1. Dieses hat in Kombination mit dem ihm zugeordneten Gewicht die Funktion eines Schwellenwerts, auf dessen Bedeutung noch näher eingegangen wird. Für den Vektor der Ausgangsgrößen exklusive des *Biases* wird die Bezeichnung $\bar{\mathbf{y}}^{(L)}$ verwendet. Die Verbindung eines Neurons $y_l^{(L-1)}$ der Schicht $L-1$ mit dem Neuron $y_m^{(L)}$ einer nachfolgenden Schicht L erfolgt über das Gewicht $w_{ml}^{(L)}$. Damit gilt die Konvention des ersten Indizes für das Ziel- und des zweiten Indizes für das Startneuron. Die der Verbindung einer Schicht L mit der vorherigen Schicht von Neuronen $L-1$ zugeordneten Gewichte werden in der Matrix

$$\mathbf{w}^{(L)} = \begin{bmatrix} w_{10}^{(L)} & w_{11}^{(L)} & \dots & w_{1n_{L-1}}^{(L)} \\ w_{20}^{(L)} & w_{21}^{(L)} & \dots & w_{2n_{L-1}}^{(L)} \\ \vdots & \vdots & w_{ml}^{(L)} & \vdots \\ w_{n_L0}^{(L)} & w_{n_L1}^{(L)} & \dots & w_{n_Ln_{L-1}}^{(L)} \end{bmatrix} \in \mathbb{R}^{n_L \times (n_{L-1}+1)} \quad (6.14)$$

eingesortiert. Die *Bias*-Neurone sind dabei nicht mit der ihr vorangehenden Schicht verbunden, womit der erste Index der Gewichte immer mit 1 beginnt. Analog zu der bereits verwendeten Notation für den Vektor der Neuronenausgaben, bezeichnet $\bar{\mathbf{w}}^{(L)} \in \mathbb{R}^{n_L \times (n_{L-1})}$ die Gewichtsmatrix ohne die *Bias*-Gewichte. Der bereits eingeführte Vektor $\mathbf{w}$ aller Gewichte wird wie folgt aufgebaut:

$$\mathbf{w} = [w_{10}^{(1)}, \dots, w_{1n_i}^{(1)}, \dots, w_{ml}^{(L)}, \dots, w_{n_o0}^{(n_h+1)}, \dots, w_{n_on_{n_h}}^{(n_h+1)}]^T \in \mathbb{R}^{n_w}. \quad (6.15)$$

Die Initialisierung der Gewichte erfolgt über eine Gleichverteilung $\mathcal{U}(-1, 1)$ im Intervall von $[-1, 1]$ und anschließender Wichtung mit einem Faktor²⁸ $2,4/n_i$:

$$w_{ml}^{(L)} = \mathcal{U}(-1, 1) \cdot \frac{2,4}{n_i}. \quad (6.16)$$

Die Ausgänge der vorherigen Schicht multipliziert mit dem zugehörigen Gewicht wird als gewichtete Summe

$$s_m^{(L)} = \sum_{l=0}^{n_{L-1}} w_{ml}^{(L)} y_l^{(L-1)} \quad (6.17)$$

²⁸Die Skalierung führt empirisch zu guten Startwerten. Da der Eingang auf das Intervall $[-1, 1]$ skaliert wird, kann die o.g. Wichtung als Zuweisung eines Intervalls der Breite $2/n_i$ gesehen werden, womit der gewählte Faktor eine gewisse Überlappung der einzelnen Intervalle zulässt.

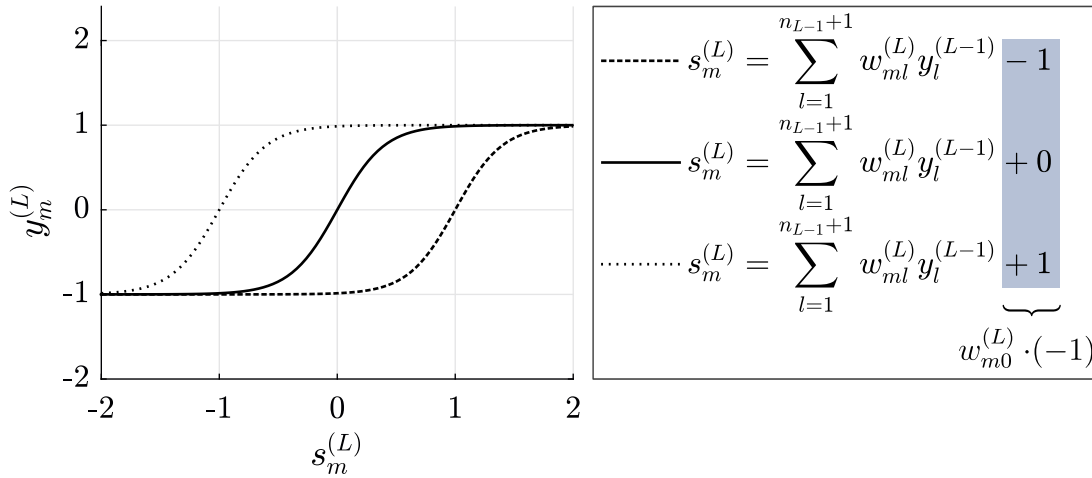


Abbildung 6.2: Darstellung des Tangens hyperbolicus für drei verschiedene Schwellenwerte, die die Aktivität des Neurons steuern

eines Neurons eingeführt. Im Falle der Identität $y_m^{(L)} = s_m^{(L)}$ wäre das KNN eine rein lineare Abbildung zwischen den Ein- und Ausgangsgrößen, was die Approximationsfähigkeit stark einschränken würde. Die Nichtlinearität wird dahingehend über eine zusätzliche Funktion – die Aktivierungsfunktion $g(s_m^{(L)})$ – in das Netz integriert:

$$y_m^{(L)} = g(s_m^{(L)}) . \quad (6.18)$$

Für die Aktivierungsfunktion steht eine große Bandbreite an geeigneten Funktionen zur Verfügung. Im Rahmen der Arbeit wird der Tangens hyperbolicus in der Darstellung

$$g(s) = \frac{1 - e^{-Ts}}{1 + e^{-Ts}} \quad (6.19)$$

verwendet. T ist dabei einer der zuvor genannten Hyperparameter, der hier zu $T = 5$ gesetzt wird. Dies führt dazu, dass sich die Bereiche großer Sensitivität immer innerhalb eines Intervalls von ungefähr $[(a - 1), (a + 1)]$ befinden. a ist dabei der zuvor genannte Schwellenwert, der über das Gewicht $w_{m0}^{(L)}$ eines *Bias*-Neurons während des KNN-Trainings angepasst wird. In Abbildung 6.2 ist der Einfluss für die Schwellenwerte $a \in \{-1, 0, 1\}$ auf den Verlauf der Aktivierungsfunktion dargestellt. Ohne das *Bias*-Neuron würde diese immer durch den Ursprung verlaufen. Die Aktivität oder Inaktivität des Ziel-Neurons kann damit über die Einführung des modifizierbaren Schwellenwerts durch das KNN im Trainingsprozess gesteuert werden.

Es wird angemerkt, dass in der Literatur (vgl. z.B. SILVA [27] oder GOODFELLOW ET AL. [42]) eine ganze Reihe weiterer Aktivierungsfunktionen genannt werden, die in Abhängigkeit der zu approximierenden Problemstellung von Vorteil sein

Eingangstransformation:	$\mathbf{x} \mapsto \hat{\mathbf{x}} \Rightarrow \mathbf{x}_{\min}, \mathbf{x}_{\max}$
	$\mathbf{z} \mapsto \hat{\mathbf{z}} \Rightarrow \mathbf{t}_{\min}, \mathbf{t}_{\max}$
Eingangsvektor:	$\mathbf{y}^{(0)} = \hat{\mathbf{x}}$
Schleife über alle Schichten:	for $L = 1, \dots, n_h$ $\mathbf{s}^{(L)} = \mathbf{w}^{(L)} \mathbf{y}^{(L-1)}$ $\bar{\mathbf{y}}^{(L)} = g(\mathbf{s}^{(L)})$ $\mathbf{y}^{(L)} = [-1; \bar{\mathbf{y}}^{(L)}]$ end
Ausgangsschicht:	$\mathbf{s}^{(n_h+1)} = \mathbf{w}^{(n_h+1)} \mathbf{y}^{(n_h)}$ $\hat{\mathbf{z}} = \bar{\mathbf{y}}^{(n_h+1)} = \mathbf{s}^{(n_h+1)}$
Ausgangstransformation:	$\hat{\mathbf{z}} \mapsto \mathbf{z}$

Algorithmus 6.1: Berechnungsvorschrift für die FFN-Vorwärtsrechnung

können²⁹. In der Ausgabeschicht wird keine nichtlineare Funktion zwischen geschaltet, hier gilt die Identität zwischen gewichteter Summe und Ausgang der Neurone. Damit folgt die Unterscheidung

$$y_m^{(L)} = \begin{cases} s_m^{(L)}, & L = n_h + 1 \\ g(s_m^{(L)}), & \text{sonst.} \end{cases} \quad (6.20)$$

Die Gleichungen (6.17) und (6.18) werden abschließend noch in Vektor-Matrix Notation gegeben, die sich im Rahmen der Implementierung der KNN-Vorwärtsrechnung anbietet:

$$\mathbf{s}^{(L)} = \mathbf{w}^{(L)} \mathbf{y}^{(L-1)} \quad \text{und} \quad (6.21)$$

$$\bar{\mathbf{y}}^{(L)} = g(\mathbf{s}^{(L)}). \quad (6.22)$$

Dabei wurde der Vektor der gewichteten Summen der Schicht L

$$\mathbf{s}^{(L)} = [s_1^{(L)}, \dots, s_m^{(L)}, \dots, s_{n_L}^{(L)}]^T \in \mathbb{R}^{n_L} \quad (6.23)$$

eingeführt. Die entsprechenden Schritte sind in Algorithmus 6.1 zusammengefasst. Angemerkt sei dazu, dass keine Schleife über die Datenpunkte k benötigt wird und die Vorwärtsrechnung für alle Datenpunkte im Rahmen der vorgenommenen Implementierung durch einmaliges Ausführen des Algorithmus möglich ist.

²⁹Durch gezielte Wahl der Aktivierungsfunktion können dem KNN auch a priori gewisse Eigenschaften zugesprochen werden. Exemplarisch genannt seien sogenannte *Input Convex Neural Networks* (vgl. AMOS [6]), bei denen die Aktivierungsfunktionen konvex und monoton wachsend sein müssen. Dies lässt sich beispielsweise mit der *SoftPlus* Funktion $g(s) = \ln(1 + e^s)$ realisieren.

6.1.3 Definition des Fehlers und Rückwärtsrechnung

Im vorherigen Abschnitt wurde das „Nervensystem“ als Analogie für die Vorwärtsrechnung verwendet. Entsprechend kann die „Neuroplastizität“ als biologische Metapher für die Rückwärtsrechnung interpretiert werden. Diese wird beim KNN für die Anpassung der Gewichte – also das Stärken und Schwächen der Synapsen – benötigt. Wie dieser Prozess konkret abläuft, wird in Kapitel 6.2 anhand verschiedener Optimierungsverfahren diskutiert. Für alle sind jedoch zwei Größen notwendig: ein Fehlermaß und dessen Gradient, der die Sensitivität bezüglich der Gewichte beschreibt.

Das Fehlermaß dient zur Quantifizierung der Abweichung eines Ausgangs des KNNs z_{kj} für eine Eingangsgröße $\mathbf{x}_k$ von der Zielgröße t_{kj} . Hierfür wird zunächst der physikalische Fehler eines einzelnen Ausgangsneurons definiert:

$$\bar{e}_{kj}(\mathbf{w}) := z_{kj}(\mathbf{x}_k, \mathbf{w}) - t_{kj}. \quad (6.24)$$

Der mittlere quadratische Fehler ist eine skalare Größe und quadriert die Fehlerterme $\bar{e}_{kj}$ zunächst, um diese anschließend über alle Ausgangsneurone und Trainingsdaten $k = 1, \dots, P$ zu summieren und über letztere zu mitteln:

$$\bar{\mathcal{F}}(\mathbf{w}) := \frac{1}{2P} \sum_{k=1}^P \sum_{j=1}^{n_o} (z_{kj}(\mathbf{x}_k, \mathbf{w}) - t_{kj})^2 = \frac{1}{2P} \sum_{k=1}^P \sum_{j=1}^{n_o} \bar{e}_{kj}(\mathbf{w})^2. \quad (6.25)$$

Mit der Einführung des Fehlervektors für einen Datenpunkt k

$$\bar{\mathbf{e}}_k := \mathbf{z}_k(\mathbf{x}_k, \mathbf{w}) - \mathbf{t}_k = [\mathbf{z}(\mathbf{x}, \mathbf{w}) - \mathbf{t}]_k \quad (6.26)$$

lässt sich Gleichung (6.25) auch kompakt schreiben als

$$\bar{\mathcal{F}}(\mathbf{w}) = \frac{1}{2P} \sum_{k=1}^P \|\mathbf{z}_k(\mathbf{x}_k, \mathbf{w}) - \mathbf{t}_k\|^2 = \frac{1}{2P} \sum_{k=1}^P \|\bar{\mathbf{e}}_k\|^2. \quad (6.27)$$

Der Faktor $1/2$ wird dabei lediglich aus ästhetischen Gründen – die bei der Berechnung des Gradienten erkenntlich werden – eingeführt. Der waagrechte Strich indiziert den Fehler bezüglich der in dem physikalischen Raum definierten Größen. Das KNN-Training erfolgt auf Basis des normalisierten Fehlers:

$$\mathcal{F}(\mathbf{w}) := \frac{1}{2P} \sum_{k=1}^P \|\hat{\mathbf{z}}_k(\hat{\mathbf{x}}_k, \mathbf{w}) - \hat{\mathbf{t}}_k\|^2 = \frac{1}{2P} \sum_{k=1}^P \|\mathbf{e}_k\|^2. \quad (6.28)$$

Das Ziel des Trainings ist die Approximation $\hat{\mathbf{z}}_k \approx \hat{\mathbf{t}}_k$ was, sich äquivalent formulieren lässt als die Minimierung von $\mathcal{F}(\mathbf{w})$. Mathematisch definiert sich dies als

$$\tilde{\mathbf{w}} = \arg \min \mathcal{F}(\mathbf{w}). \quad (6.29)$$

Im Weiteren wird der Einfluss eines Gewichts auf die Fehlerfunktion benötigt. Dieser lässt sich über den Gradienten

$$\nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}) := \frac{\partial \mathcal{F}}{\partial \mathbf{w}} = \left[\frac{\partial \mathcal{F}}{\partial w_i} \right]_{i=1}^{n_w} \quad (6.30)$$

mit

$$\frac{\partial \mathcal{F}}{\partial \mathbf{w}} = \left[\frac{\partial \mathcal{F}}{\partial w_{10}^{(1)}}, \dots, \frac{\partial \mathcal{F}}{\partial w_{1n_i}^{(1)}}, \dots, \frac{\partial \mathcal{F}}{\partial w_{ml}^{(L)}}, \dots, \frac{\partial \mathcal{F}}{\partial w_{n_o0}^{(n_h+1)}}, \dots, \frac{\partial \mathcal{F}}{\partial w_{n_on_{n_h}}^{(n_h+1)}} \right]^T \in \mathbb{R}^{n_w} \quad (6.31)$$

definieren. Die Sortierung entspricht dabei der des Gewichtsvektors (6.15). Bei gewissen Verfahren werden nicht nur die ersten, sondern auch die zweiten partiellen Ableitungen der Fehlerfunktion benötigt. Zweckmäßig ist dahingehend hier auch die Einführung der Hesse-Matrix

$$\nabla_{\mathbf{w}}^2 \mathcal{F}(\mathbf{w}) := \frac{\partial^2 \mathcal{F}}{\partial \mathbf{w} \partial \mathbf{w}^T} = \left[\frac{\partial^2 \mathcal{F}}{\partial w_i \partial w_j} \right]_{i,j=1}^{n_w} \quad (6.32)$$

mit der Anordnung

$$\nabla_{\mathbf{w}}^2 \mathcal{F}(\mathbf{w}) = \begin{bmatrix} \frac{\partial^2 \mathcal{F}}{\partial w_{10}^{(1)} \partial w_{10}^{(1)}} & \frac{\partial^2 \mathcal{F}}{\partial w_{10}^{(1)} \partial w_{11}^{(1)}} & \cdots & \frac{\partial^2 \mathcal{F}}{\partial w_{10}^{(1)} \partial w_{n_on_{n_h}}^{(n_h+1)}} \\ \frac{\partial^2 \mathcal{F}}{\partial w_{20}^{(1)} \partial w_{10}^{(1)}} & \frac{\partial^2 \mathcal{F}}{\partial w_{20}^{(1)} \partial w_{11}^{(1)}} & \cdots & \frac{\partial^2 \mathcal{F}}{\partial w_{20}^{(1)} \partial w_{n_on_{n_h}}^{(n_h+1)}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \mathcal{F}}{\partial w_{n_on_{n_h}}^{(n_h+1)} \partial w_{10}^{(1)}} & \frac{\partial^2 \mathcal{F}}{\partial w_{n_on_{n_h}}^{(n_h+1)} \partial w_{11}^{(1)}} & \cdots & \frac{\partial^2 \mathcal{F}}{\partial w_{n_on_{n_h}}^{(n_h+1)} \partial w_{n_on_{n_h}}^{(n_h+1)}} \end{bmatrix} \in \mathbb{R}^{n_w \times n_w} \quad (6.33)$$

vorzunehmen. Dem Kapitel 6.2 vorweggreifend sei an dieser Stelle bereits erwähnt, dass die zweiten Ableitungen – im Rahmen dieser Arbeit – nicht explizit berechnet werden müssen, sondern anderweitig approximiert werden können.

Im Weiteren sind die Einträge des Gradienten (6.31) von Interesse. Mit Gleichung (6.28) folgt

$$\nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}) = \frac{1}{P} \sum_{k=1}^P \left[\left(\hat{\mathbf{z}}(\hat{\mathbf{x}}, \mathbf{w}) - \hat{\mathbf{t}} \right) \nabla_{\mathbf{w}} \hat{\mathbf{z}} \right]_k. \quad (6.34)$$

Unter Anwendung der Kettenregel werden zunächst die Ableitungen $\nabla_{\mathbf{w}} \hat{\mathbf{z}}$ nach den der Ausgangsschicht zugeordneten Gewichten $w_{jo}^{(n_h+1)}$ komponentenweise betrachtet³⁰:

$$\frac{\partial \hat{z}_j}{\partial w_{jo}^{(n_h+1)}} = \frac{\partial y_j^{(n_h+1)}}{\partial w_{jo}^{(n_h+1)}} = \frac{\partial y_j^{(n_h+1)}}{\partial s_j^{(n_h+1)}} \frac{\partial s_j^{(n_h+1)}}{\partial w_{jo}^{(n_h+1)}}. \quad (6.35)$$

³⁰Vgl. Abbildung 6.1 für die hier verwendete Notation der Neuronen.

Damit kann die Ableitung des Fehlers nach einem Gewicht der Ausgangsschicht berechnet werden:

$$\frac{\partial \mathcal{F}}{\partial w_{jo}^{(n_h+1)}} = \frac{\partial \mathcal{F}}{\partial y_j^{(n_h+1)}} \frac{\partial y_j^{(n_h+1)}}{\partial s_j^{(n_h+1)}} \frac{\partial s_j}{\partial w_{jo}^{(n_h+1)}} \quad (6.36)$$

$$= (z_j - t_j) \cdot 1 \cdot y_o^{(n_h)} \quad (6.37)$$

$$= \delta_j^{(n_h+1)} \cdot y_o^{(n_h)} \quad (6.38)$$

Dabei wurden mit $\delta_j^{(n_h+1)}$ die sogenannten Delta-Werte eingeführt, die sich auch allgemein durch

$$\delta_j^{(L)} := \frac{\partial \mathcal{F}}{\partial y_j^{(L)}} \frac{\partial y_j^{(L)}}{\partial s_j^{(L)}} = \frac{\partial \mathcal{F}}{\partial s_j^{(L)}} \quad (6.39)$$

definieren lassen. Für lineare, sogenannte Single-Layer-Perceptrons geht die Einführung dieser auf die Arbeiten von WIDROW und HOFF [150] zurück. Inhaltlich steht hinter der dazu assoziierten Delta-Regel die Aussage, dass der Einfluss eines Gewichts auf den Fehler proportional zur Differenz zwischen Ausgabe- und Zielwert des entsprechenden Neurons ist. Diese können im Vektor $\boldsymbol{\delta}^{(L)} \in \mathbb{R}^{n_L}$ zusammengefasst werden, womit sich (6.38) auch kompakt schreiben lässt als

$$\begin{aligned} \frac{\partial F}{\partial \mathbf{w}^{(n_h+1)}} &= \left[(\hat{\mathbf{z}}(\hat{\mathbf{x}}, \mathbf{w}) - \hat{\mathbf{t}}) \odot \mathbf{1}_{n_o} \right] \mathbf{y}^{(n_h)T} \\ &= [\mathbf{e}_k \odot \mathbf{1}_{n_o}] \mathbf{y}^{(n_h)T} \\ &= \boldsymbol{\delta}^{(n_h+1)} \mathbf{y}^{(n_h)T} \in \mathbb{R}^{n_o \times n_h+1}. \end{aligned} \quad (6.40)$$

Der Einheitsvektor $\mathbf{1}_{n_o}$ wurde dabei nur aus formalen Gründen mitgeführt, da er die Ableitung der Aktivierungsfunktion in der Ausgangsschicht darstellt. Die Berechnungsvorschrift (6.40) führt auf die Matrix der Gradienteneinträge

$$\frac{\partial F}{\partial \mathbf{w}^{(n_h+1)}} = \begin{bmatrix} \frac{\partial F}{w_{10}^{(n_h+1)}} & \frac{\partial F}{w_{11}^{(n_h+1)}} & \cdots & \frac{\partial F}{w_{1n_h}^{(n_h+1)}} \\ \frac{\partial F}{w_{20}^{(n_h+1)}} & \frac{\partial F}{w_{21}^{(n_h+1)}} & \cdots & \frac{\partial F}{w_{2n_h}^{(n_h+1)}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F}{w_{n_o0}^{(n_h+1)}} & \frac{\partial F}{w_{n_o1}^{(n_h+1)}} & \cdots & \frac{\partial F}{w_{n_on_h}^{(n_h+1)}} \end{bmatrix} \in \mathbb{R}^{n_o \times (n_h+1)}, \quad (6.41)$$

die in den Vektor (6.31) einsortiert werden können. Für die Berechnung der partiellen Ableitungen des Fehlers nach einem Gewicht der Zwischenschichten folgt – ebenfalls durch Anwendung der Kettenregel – komponentenweise:

$$\frac{\partial \mathcal{F}}{\partial w_{ml}^{(L)}} = \frac{\partial \mathcal{F}}{\partial y_m^{(L)}} \frac{\partial y_m^{(L)}}{\partial s_m^{(L)}} \frac{\partial s_m^{(L)}}{\partial w_{ml}^{(L)}}. \quad (6.42)$$

Mit der Definition für die Ableitung der Aktivierungsfunktion nach der gewichteten Summe

$$g'(s_m^{(L)}) := \frac{\partial y_m^{(L)}}{\partial s_m^{(L)}} = \frac{\partial g(s_m^{(L)})}{\partial s_m^{(L)}} \quad (6.43)$$

wird die Beziehung

$$\frac{\partial \mathcal{F}}{\partial w_{ml}^{(L)}} = \frac{\partial \mathcal{F}}{\partial y_m^{(L)}} g'(s_m^{(L)}) y_l^{(L-1)} \quad (6.44)$$

erhalten. An dieser Stelle stellt sich die Frage nach der Berechnung des ersten Ableitungsterms der rechten Seite von Gleichung (6.44). Hierfür muss die Überlegung angestellt werden, dass sich die Abhängigkeit des Fehlers zum Ausgang einer Schicht auch verkettet über die gewichtete Summe der nachfolgenden Schicht darstellen lässt:

$$\mathcal{F}(y_m^{(L)}) = \mathcal{F}(s_1^{(L+1)}(y_m^{(L)}), s_2^{(L+1)}(y_m^{(L)}), \dots, s_n^{(L+1)}(y_m^{(L)}), \dots, s_{n_L+1}^{(L+1)}(y_m^{(L)})) . \quad (6.45)$$

Damit folgt

$$\frac{\partial \mathcal{F}}{\partial y_m^{(L)}} = \sum_{n=1}^{n_{L+1}} \frac{\partial \mathcal{F}}{\partial s_n^{(L+1)}} \frac{\partial s_n^{(L+1)}}{\partial y_m^{(L)}} = \sum_{n=1}^{n_{L+1}} \frac{\partial \mathcal{F}}{\partial s_n^{(L+1)}} w_{nm}^{(L+1)} . \quad (6.46)$$

Die erneute Anwendung der Kettenregel auf den ersten Term der rechten Seite führt auf

$$\frac{\partial \mathcal{F}}{\partial s_n^{(L+1)}} = \frac{\partial \mathcal{F}}{\partial y_n^{(L+1)}} \frac{\partial y_n^{(L+1)}}{\partial s_n^{(L+1)}} = \delta_n^{(L+1)} . \quad (6.47)$$

Aus (6.47) wird ersichtlich, dass für die Berechnung der Gradienteneinträge bezüglich der Gewichte einer Zwischenlicht L die Delta-Werte der nachfolgenden Schicht benötigt werden. Dahingehend wird häufig die Bezeichnung generalisierte Delta-Werte verwendet, da sich hier das Prinzip der Delta-Regel auch für MLPs erkennen lässt. Das Prinzip des Fehlerrückführens – auch *Backpropagation of Error* genannt – wird dabei den Arbeiten von RUMELHART ET AL. [115] zugeschrieben.

Unter Einsetzen der Beziehungen (6.46) und (6.47) in (6.44) folgt letztlich

$$\frac{\partial \mathcal{F}}{\partial w_{ml}^{(L)}} = \sum_{n=1}^{n_{L+1}} \delta_n^{(L+1)} w_{nm}^{(L+1)} g'(s_m^{(L)}) y_l^{(L-1)} , \quad (6.48)$$

wobei sich hier wiederum die generalisierten Delta-Werte der Schicht L identifizieren lassen:

$$\delta_m^{(L)} := \sum_{n=1}^{n_{L+1}} \delta_n^{(L+1)} w_{nm}^{(L+1)} g'(s_m^{(L)}) . \quad (6.49)$$

Vorwärtsrechnung (Alg. 6.1): $\hat{\mathbf{x}} \mapsto \hat{\mathbf{z}}$, für Datenpunkt k
Speicherung: $\hat{\mathbf{z}}, \mathbf{y}^{(L)}, \mathbf{s}^{(L)}$
Ausgangsschicht: $\mathbf{e}_k = \hat{\mathbf{z}} - \hat{\mathbf{t}}$
 $\boldsymbol{\delta}^{(n_h+1)} = \mathbf{e} \odot \mathbf{g}'(\hat{\mathbf{z}})$
 $\nabla_{\mathbf{w}^{(n_h+1)}} \mathcal{F} = \boldsymbol{\delta}^{(n_h+1)} \mathbf{y}^{(n_h)}$
Schleife über alle Schichten: **for** $L = n_h, \dots, 1$
 $\boldsymbol{\delta}^{(L)} = \left(\boldsymbol{\delta}^{(L+1)^T} \bar{\mathbf{w}}^{(L+1)} \right)^T \odot g'(\mathbf{s}^{(L)})$
 $\nabla_{\mathbf{w}^{(L)}} \mathcal{F} = \boldsymbol{\delta}^{(L)} \mathbf{y}^{(L-1)}$
end
Sortierung (optional): $\nabla_{\mathbf{w}^{(L)}} \mathcal{F} \rightarrow \nabla_{\mathbf{w}} \mathcal{F}, L = 1, \dots, n_h + 1$

Algorithmus 6.2: Berechnungsvorschrift für die FFN-Rückwärtsrechnung

Erneut wird die Vektor-Matrix Notation angewendet, um direkt die Matrix der Gradienteneinträge der Schicht L

$$\frac{\partial \mathcal{F}}{\partial \mathbf{w}^{(L)}} = \left[\left(\boldsymbol{\delta}^{(L+1)^T} \bar{\mathbf{w}}^{(L+1)} \right)^T \odot g'(\mathbf{s}^{(L)}) \right] \mathbf{y}^{(L-1)^T} \in \mathbb{R}^{n_L \times (n_{L-1}+1)} \quad (6.50)$$

zu erhalten, sowie den Vektor

$$\boldsymbol{\delta}^{(L)} := \left(\boldsymbol{\delta}^{(L+1)^T} \bar{\mathbf{w}}^{(L+1)} \right)^T \odot g'(\mathbf{s}^{(L)}) \in \mathbb{R}^{n_L}. \quad (6.51)$$

Die beschriebenen Schritte können zusammenfassend für eine Implementierung aus Algorithmus 6.2 entnommen werden³¹. Zu beachten ist dabei, dass die Herleitungen für *einen* Datenpunkt erfolgten und eine entsprechende Schleife über alle Datenpunkte P übergeordnet ergänzt werden muss.

6.1.4 Partielle Ableitungen der Ausgangsgrößen

In bestimmten Fällen werden nicht nur die partiellen Ableitungen nach den Gewichten, sondern auch jene nach den Eingängen des KNN benötigt. Konkret sind im Folgenden die Einträge der KNN-Tangente³² im physikalischen Raum

$$\bar{\mathbf{c}}_T^{\text{KNN}} := \frac{\partial \mathbf{z}}{\partial \mathbf{x}} = \left[(\mathbf{t}_{\max} - \mathbf{t}_{\min}) \oslash (\mathbf{x}_{\max} - \mathbf{x}_{\min})^T \right] \circ \frac{\partial \hat{\mathbf{z}}}{\partial \hat{\mathbf{x}}} \in \mathbb{R}^{n_o \times n_i} \quad (6.52)$$

³¹Gerade bei komplexen KNN-Architekturen wird die Berechnung partieller Ableitungen häufig mittels automatischer Differenzierung vorgenommen. Auch hier kann nur die Aussage wiederholt werden, dass die Herleitung und Implementierung aus Sicht des Autos zu einer flexiblen Programmstruktur und – persönlicher – Entschärfung des „Black Box“-Charakters geführt haben.

³²Formal wäre hier auch der Begriff KNN Jacobi-Matrix korrekt. Dieser wird hier jedoch für die Bezeichnung der partiellen Ableitungen $\partial \mathbf{e}_k / \partial \mathbf{w}$ verwendet, die später für das KNN-Training benötigt werden.

gesucht. Dabei wurde das Hadamard-Produkt $\circ$ zur komponentenweise angewendeten Multiplikation der Einträge zweier *Matrizen* eingeführt, sowie die Transformationsbeziehungen (6.10) und (6.11) angewendet. Entsprechend wird die Tangente im normalisierten Raum

$$\mathbf{C}_T^{\text{KNN}} := \frac{\partial \hat{\mathbf{z}}}{\partial \hat{\mathbf{x}}} = \begin{bmatrix} \frac{\partial \hat{z}_1}{\partial \hat{x}_1} & \frac{\partial \hat{z}_1}{\partial \hat{x}_2} & \cdots & \frac{\partial \hat{z}_1}{\partial \hat{x}_{n_i}} \\ \frac{\partial \hat{z}_2}{\partial \hat{x}_1} & \frac{\partial \hat{z}_2}{\partial \hat{x}_2} & \cdots & \frac{\partial \hat{z}_2}{\partial \hat{x}_{n_i}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \hat{z}_{n_o}}{\partial \hat{x}_1} & \frac{\partial \hat{z}_{n_o}}{\partial \hat{x}_2} & \cdots & \frac{\partial \hat{z}_{n_o}}{\partial \hat{x}_{n_i}} \end{bmatrix} \in \mathbb{R}^{n_o \times n_i} \quad (6.53)$$

definiert. Die KNN-Tangente wird hier für zwei Zwecke benötigt:

- **Erweiterung des Trainings:**

Neben den Primärdaten $(\mathbf{x}_k, \mathbf{t}_k)$ können – sofern die partiellen Ableitungen $(\mathbf{x}_k, \partial \mathbf{t}_k / \partial \mathbf{x}_k)$ bekannt sind – auch diese in den Lernprozess einbezogen werden. Dieses Vorgehen wird als Sobolev Training bezeichnet und in Kapitel 6.3.2 ausführlich diskutiert.

- **Einsatz als Materialmodell:**

Wird das KNN als Materialmodell eingesetzt, so ist im Rahmen der Linearisierung – vgl. auch (2.44) – die Materialtangente $\mathbf{C}_T$ erforderlich. Diese ist sowohl für die FEM als auch für synthetische Materialpunkt-Versuche von Bedeutung. Beides wird in Kapitel 7 behandelt.

Die folgende Berechnungsvorschrift der Einträge der normalisierten Tangente orientiert sich an der Herleitung in WEBER [145], die direkt auf eine kompakte Vektor-Matrix Notation führt. Unter Anwendung der Kettenregel erhält man die Rekursionsformel

$$\frac{\partial \hat{\mathbf{z}}}{\partial \hat{\mathbf{x}}} = \frac{\partial \bar{\mathbf{y}}^{(n_h+1)}}{\partial \bar{\mathbf{y}}^{(0)}} = \prod_{L=n_h+1}^1 \frac{\partial \bar{\mathbf{y}}^{(L)}}{\partial \bar{\mathbf{y}}^{(L-1)}}. \quad (6.54)$$

Die Betrachtung der Komponente der Ableitung des Ausgangs eines Neurons der Schicht L nach einem der Schicht $L-1$ liefert

$$\frac{\partial \bar{y}_m^{(L)}}{\partial \bar{y}_n^{(L-1)}} = \frac{\partial \bar{y}_m^{(L)}}{\partial s_m^{(L)}} \frac{\partial s_m^{(L)}}{\partial \bar{y}_n^{(L-1)}} = g'(s_m^{(L)}) w_{mn}^{(L)}. \quad (6.55)$$

Ähnlich wie die zuvor eingeführten Delta-Werte, werden in diesem Fall die Ableitungen der Aktivierungsfunktionen Schicht für Schicht zurück „propagiert“. Dies lässt sich kompakt über

$$\frac{\partial \bar{\mathbf{y}}^{(L)}}{\partial \bar{\mathbf{y}}^{(L-1)}} = \left[g'(\mathbf{s}^{(L)}) \mathbf{1}_{n_{L-1}}^T \right] \circ \bar{\mathbf{w}}^{(L)} \quad (6.56)$$

Vorwärtsrechnung (Alg. 6.1):	$\hat{\mathbf{x}} \mapsto \hat{\mathbf{z}}$
Speicherung:	$\mathbf{s}^{(L)}, \quad L = 1, \dots, n_h + 1$
Ausgangsschicht:	$\mathbf{g}'(\mathbf{s}^{(n_h+1)}) = \mathbf{1}_{n_o} \mathbf{1}_{n_h}^T$ $\left[\frac{\partial \hat{\mathbf{z}}}{\partial \hat{\mathbf{x}}} \right]^{(n_h+1)} = \mathbf{g}'(\mathbf{s}^{(n_h+1)}) \circ \bar{\mathbf{w}}^{(n_h+1)}$
Schleife über alle Schichten:	for $L = n_h, \dots, 1$ $\left[\frac{\partial \hat{\mathbf{z}}}{\partial \hat{\mathbf{x}}} \right]^{(L)} = \mathbf{g}'(\mathbf{s}^{(n_h+1)}) \circ \bar{\mathbf{w}}^{(L)}$ $\left[\frac{\partial \hat{\mathbf{z}}}{\partial \hat{\mathbf{x}}} \right] = \left[\frac{\partial \hat{\mathbf{z}}}{\partial \hat{\mathbf{x}}} \right]^{(L+1)} \left[\frac{\partial \hat{\mathbf{z}}}{\partial \hat{\mathbf{x}}} \right]^{(L)}$ end
Transformation:	$\frac{\partial \mathbf{z}}{\partial \mathbf{x}} = \dots$ $\left[(\mathbf{t}_{\max} - \mathbf{t}_{\min}) \oslash (\mathbf{x}_{\max} - \mathbf{x}_{\min})^T \right] \circ \frac{\partial \hat{\mathbf{z}}}{\partial \hat{\mathbf{x}}}$

Algorithmus 6.3: Berechnungsvorschrift für die KNN-Tangente $\partial \mathbf{z} / \partial \mathbf{x}$

für alle partiellen Ableitungen der Ausgänge einer Schicht nach den vorherigen darstellen. Einsetzen der Beziehung in (6.54) liefert

$$\frac{\partial \hat{\mathbf{z}}}{\partial \hat{\mathbf{x}}} = \prod_{L=n_h+1}^1 \left[\mathbf{g}'(\mathbf{s}^{(L)}) \mathbf{1}_{n_{L-1}}^T \right] \circ \bar{\mathbf{w}}^{(L)} = \prod_{L=n_h+1}^1 \mathbf{g}'(\mathbf{s}^{(L)}) \circ \bar{\mathbf{w}}^{(L)}, \quad (6.57)$$

mit der Matrix der Ableitungen der gewichteten Summen einer Schicht L

$$\mathbf{g}'(\mathbf{s}^{(L)}) = \begin{bmatrix} g'(s_1^{(L)}) & \dots & g'(s_1^{(L)}) \\ \vdots & \ddots & \vdots \\ g'(s_{n_L}^{(L)}) & \dots & g'(s_{n_L}^{(L)}) \end{bmatrix} \in \mathbb{R}^{n_L \times n_{L-1}} \quad (6.58)$$

Die gewichteten Summen $\mathbf{s}^{(L)}$ werden durch eine Vorwärtsrechnung bestimmt und abgespeichert. Alle vorzunehmenden Schritte zur Berechnung der KNN-Tangente sind in Algorithmus 6.3 zusammengefasst.

6.2 Optimierungsverfahren - Das Training des KNN

Die noch ausstehende Fragestellung, die sich nach den ausführlichen Herleitungen der vorherigen Kapitel stellt, ist die nach der Bestimmung eines geeigneten Satzes an Gewichten. Hierfür wird die eigentliche Aufgabe des KNN-Trainings nochmals

aufgriffen (vgl. Gleichung (6.29)). Mit der Definition von

$$\tilde{\mathbf{w}} = \arg \min \mathcal{F}(\mathbf{w}) . \quad (6.59)$$

wird ein geeigneter Satz an Gewichten $\tilde{\mathbf{w}}$ zur Minimierung von $\mathcal{F}(\mathbf{w})$ gesucht. In der dargestellten Weise ist dies die klassische Formulierung eines Optimierungsproblems. Dass es sich beim KNN-Training jedoch nicht um klassische Optimierungsprobleme handelt, wird im Folgenden ersichtlich. Ausgangspunkt dafür ist, dass an KNNs die Anforderung gestellt wird, nicht nur Daten auswendig zu lernen, sondern die Fähigkeit zur Generalisierung aufzuweisen. Um die Erfüllung dieser Anforderung zu prüfen, wird der Datensatz (6.3) in eine Teilmenge an Trainingsdaten

$$\mathcal{D}^{\text{Tr}} = \{(\mathbf{x}_k, \mathbf{t}_k) \mid \mathbf{x}_k \in \mathcal{X} \subset \mathbb{R}^{n_i}, k = 1, 2, \dots, P^{\text{Tr}}\} \quad (6.60)$$

und eine Teilmenge an Testdaten

$$\mathcal{D}^{\text{Test}} = \{(\mathbf{x}_k, \mathbf{t}_k) \mid \mathbf{x}_k \in \mathcal{X} \subset \mathbb{R}^{n_i}, k = 1, 2, \dots, P^{\text{Test}}\} \quad (6.61)$$

unterteilt, sodass

$$\mathcal{D}^{\text{Tr}} \cup \mathcal{D}^{\text{Test}} = \mathcal{D} \quad \text{und} \quad \mathcal{D}^{\text{Tr}} \cap \mathcal{D}^{\text{Test}} = \emptyset . \quad (6.62)$$

Es wird also anhand eines Satzes an Daten trainiert und die Fähigkeit zur Generalisierung an einem unbekannten Datensatz geprüft. Entsprechend erfolgt die Definition zweier Fehlermaße, dem bezüglich der normalisierten Trainingsdaten

$$\mathcal{F}^{\text{Tr}}(\mathbf{w}) := \frac{1}{2P^{\text{Tr}}} \sum_{k=1}^{P^{\text{Tr}}} \|\hat{\mathbf{z}}_k(\hat{\mathbf{x}}_k, \mathbf{w}) - \hat{\mathbf{t}}_k\|^2 \quad (6.63)$$

und dem bezüglich der normalisierten Testdaten (Test-MSE)

$$\mathcal{F}^{\text{Test}}(\mathbf{w}) := \frac{1}{2P^{\text{Test}}} \sum_{k=1}^{P^{\text{Test}}} \|(\hat{\mathbf{z}}_k(\hat{\mathbf{x}}_k, \mathbf{w}) - \hat{\mathbf{t}}_k)\|^2 . \quad (6.64)$$

Damit unterscheidet sich das Training von KNNs von klassischen Optimierungsproblemen, bei denen die Zielfunktion zugleich den Erfolg der Optimierung quantifiziert (vgl. GOODFELLOW ET AL. [42]). Da für das Training im Weiteren nur der Gradient und die Hesse-Matrix von $\mathcal{F}^{\text{Tr}}(\mathbf{w})$ benötigt werden, gilt

$$\nabla_{\mathbf{w}} \mathcal{F} := \nabla_{\mathbf{w}} \mathcal{F}^{\text{Tr}} \quad \text{und} \quad \nabla_{\mathbf{w}}^2 \mathcal{F} = \nabla_{\mathbf{w}}^2 \mathcal{F}^{\text{Tr}} . \quad (6.65)$$

Im Folgenden werden drei der implementierten Verfahren vorgestellt, die zur Minimierung von $\mathcal{F}^{\text{Tr}}$ verwendet werden können. Grundsätzlich lassen sich diese zwei verschiedenen Kategorien zuordnen:

- **Verfahren 1. Ordnung** verwenden nur die Informationen des Fehlergradienten $\nabla_{\mathbf{w}}\mathcal{F}$ zur Aktualisierung der Gewichte. In Kombination mit einem Skalierungsfaktor sind verschiedene Abwandlungen dieser Verfahren vor allem für hochdimensionale Problemstellung geeignet. Sie berücksichtigen allerdings die Krümmungsinformation – die Änderung des Gradienten – der Fehlerfläche nicht.
- **Verfahren 2. Ordnung** beziehen zusätzlich Krümmungsinformationen über die Hesse-Matrix $\nabla_{\mathbf{w}}^2\mathcal{F}$ in die Aktualisierungsvorschrift mit ein. Dies führt zu einem deutlich besseren Konvergenzverhalten. Die – häufig nur approximative – Bestimmung der Hesse-Matrix sowie die notwendigen Schritte bei der Aktualisierung gehen allerdings mit einem erhöhten Berechnungsaufwand einher.

Abschließend sei angemerkt, dass die hier vorgestellten Verfahren nur einen unvollständigen Teil des breiten Spektrums an existierenden Methoden darstellen. In Abhängigkeit der Problemstellung können eine ganze Reihe weiterer Verfahren geeignet sein, wofür der Leser an die bereits genannte Literatur verwiesen wird.

6.2.1 Das Verfahren des steilsten Abstiegs

Das Verfahren des steilsten Abstiegs – auch als Gradientenabstiegsverfahren bekannt und durch CAUCHY [19] bereits 1847 eingeführt – ist das wohl einfachste Verfahren zur iterativen Bestimmung des Minimums einer Funktion. Im mehrdimensionalen Fall wird hierfür zunächst die Richtungsableitung (vgl. (2.35)) von $\mathcal{F}(\mathbf{w})$ in Richtung eines beliebigen Vektors $\mathbf{v}$ mit $|\mathbf{v}| = 1$ betrachtet:

$$\left. \frac{d}{d\alpha} \mathcal{F}(\mathbf{w} + \alpha \cdot \mathbf{v}) \right|_{\alpha=0} = \mathbf{v}^T \nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}) = |\mathbf{v}| |\nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w})| \cos(\theta), \quad (6.66)$$

wobei die Definition des Skalarprodukts angewendet wurde. Lediglich der zwischen den beiden Vektoren eingeschlossenen Winkel θ ist hier die unabhängige Variable. Gleichung (6.66) wird also gerade dann minimal, wenn $\mathbf{v}$ und $\nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w})$ in entgegengesetzte Richtung zeigen ($\cos(\pi) = -1$). Der negative Gradient ist dementsprechend die Richtung des stärksten Abstiegs. Damit lautete die allgemeine Form der Aktualisierungsvorschrift für die Gewichte einer Schicht (L):

$$\mathbf{w}_{[n+1]}^{(L)} = \mathbf{w}_{[n]}^{(L)} - \eta \nabla_{\mathbf{w}^{(L)}} \mathcal{F}(\mathbf{w}). \quad (6.67)$$

Der Index n zählt dabei die Gewichtsaktualisierungen während des Trainings. Dies vervollständigt den Algorithmus der Fehlerrückführung nach RUMELHART ET AL. [115]. Der skalare Faktor η , die sogenannte Lernrate, steuert die Schrittweite der Gewichtsaktualisierung und stellt einen Hyperparameter des KNN-Trainings

dar, der zu unterscheiden ist von den Hyperparametern der Topologie, wie bspw. T der Aktivierungsfunktion. Typischerweise wird sie bei Netzen mit moderater Anzahl an Zwischenschichten im Bereich von $[0,001, 0,1]$ gewählt. Eine geeignete Wahl der Lernrate kann zudem dem Problem des *verschwindenden Gradienten*³³ entgegenwirken, das dadurch entsteht, dass der Einfluss der Gewichte auf den Fehler in den eingangsnahen Schichten zunehmend abnimmt. Das Gewichtsupdate kann entweder wie dargestellt Schicht für Schicht³⁴, oder durch Anordnung der Schrittweiten in dem Vektor $\boldsymbol{\eta} \in \mathbb{R}^{n_w}$ direkt für alle Gewichte durchgeführt werden:

$$\mathbf{w}_{[n+1]} = \mathbf{w}_{[n]} - \boldsymbol{\eta} \nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}). \quad (6.68)$$

Eine Epoche bezeichnet einen vollständigen Durchlauf über alle Trainingsdaten. Die Trainingsverfahren unterscheiden sich darin, wie viele Daten zur Berechnung des Gradienten herangezogen werden. Beim *Batch*-Verfahren fließen alle Daten $k = 1, \dots, P^{\text{train}}$ in die Gradientenberechnung ein, was dem exakten Gradienten entspricht. Beim Einsatz *stochastischer Gradientenverfahren* wird der Gradienten approximativ anhand zufälliger Teilmengen der Daten (*Minibatches*) berechnet. Im *online*-Training erfolgt die Aktualisierung nach jedem einzelnen Datenpunkt k . Auf eine Implementierungsvorschrift wird an dieser Stelle verzichtet; hierfür kann Algorithmus 6.2 herangezogen werden, ergänzt um eine äußere Schleife über die Trainingsdaten zur Gradientenberechnung.

6.2.2 Der ADAM Optimierungsalgorithmus

Der einfache Gradientenabstieg gemäß (6.68) weist aufgrund der konstanten Lernrate einige Einschränkungen auf. Die Wahl geeigneter Hyperparameter erfordert zum einen Erfahrung und zum anderen fehlt eine dynamische Anpassung an die Form der Fehlerfläche. Zahlreiche Erweiterungen des Verfahrens – etwa RPROP (vgl. RIEDMILLER [110]), ADAGRAD (vgl. DUCHI [29]) oder RMSPROP – verfolgen daher adaptive Strategien. Besonders vielversprechend ist in diesem Zusammenhang der Adaptive-Moments-Algorithmus (ADAM) von KINGMA und BA [66], der sich insbesondere im Bereich des *Deep Learning*, also bei KNN-Architekturen mit vielen verdeckten Schichten, großer Beliebtheit erfreut.

Der ADAM-Algorithmus berücksichtigt den Einfluss vergangener Gradienten über einen gleitenden Mittelwert, was auch als *Momentum* bezeichnet wird. Zusätzlich werden große Gradienten über das geschätzte zweite Moment skaliert, sodass

³³Das grundsätzliche Auftreten des Phänomens ist mit der Wahl der Aktivierungsfunktion assoziiert. Durch die Plateaubereiche des Tangens hyperbolicus können die Ableitungen der Aktivierungsfunktionen Werte nahe 0 annehmen. Durch das Zurückführen der Delta-Werte schaukelt sich das Problem hoch und wirkt sich vor allem auf die eingangsnahen Gewichte aus.

³⁴Optional kann noch entschieden werden, ob die Rückführung der Delta-Werte mit den bereits aktualisierten Gewichten erfolgt.

Wahl der Parameter:	$\eta, \beta_1, \beta_2, \epsilon$ Anzahl Minibatches n_B
Initialisierung:	$\mathbf{m}_{[0]} = \mathbf{0}, \mathbf{v}_{[0]} = \mathbf{0}, \mathbf{w}_{[0]}, n = 0, t = 0$
Training:	while (kein Abbruch) do $n = n + 1$ $\mathcal{D}^B := \mathcal{D}^{\text{Train}}$ for $i = 1, \dots, n_B$ $P_i^B = P^{\text{Train}} / n_B$ $\mathcal{D}^{\text{Train}} = \{(\mathbf{x}_k, \mathbf{t}_k) \mid k = 1, \dots, P_i^B\} \subset \mathcal{D}^B$ $\mathcal{D}^B := \mathcal{D}^{\text{Train}} \setminus \mathcal{D}^B$ Gradient (Alg. 6.2): $\nabla_{\mathbf{w}} \mathcal{F}^B$ von Batch $\mathcal{D}^B$ $\mathbf{m}_{[t+1]} = \beta_1 \mathbf{m}_{[t]} + (1 - \beta_1) \nabla_{\mathbf{w}} \mathcal{F}$ $\mathbf{v}_{[t+1]} = \beta_2 \mathbf{v}_{[t]} + (1 - \beta_2) \nabla_{\mathbf{w}} \mathcal{F} \odot \nabla_{\mathbf{w}} \mathcal{F}$ $\hat{\mathbf{m}}_{[t+1]} = \mathbf{m}_{[t+1]} / (1 - \beta_1^{(t+1)})$ $\hat{\mathbf{v}}_{[t+1]} = \mathbf{v}_{[t+1]} / (1 - \beta_2^{(t+1)})$ $\Delta \mathbf{w} = \eta \hat{\mathbf{m}}_{[t+1]} \oslash (\sqrt{\hat{\mathbf{v}}_{[t+1]}} + \epsilon \mathbf{1}_{nw})$ $\mathbf{w}_{[t+1]} = \mathbf{w}_{[t]} - \Delta \mathbf{w}$ $t = t + 1$ end end

Algorithmus 6.4: Berechnungsvorschrift für den ADAM-Algorithmus

einzelne Parameter mit großen Gradienten die Richtung der Aktualisierung nicht dominieren. Da die Mittelwerte und zweiten Momente mit Null initialisiert werden, sind die Schätzungen zu Beginn verzerrt (Bias). Dieser Anfangsbias wird durch eine exponentielle Korrektur in Abhängigkeit vom Aktualisierungsschritt t kompensiert.

Die Berechnung des gleitenden Mittelwerts erfolgt zu Schritt $t + 1$ mit

$$\mathbf{m}_{[t+1]} = \beta_1 \mathbf{m}_{[t]} + (1 - \beta_1) \nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}), \quad (6.69)$$

sowie des zweiten Moments mit

$$\mathbf{v}_{[t+1]} = \beta_2 \mathbf{v}_{[t]} + (1 - \beta_2) \nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}) \odot \nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}). \quad (6.70)$$

Die Parameter $\beta_1, \beta_2 \in [0, 1)$ sowie die Startwerte $\mathbf{m}_{[0]}$ und $\mathbf{v}_{[0]}$ müssen vom Anwender initialisiert werden, wobei letztere üblicherweise Nullvektoren sind. Um den Initialisierungsbias zu korrigieren, werden die Schätzer normalisiert zu

$$\hat{\mathbf{m}}_{[t+1]} = \frac{\mathbf{m}_{[t+1]}}{1 - \beta_1^{(t+1)}}, \quad \hat{\mathbf{v}}_{[t+1]} = \frac{\mathbf{v}_{[t+1]}}{1 - \beta_2^{(t+1)}}. \quad (6.71)$$

Für $t = 0$ hebt die Korrektur die Verzerrung exakt auf, während sie für große t keinen Einfluss mehr hat. Das Gewichtsupdate lautet schließlich

$$\mathbf{w}_{[t+1]} = \mathbf{w}_{[t]} - \eta \hat{\mathbf{m}}_{[t+1]} \oslash \left(\sqrt{\hat{\mathbf{v}}_{[t+1]}} + \epsilon \mathbf{1}_{nw} \right). \quad (6.72)$$

Sowohl die Lernrate η als auch der Faktor ϵ , der zur numerischen Stabilisierung dient, müssen vorab festgelegt werden. Eine gängige Wahl der Parameter für den ist in Tabelle 6.5 angegeben. Der ADAM-Algorithmus fällt in die Kategorie der *stochastischen Gradientenabstiegsverfahren*. Dahingehend wird in der Regel neben den Hyperparametern auch eine Anzahl von *Minibatches* n_B gewählt werden. Die Implementierung kann Algorithmus 6.4 entnommen werden.

Parameter	Typischer Wert	Bemerkung
η (Lernrate)	10^{-3}	Kann zwischen 10^{-4} und 10^{-2} variiert werden
β_1	0.9	Steuert Glättung des ersten Moments
β_2	0.999	Steuert Glättung des zweiten Moments
ϵ	10^{-8}	Numerische Stabilisierung

Tabelle 6.5: Empfohlene Wahl der Hyperparameter nach [66] für den ADAM-Algorithmus

6.2.3 Das Levenberg-Marquardt-Verfahren

Wie eingangs bereits erwähnt, beziehen Verfahren 2. Ordnung die Krümmungseigenschaften der Fehlerfläche mit ein. Ausgangspunkt all dieser Verfahren ist das NEWTON-Verfahren, das bereits in Kapitel 2.2.3 Anwendung gefunden hat. Die Bedingung eines Minimums der Fehlerfunktion entspricht der Forderung

$$\nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}) \stackrel{!}{=} \mathbf{0}. \quad (6.73)$$

Diese nichtlineare Gleichung in $\mathbf{w}$ kann unter entsprechender Linearisierung

$$L[\nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w})]_{\mathbf{w}=\mathbf{w}_{[n]}} = \nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}_{[n]}) + \nabla_{\mathbf{w}}^2 \mathcal{F}(\mathbf{w}_{[n]}) \Delta \mathbf{w} \quad (6.74)$$

und Einsetzen der Beziehung $\Delta \mathbf{w} = \mathbf{w}_{[n+1]} - \mathbf{w}_n$ umgeformt werden zu

$$\mathbf{w}_{[n+1]} = \mathbf{w}_{[n]} - \nabla_{\mathbf{w}}^2 \mathcal{F}(\mathbf{w}_{[n]})^{-1} \nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}_{[n]}). \quad (6.75)$$

Die Aktualisierungsvorschrift (6.75) macht unmittelbar einige Nachteile in der praktischen Anwendung deutlich: zum einen muss die $n_w \times n_w$ dimensionale Hesse-Matrix berechnet werden, was grundsätzlich – unter Voraussetzung der entsprechenden Differenzierbarkeit der Fehlerfunktion – möglich ist, jedoch mit einem

erhöhten Berechnungsaufwand einhergeht. Zum anderen muss diese in jedem Iterationsschritt invertiert werden. Etablierte Verfahren 2. Ordnung im Kontext des KNN-Trainings versuchen entsprechend die Informationen über die zweiten Ableitungen zu berücksichtigen, ohne die Hesse-Matrix direkt berechnen zu müssen. Weitere Vertreter hiervon sind neben dem LEVENBERG-MARQUARDT-Verfahren (LMV) u. a. noch das Verfahren der konjugierten Gradienten (vgl. MØLLER [98]) oder das QUASI-NEWTON-Verfahren³⁵. Letzteres verzichtet sogar auf die Invertierung und approximiert direkt die Inverse der Hesse-Matrix. All diese Methoden gehören zur Kategorie der *Batch* Verfahren, die alle Informationen der Fehlerfläche berücksichtigen und damit eine hohe Präzision in der Approximation liefern (vgl. BATTITI [13]).

Aus den o. g. Gründen sind Verfahren 2. Ordnung schon früh als ungeeignet für das KNN-Training eingestuft worden (vgl. LECUN ET AL. [79])³⁶. Für das Training heutiger KNN-Architekturen, die häufig aus mehreren Millionen Freiheitsgraden (Gewichten) bestehen, ist dies sicherlich weiter zutreffend. Die im Rahmen dieser Arbeit verwendeten Netze reduzieren sich jedoch auf einige Tausend Freiheitsgrade, für die sich das LMV als ein sehr robustes und mit einer hohen Approximationsgüte einhergehende Methode herausgestellt hat. Letzteres ist vor allem bei der Materialmodellierung ein wichtiger Aspekt, da dadurch das Risiko für unphysikalisches Materialverhalten geringer und für die Numerik außerdem auch die Approximationsgüte der Materialtangente entscheidend ist.

Das LMV ist zunächst auf die Minimierung quadratischer Fehlerfunktionen beschränkt und wird in diesem Kontext den Arbeiten von LEVENBERG [83] und MARQUARDT [89] zugeschrieben. Im Kontext des KNN-Trainings wurde es erstmalig von HAGAN und MENHAJ [48] angewendet. Weitere anschauliche Herleitungen sind in WILAMOWSKI und IRWIN [151] zu finden. Dabei wird sich zunächst des GAUSS-NEWTON-Verfahrens bedient. Mit Gleichung (6.28) lässt sich die Hesse-Matrix auch schreiben als

$$\nabla_{\mathbf{w}}^2 \mathcal{F}(\mathbf{w}) = \frac{1}{P} \sum_{k=1}^P \nabla_{\mathbf{w}}^2 ||\mathbf{e}_k||^2 = \frac{1}{P} \sum_{k=1}^P \left[\nabla_{\mathbf{w}} \mathbf{e}_k^T \nabla_{\mathbf{w}} \mathbf{e}_k + \mathcal{H}_k \mathbf{e}_k \right]. \quad (6.76)$$

Dabei wurde die Matrix³⁷ dritter Stufe

$$\mathcal{H}_k := \nabla_{\mathbf{w}}^2 \mathbf{e}_k \in \mathbb{R}^{n_o \times n_w \times n_w} \quad \text{mit} \quad [\mathcal{H}_k \mathbf{e}_k] := \sum_j^{n_0} e_{kj} \frac{\partial^2 e_{kj}}{\partial w_m \partial w_l} \quad (6.77)$$

³⁵Für umfassende Ausführungen zum QUASI-NEWTON-Verfahren für das KNN-Training, insbesondere im Kontext der Materialmodellierung, sei ausdrücklich auf WEBER [145] verwiesen.

³⁶Hier sollte noch beachtet werden, dass vor 30 Jahren die Rechenleistung noch ein entscheidenderer Faktor war. Die Aussage von LECUN ET AL. ist damit auch in den historischen Kontext einzuordnen.

³⁷Der Begriff Tensor wird hier gezielt vermieden, da Tensoren gewisse Transformationseigenschaften erfüllen.

eingeführt. Das GAUSS-NEWTON-Verfahren vernachlässigt dabei den letzten Summanden aus Gleichung (6.76) unter der Annahme, dass im Falle schwacher Nichtlinearitäten im Bereich der Minima der Einfluss der zweiten Ableitungen gering ist³⁸. Die Hesse-Matrix wird damit allein über die 1. Ableitungen approximiert. Weiter lässt sich durch Einführen des Fehlervektors für alle Datenpunkte

$$\mathbf{e} = [e_{11}, e_{12}, \dots, e_{1n_o}, \dots, e_{kj}, \dots, e_{P1}, \dots, e_{Pn_o}]^T \in \mathbb{R}^{(P \cdot n_o)} \quad (6.78)$$

die Summation über k in Gleichung (6.76) durch Einführen der Jacobi-Matrix

$$\mathbf{J} := \nabla_{\mathbf{w}} \mathbf{e} = \begin{bmatrix} \frac{\partial e_{11}}{\partial w_{10}^{(1)}} & \frac{\partial e_{11}}{\partial w_{20}^{(1)}} & \cdots & \frac{\partial e_{11}}{\partial w_{n_o n_{n_h}}^{(n_h+1)}} \\ \frac{\partial e_{12}}{\partial w_{10}^{(1)}} & \frac{\partial e_{12}}{\partial w_{20}^{(1)}} & \cdots & \frac{\partial e_{12}}{\partial w_{n_o n_{n_h}}^{(n_h+1)}} \\ \vdots & \vdots & \frac{\partial e_{kj}}{\partial w_{ml}^{(L)}} & \vdots \\ \frac{\partial e_{Pn_o}}{\partial w_{10}^{(1)}} & \frac{\partial e_{Pn_o}}{\partial w_{20}^{(1)}} & \cdots & \frac{\partial e_{Pn_o}}{\partial w_{n_o n_{n_h}}^{(n_h+1)}} \end{bmatrix} \in \mathbb{R}^{(P \cdot n_o) \times n_w} \quad (6.79)$$

kompakt zusammenfassen. Die Hesse-Matrix im Rahmen des GAUSS-NEWTON-Verfahrens lässt sich entsprechend mit Gleichung (6.76) über

$$\nabla_{\mathbf{w}}^2 \mathcal{F}(\mathbf{w}) \approx \frac{1}{P} \mathbf{J}^T \mathbf{J} \quad (6.80)$$

approximieren. Weiterhin gilt

$$\nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}) = \frac{1}{P} \sum_{k=1}^P \nabla_{\mathbf{w}} ||\mathbf{e}_k||^2 = \frac{1}{P} \sum_{k=1}^P \nabla_{\mathbf{w}} \mathbf{e}_k^T \mathbf{e}_k = \frac{1}{P} \mathbf{J}^T \mathbf{e}, \quad (6.81)$$

wodurch mit den entsprechenden Annahmen für die Aktualisierungsvorschrift (6.75) folgt³⁹

$$\mathbf{w}_{[n+1]} = \mathbf{w}_{[n]} - (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{e}. \quad (6.82)$$

Die Fehlerfläche beim Training eines KNN ist im Allgemeinen weder quadratisch noch konvex. Anschaulich lässt sie sich eher als mehrdimensionale zerklüftete Landschaft mit Bergen, Tälern und Hochebenen beschreiben. In Bereichen von

³⁸Dies entspricht der Annahme, dass die Fehlerfunktion eine lineare Funktion in den Gewichten ist (vgl. LECUN ET. AL [79]). Durch die Verwendung nichtlinearer Aktivierungsfunktionen ist dies hier offensichtlich nicht gegeben.

³⁹Aus Gleichung (6.82) wird ersichtlich, dass sich der Faktor $1/P$ kürzt. Häufig wird dahingehend in der Literatur (vgl. z.B. WILAMOWSKI UND IRWIN [151]) bei der Herleitung direkt von der Fehlersumme der Quadrate ausgegangen. Im Rahmen der Arbeit und der Verwendung des mittleren quadratischen Fehlers ist die hier aufgeführte Herleitung jedoch konsistenter.

Parameter	Typischer Wert	Bemerkung
μ	10^{-4}	Startwert, steuert numerische Dämpfung
β	10	Skaliert μ im Rahmen der Iteration
μ_{max}	1^{10}	Beschränkung von μ
t_{max}	10	Vermeidet Endlos-Schleife

Tabelle 6.6: Verwendete Wahl der Hyperparameter für den LMV-Algorithmus

lokalen Minima ist das GAUSS-NEWTON-Verfahren bezüglich der Konvergenzgeschwindigkeit dem einfachen Gradientenabstiegsverfahren überlegen. Im Bereich von Sattelpunkten kann es aufgrund von negativen Eigenwerten der Hesse-Matrix jedoch sogar dazu kommen, dass das Verfahren falsche Richtungen vorgibt (vgl. GOODFELLOW ET AL. [42]). Entsprechend erfolgt eine weitere Modifikation der Form

$$\mathbf{w}_{[n+1]} = \mathbf{w}_{[n]} - (\mathbf{J}^T \mathbf{J} + \mu \mathbf{I})^{-1} \mathbf{J}^T \mathbf{e} \quad \text{mit} \quad \mathbf{I} \in \mathbb{R}^{n_w \times n_w}. \quad (6.83)$$

Dies stellt die Aktualisierungsvorschrift für das LMV dar. Dabei werden auf die Hauptdiagonale der Hesse-Matrix numerische Dämpfungsterme μ addiert, die zumindest bei negativen Eigenwerten nahe Null das beschriebene Problem der falschen Abstiegsrichtung beheben. Dahingehend ist es sinnvoll, den Parameter μ adaptiv zu steuern. Steigt der Fehler in einem Aktualisierungsschritt, wird der Einfluss von μ gestärkt, andernfalls geschwächt. Die Skalierung erfolgt über einen weiteren Parameter β . Für große μ geht das LMV damit in ein Gradientenabstiegsverfahren mit Lernrate $1/\mu$ über. Die verwendete Wahl der eingeführten Parameter sowie weitere Größen, die im Rahmen der Implementierung benötigt werden, sind in Tabelle 6.6 gegeben. Die Größen t_{max} und μ_{max} dienen dabei als Abbruchkriterien, falls keine erfolgreiche Aktualisierung der Gewichte vorgenommen werden kann.

Ausstehend ist noch die Berechnung der Einträge der Jacobi-Matrix. Dafür wird in Anlehnung an die Vorgehensweise in (6.42) die Komponente

$$\frac{\partial e_{jk}}{\partial w_{ml}^{(L)}} = \frac{\partial e_{jk}}{\partial y_m^{(L)}} \frac{\partial y_m^{(L)}}{\partial s_m^{(L)}} \frac{\partial s_m^{(L)}}{\partial w_{ml}^{(L)}} \quad (6.84)$$

betrachtet. Analog zu den Überlegungen, die zu den Gleichungen (6.44)-(6.47) führten, wird die allgemeine Beziehung

$$\frac{\partial e_{jk}}{\partial w_{ml}^{(L)}} = \sum_{n=1}^{n_{L+1}} \delta_n^{(L+1)} w_{nm}^{(L+1)} g'(s_m^{(L)}) y_l^{(L-1)} \quad (6.85)$$

erhalten. Die Berechnung der Einträge der Jacobi-Matrix unterscheidet sich damit für die Komponenten bezüglich der Gewichte der Zwischenschichten nicht zu

denen des Gradienten. Lediglich in der Ausgangsschicht erfolgt eine Modifikation:

$$\frac{\partial e_{jk}}{\partial w_{jo}^{(n_h+1)}} = \frac{\partial e_{jk}}{\partial z_j} \frac{\partial z_j}{\partial s_j^{(n_h+1)}} \frac{\partial s_j^{(n_h+1)}}{\partial w_{jo}^{(n_h+1)}} = 1 \cdot 1 \cdot y_o^{(L-1)} = \bar{\delta}_{jk}^{(n_h+1)} y_o^{(L-1)}. \quad (6.86)$$

Der Delta-Wert der Ausgangsschicht $\bar{\delta}_{jk}^{(n_h+1)}$ ist in diesem Fall 1. Aus formalen Gründen wird er jedoch weiter mitgeführt. In Vektor-Matrix Notation können direkt die Einträge der Jacobi-Matrix der Zeile jk bezüglich der Gewichtskomponenten $\mathbf{w}^{(n_h+1)}$ berechnet werden:

$$\frac{\partial e_{jk}}{\partial \mathbf{w}^{(n_h+1)}} = \bar{\delta}_{jk}^{(n_h+1)} \mathbf{y}^{(n_h)T} \in \mathbb{R}^{1 \times n_h+1}. \quad (6.87)$$

Für die entsprechende Rückführung der Delta-Werte der Ausgangsschicht $\bar{\delta}_{jk}^{(n_h+1)}$ an die vorherige Schicht muss weiterhin die Gleichung (6.51) angepasst werden:

$$\boldsymbol{\delta}^{(n_h)} = \left(\bar{\delta}_{jk}^{(n_h+1)} \bar{\mathbf{w}}_{(j,:)}^{(n_h+1)} \right)^T \odot g'(\mathbf{s}^{(n_h)}) \in \mathbb{R}^{n_h}. \quad (6.88)$$

Die Index $(j, :)$ wird dabei zur Bezeichnung der j -ten Zeile verwendet. Die in Matrizenform vorliegenden Einträge der Jacobi-Matrix

$$\frac{\partial e_{jk}}{\partial \mathbf{w}^{(L)}} = \left[\left(\boldsymbol{\delta}^{(L+1)T} \bar{\mathbf{w}}^{(L+1)} \right)^T \odot g'(\mathbf{s}^{(L)}) \right] \mathbf{y}^{(L-1)T} \in \mathbb{R}^{n_L \times (n_{L-1}+1)}, L \neq n_h + 1 \quad (6.89)$$

der Zwischenschichten müssen dabei entsprechend in die Zeilen einsortiert werden. Alle beschriebenen Schritte sind als Implementierungsvorschrift in Algorithmus 6.7 zusammengefasst.

6.2.4 Beispiel zu Funktionsapproximation

Im Folgenden werden die Approximationseigenschaften von KNNs, welche auf einem identischen Datensatz mit dem ADAM- und dem LEVENBERG-MARQUARDT-Algorithmus trainiert wurden, demonstriert. Dabei soll auch das Einfluss der entsprechenden Trainings-Algorithmen diskutiert werden. Dabei dient die sogenannte Sinc-Funktion

$$\tilde{f}(\mathbf{x}) = \frac{\sin(r)}{r}, \quad r = \sqrt{x^2 + y^2} \quad \text{mit} \quad [x, y]^T \in \left[-\frac{5}{2}\pi, \frac{5}{2}\pi\right]^2 \quad (6.90)$$

als zu approximierende Referenzlösung, vgl. auch Gleichung (6.2). Aufgrund des oszillativen Charakters wird diese gerne als Maßstab (engl.: *Benchmark*) verwendet, s. z. B. HAGAN und MENHAJ [48]. Die Referenzlösung wird genutzt, um 100 zufällige Datenpunkte mit der *Latin Hypercube Sampling* Methode (LHS) (vgl. z.B. TANG [127]) zu erzeugen. Diese dienen als Datenpaare $(\mathbf{x}_k, \mathbf{z}_k)$ für das Training des KNNs. Die Funktion und die verwendeten Daten sind in Abbildung 6.3

```

Initialisierung:  $\mu, \beta, \mu_{max}, t_{max}$ , setze  $n = 0$ 
Training: while (kein Abbruch) do
     $n = n + 1$ 
Berechne (Alg. 6.1):  $\hat{\mathbf{x}} \mapsto \hat{\mathbf{z}}$  für alle Daten  $\rightarrow \mathbf{e}$  (Fehlervektor)
Aktueller Fehler:  $\mathcal{F}(\mathbf{w})_{[n]}$ , Initialisiere  $\mathcal{F}(\mathbf{w})_{[n+1]} = \mathcal{F}(\mathbf{w})_{[n]}$ 
Daten-Schleife: for  $k = 1, \dots, P^{\text{Train}}$ 
Ausgangs-Schleife: for  $j = 1, \dots, n_o$ 
     $\bar{\delta}_{jk}^{(n_h+1)} = 1 \cdot g'(s_j^{(n_h+1)})$ 
     $\nabla_{\mathbf{w}^{(n_h+1)}} e_{jk} = \bar{\delta}_{jk}^{(n_h+1)} \mathbf{y}^{(n_h)T}$ 
     $\boldsymbol{\delta}^{(n_h)} = \left( \bar{\delta}_{jk}^{(n_h+1)} \bar{\mathbf{w}}_{(j,:)}^{(n_h+1)} \right)^T \odot g'(\mathbf{s}^{(n_h)})$ 
     $\nabla_{\mathbf{w}^{(n_h)}} e_{jk} = \boldsymbol{\delta}^{(n_h)} \mathbf{y}^{(n_h-1)T}$ 
Restliche Schichten: for  $L = n_h - 1, \dots, 1$ 
     $\boldsymbol{\delta}^{(L)} = \left( \boldsymbol{\delta}^{(L+1)T} \bar{\mathbf{w}}^{(L+1)} \right)^T \odot g'(\mathbf{s}^{(L)})$ 
     $\nabla_{\mathbf{w}^{(L)}} e_{jk} = \boldsymbol{\delta}^{(L)} \mathbf{y}^{(L-1)T}$ 
    end
end
end
Sortierung:  $\nabla_{\mathbf{w}^{(L)}} e_{jk} \rightarrow \mathbf{J}, L = 1, \dots, n_h + 1$ 
Berechne:  $\nabla_{\mathbf{w}} \mathcal{F} = \mathbf{J}^T \mathbf{e}, \nabla_{\mathbf{w}}^2 \mathcal{F} = \mathbf{J}^T \mathbf{J}$ 
Initialisierung:  $t = 0$ 
Starte LMV: while  $\mathcal{F}(\mathbf{w})_{[n+1]} > \mathcal{F}(\mathbf{w})_{[n]}$  and  $t \leq t_{max}$ 
     $t = t + 1$ 
Aktualisierung:  $\mathbf{w}_{[t]} = \mathbf{w}_{[n]} - (\mathbf{J}^T \mathbf{J} + \mu \mathbf{I})^{-1} \mathbf{J}^T \mathbf{e}$ 
Neuer Fehler:  $\hat{\mathbf{x}} \mapsto \hat{\mathbf{z}} \rightarrow \mathcal{F}(\mathbf{w})_{[t]}$ 
    if  $\mathcal{F}(\mathbf{w})_{[t]} < \mathcal{F}(\mathbf{w})_{[n]}$ 
         $\mu = \mu / \beta$ 
         $\mathbf{w}_{[n+1]} = \mathbf{w}_{[t]}, \mathcal{F}(\mathbf{w})_{[n+1]} = \mathcal{F}(\mathbf{w})_{[t]}$ 
    else
         $\mu = \mu \cdot \beta$ 
    end
end

```

Algorithmus 6.7: Berechnungsvorschrift für den LMV-Algorithmus

dargestellt. Zusätzlich werden zehn weitere Datenpunkte für die Berechnung des Testfehlers erzeugt. Grundsätzlich sei hierbei angemerkt, dass bei einer vorhandenen Referenzlösung unendliche viele Daten zur Verfügung stehen. Im Rahmen der Arbeit wird jedoch das Ziel „so viel wie nötig, so wenig wie möglich“ verfolgt, um später möglichst effizient KNN-Materialmodelle zu konstruieren.

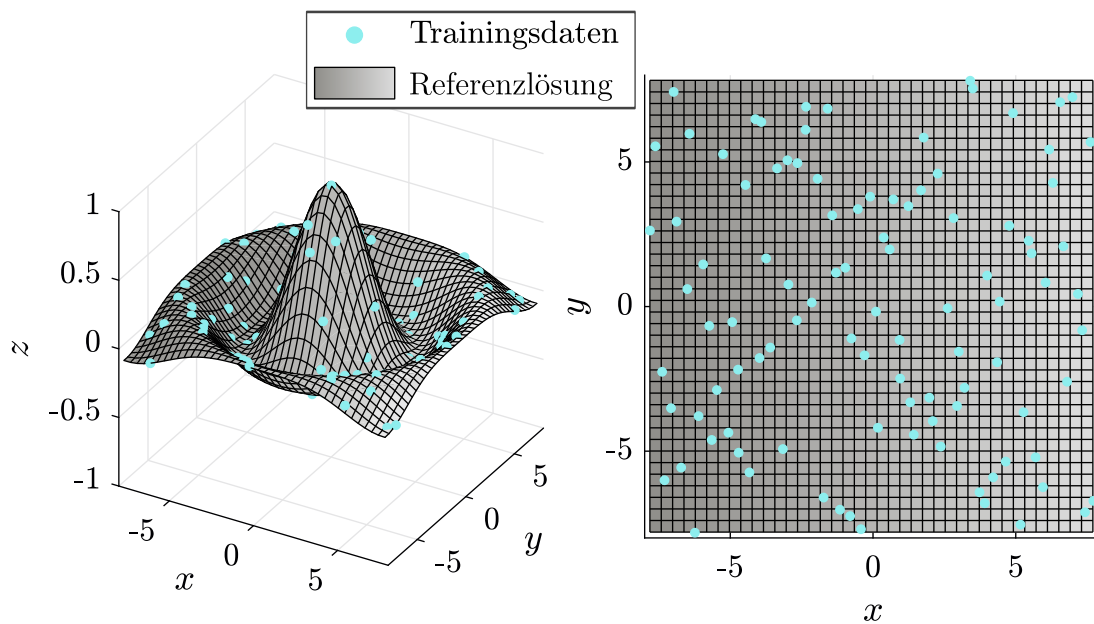


Abbildung 6.3: Perspektivische Ansicht (links) und Draufsicht (rechts) der Sinc-Funktion sowie der mit LHS generierten Trainingsdaten

Die für die Initialisierung des ADAM und des LMV verwendeten Parameter sind in den Tabellen 6.5 und 6.6 gegeben. Weiterhin wurden $n_B = 10$ Minibatches für den ADAM-Algorithmus verwendet. Mit beiden Verfahren wurden jeweils zehn KNN auf dem gleichen Datensatz trainiert, womit sich die Trainingsdurchläufe nur in den initialisierten Gewichten unterscheiden. Das KNN definiert die Abbildungsvorschrift $f : \mathbb{R}^2 \mapsto \mathbb{R}$ und besitzt somit $n_i = 2$ Eingangs- und $n_o = 1$ Ausgangsneurone. Es wird eine verdeckte Schicht mit 40 Neuronen verwendet, die sich im Rahmen von Vorstudien für beide Trainings-Verfahren als geeignet erwiesen hat. Damit lautet die KNN-Architektur $2 - 40 - 1$, die in $n_w = 161$ Gewichten resultiert.

Die Verläufe des mittleren quadratischen Testfehlers $\mathcal{F}^{\text{Test}}(\mathbf{w})$ über die Trainingsepochen ist für beide Verfahren und alle zehn Trainingsdurchläufe in Abbildung 6.4 dargestellt. Für beide Verfahren ist unabhängig von der Initialisierung der Gewichte eine Konvergenzverhalten zu Fehlermaßen in vergleichbaren Größenordnungen erkennbar. Beim LMV stellt sich diese Konvergenz bereits nach ≈ 100 Epochen ein, während der ADAM ≈ 5000 Epochen benötigt, woran sich gut der Unterschied zwischen Minibatch- und Batch-basierten Verfahren erkennen lässt. Im Rahmen der Arbeit wird in der Regel jedoch nicht allein der Testfehler als Qualitätsmerkmal herangezogen, da er lediglich einen Erwartungswert bezüglich eines kleinen Stichprobenumfangs darstellt. Ein kleiner Testfehler kann somit immer mit lokalen Ausreißern in der Approximation einhergehen, die im Falle eines KNN-Materialmodells unerwünscht sind. Dahingehend ist in Abbildung 6.4 auch

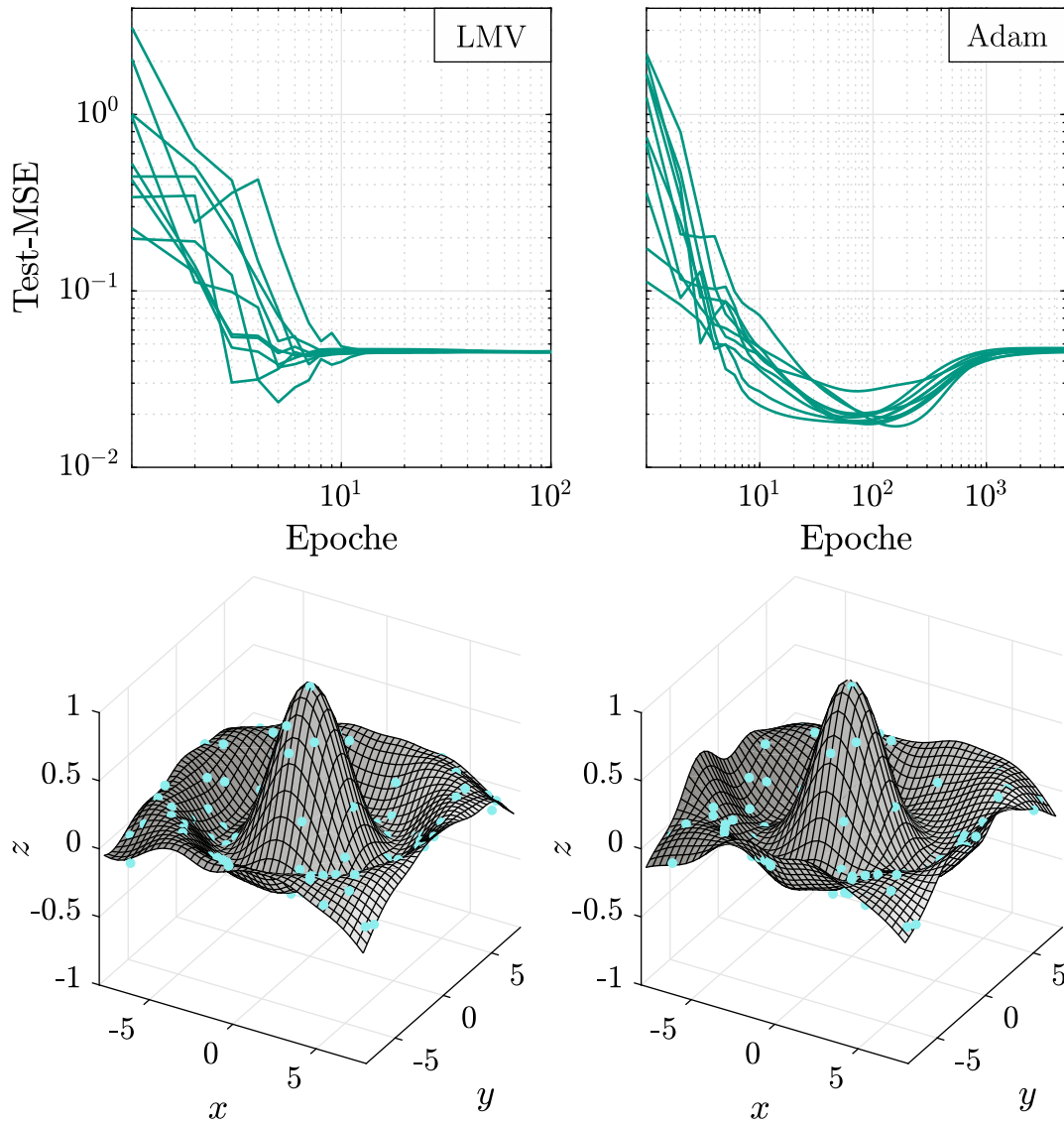


Abbildung 6.4: Verlauf des Test-MSE über die Trainingsepochen für das LMV und den ADAM und zehn Trainingsdurchläufe (oben) sowie Darstellung der Approximationen der Sinc-Funktion (unten) mit 100 LHS generierten Trainingsdaten

die Approximation der Sinc-Funktion für exemplarische KNNs nach Erreichen der Konvergenz dargestellt⁴⁰. Unabhängig von dem Trainingsdurchlauf konnten dabei rein visuell ähnliche Realisationen festgestellt werden, womit die nachfolgenden Aussagen als allgemeingültig aufgefasst werden können. In beiden Fällen

⁴⁰Offensichtlich entspricht der auskonvergierte Testfehler bei beiden Verfahren nicht zwingend dem kleinsten Fehler über die Anzahl an Epochen. Rein subjektiv sind jedoch in beiden Fällen bessere Approximationen der Sinc-Funktion im auskonvergierten Bereich erzielt worden. Dies unterstreicht die Aussage, den Erfolg des KNN-Trainings nicht allein auf Basis des Testfehlers zu beurteilen.

wird das qualitative Verhalten der Referenzfunktion erfasst. Gerade am Rande des Trainingsraums lässt sich jedoch erkennen, dass mit dem LMV bessere Approximationen erzielt werden, weshalb es das bevorzugte Trainingsverfahren im Rahmen der Arbeit darstellt.

Der Umstand, dass im vorliegenden Fall eine Referenzlösung existiert und somit beliebig viele Daten generiert werden könnten, stellt grundsätzlich einen wesentlichen Vorteil gegenüber Szenarien mit begrenzten Datensätzen dar. Dieser Weg wird hier jedoch bewusst nicht verfolgt. Stattdessen werden zusätzliche Informationen, etwa über die KNN-Tangente, in das Training integriert. Dies soll nicht nur den erforderlichen Datenumfang reduzieren, sondern zugleich die Genauigkeit der für die numerische Anwendung relevanten Größen verbessern.

6.2.5 Abschließende Anmerkungen zum KNN-Training

Das KNN-Training zeichnet sich im Allgemeinen durch eine Reihe weiterer Herausforderung aus, die hier aus Gründen der Vollständigkeit kurz erläutert werden. Grundsätzlich lassen sich zwei Problematiken im Rahmen der Approximation identifizieren: Unteranpassung (engl. *Underfitting*) und Überanpassung (engl. *Overfitting*). Unteranpassung geht mit einem hohen Training- und Testfehler einher, d. h., das KNN ist aufgrund fehlender Daten und/oder ungeeigneter Architektur nicht in der Lage, den zugrundeliegenden funktionalen Zusammenhang zufriedenstellend zu approximieren. Überanpassung wird auch als Auswendiglernen bezeichnet, da der Fehler bezüglich der Trainingsdaten kontinuierlich sinkt, der Testfehler jedoch steigt. Die Fähigkeit zu generalisieren ist entsprechend nicht gegeben.

Bei der Überanpassung stellt sich also die Frage, wann der geeignete Zeitpunkt zum Beenden des Trainings erreicht ist (vgl. hierzu auch „*Early Stopping - But When?*“, PRECHELT [107]). Die Problematik lässt sich auch am Verlauf des Testfehlers bei der Verwendung des ADAM-Algorithmus in Abbildung 6.4 erkennen. Der Minimalwert wird nach ≈ 100 Epochen erreicht, wonach ein Anstieg bis zum Erreichen der Konvergenz erkennbar ist. Im Allgemeinen kann der Verlauf des Testfehlers deutlich komplexer sein und aus alternierenden lokalen Minima und Maxima bestehen. Einige Strategien, um dabei den geeigneten Zeitpunkt zur Beendigung des Trainings zu ermitteln sind in PRECHELT [107] gegeben.

Beim LMV ist in der Regel ein anderes Verhalten zu beobachten. Wie in Abbildung 6.4 zu sehen ist, lassen sich nur in einigen wenigen KNN-Trainings lokale Minima des Testfehlers vor Erreichen der Konvergenz feststellen. Eine Strategie zur frühzeitigen Beendigung des Trainings ist dahingehend – unter der Voraussetzung eines gut gestellten Problems – nicht notwendig. Stattdessen wird ein Konvergenzkriterium definiert, wonach das Training abgebrochen wird. Dafür wird

zunächst die Änderung des Testfehlers

$$\Delta \mathcal{F}^{\text{Test}}(\mathbf{w})_{[n]} = |\mathcal{F}^{\text{Test}}(\mathbf{w})_{[n+1]} - \mathcal{F}^{\text{Test}}(\mathbf{w})_{[n]}| \quad (6.91)$$

berechnet. Damit wird das Konvergenzkriterium

$$\delta \mathcal{F}(\mathbf{w}) = \frac{1}{N} \sum_{n=N}^n \Delta \mathcal{F}^{\text{Test}}(\mathbf{w})_{[n]} \leq \text{TOL} \quad (6.92)$$

definiert, womit $\delta \mathcal{F}(\mathbf{w})$ die mittlere Änderung des Testfehler über die letzten N Epochen darstellt. Als geeignet hat sich hier $N = 10$ und eine Toleranz von $\text{TOL} = 1 \cdot 10^{-6}$ erwiesen. Weiterhin werden eine Mindestanzahl $\text{epo}_{\min} = 100$ und Maximalanzahl von Epochen $\text{epo}_{\max} = 1000$ vorgegeben, um sowohl einen frühzeitigen Trainingsabbruch als auch unnötiges Training ohne signifikante Verbesserung zu vermeiden.

6.3 KNN-Training mit Nebenbedingungen

Für das bisherige KNN-Training wurde angenommen, dass die gesuchten Gewichte $\tilde{\mathbf{w}}$ – vgl. Gleichung (6.2) – Elemente der Menge aller möglichen Gewichte $\mathcal{W}$ sind. Durch die Einführung von Nebenbedingungen kann diese Menge auf eine echte Teilmenge $\mathcal{W} \subseteq \mathcal{W}$ eingeschränkt werden. Diese Einschränkung wird als restringierte Optimierung bezeichnet. Nebenbedingungen formulieren ganz allgemein zusätzliche Optimalitätskriterien an ein mögliches Minimum der Fehlerfläche. So ist es beispielsweise wünschenswert, dass die Gewichte beim KNN-Training klein bleiben, um zum einen außerhalb der Sättigungsbereiche der Aktivierungsfunktion zu bleiben und zum anderen dem Problem der Überanpassung entgegenzuwirken⁴¹. Praktisch kann dies dadurch erreicht werden, dass der Gewichtsvektor im Rahmen der Aktualisierung – hier unter Verwendung des Gradientenabstiegsverfahrens, s. Gleichung (6.68) – mit einem Faktor $(1 - \rho \eta) \in (0, 1)$ multiplikativ skaliert wird:

$$\mathbf{w}_{[n+1]} = (1 - \rho \eta) \mathbf{w}_{[n]} - \eta \nabla_{\mathbf{w}} \mathcal{F}(\mathbf{w}). \quad (6.93)$$

Ausgehend von dieser Aktualisierung lässt sich eine Fehlerfunktion konstruieren⁴², deren Gradient exakt zu (6.93) führt:

$$\mathcal{F}^{L2}(\mathbf{w}) = \frac{1}{2P} \sum_{k=1}^P \left(\hat{\mathbf{z}}_k(\hat{\mathbf{x}}_k, \mathbf{w}) - \hat{\mathbf{t}}_k \right)^2 + \frac{\rho}{2} \|\mathbf{w}\|^2 = \mathcal{F}^D(\mathbf{w}) + \frac{\rho}{2} \|\mathbf{w}\|^2. \quad (6.94)$$

⁴¹Kleine Werte für $\mathbf{w}$ führen zu einer geringeren Sensitivität der KNN-Antwort, was sich durch kleinere Krümmungen äußert.

⁴²Die hier dargestellte Vorgehensweise der Konstruktion des Fehlers ausgehend von dessen Gradienten ist rein didaktischer Natur im Rahmen der Einführung. Üblicherweise wird *erst* die Fehlerfunktion definiert.

Dies entspricht der bekannten L2-Regularisierung, bei der die quadrierte Norm des Gewichtsvektors als Strafterm zur Fehlerfunktion bezüglich der Trainingsdaten $\mathcal{F}^D(\mathbf{w})$ addiert wird. Das Optimierungsproblem lautet dann:

$$\tilde{\mathbf{w}} = \arg \min \left(\mathcal{F}^D(\mathbf{w}) + \frac{\rho}{2} \|\mathbf{w}\|^2 \right). \quad (6.95)$$

Der Parameter ρ steuert, wie stark der zusätzliche Term in die Optimierung eingeht. Dieses Vorgehen entspricht der allgemeinen Struktur des KNN-Trainings mit Nebenbedingungen. In diesem Beispiel wurde sie zur Regularisierung eingesetzt, wobei Regularisierung das Verringern des Testfehlers ohne einhergehende Erhöhung des Trainingsfehlers bezeichnet (vgl. GOODFELLOW ET AL. [42]). Die Nebenbedingung wird hierbei nur *schwach* berücksichtigt, d. h. die Gewichte dürfen prinzipiell beliebige Werte annehmen und stellen immer eine zulässige Lösung des Problems (6.95) dar⁴³.

Eine weitere Möglichkeit ist die Formulierung von Nebenbedingungen, um zusätzliche Informationen über den zu approximierenden funktionalen Zusammenhang in das KNN-Training zu integrieren. Diese Informationen können rein physikalisch motiviert sein oder direkt aus dem zugrundeliegenden Datensatz extrahiert werden, z. B., falls die Tangente des Problems bekannt ist. Letzteres wird als SOBOLEV-Training bezeichnet und stellt den Fokus dieses Kapitels dar. Im Folgenden wird damit kurz der theoretische Rahmen zur Berücksichtigung von Nebenbedingungen bei der nichtlinearen Optimierung gegeben, um anschließend konkret die theoretische und algorithmische Behandlung der Tangenten-Nebenbedingung zu diskutieren. Ein knapper theoretischer Hintergrund zur Formulierung restringierter Optimierungsprobleme ist in GOODFELLOW ET AL. [42] zu finden, für Spezialliteratur wird auf NOCEDAL und WRIGHT [100] und GEIGER und KANZOW [33] verwiesen. Für weitere Ausführungen zur Behandlung von Nebenbedingung im Kontext des KNN-Trainings siehe WEBER [145].

6.3.1 Theorie nichtlinearer restringierter Optimierungsprobleme

Für die Beschreibung der echten Teilmenge $\bar{\mathcal{W}} \subseteq \mathcal{W}$ des restringierten Optimierungsproblems sind geeignete Bedingungen zu formulieren, sodass

$$\bar{\mathcal{W}} = \{\mathbf{w} \mid h_i(\mathbf{w}) = 0, i = 1, \dots, n_{GL} \text{ und } g_j(\mathbf{w}) \leq 0, j = 1, \dots, n_{UG}\}. \quad (6.96)$$

⁴³Eine *exakte* Forderung kleiner Gewichte wäre keine reguläre Nebenbedingung im mathematischen Sinn. Ein Beispiel für eine strikte Gleichheitsnebenbedingung wäre z.B. $\|\mathbf{w}\| - 1 = 0$, womit die zulässigen Gewichte auf die Oberfläche einer n_w -dimensionalen Kugel eingeschränkt würden. Im Falle der L2-Regularisierung kann lediglich die Region der Teilmenge $\bar{\mathcal{W}}$ über den Parameter ρ skaliert werden.

Die Nebenbedingungen werden entsprechend durch die Gleichungen $h_i(\mathbf{w}) = 0$ und die Ungleichungen⁴⁴ $g_j(\mathbf{w}) \leq 0$ beschrieben. Das dem KNN-Training zugrundeliegende restringierte Optimierungsproblem lautet somit

$$\tilde{\mathbf{w}} = \arg \min_{\mathbf{w} \in \mathcal{W}} \mathcal{F}^D(\mathbf{w}) \quad \text{u. d. N} \quad \begin{aligned} h_i(\mathbf{w}) &= 0, \quad i = 1, \dots, n_{GL} \\ g_j(\mathbf{w}) &\leq 0, \quad j = 1, \dots, n_{UG}, \end{aligned} \quad (6.97)$$

mit der Abkürzung u. d. N für *unter der Nebenbedingungen*. Um Lösungsstrategien für unrestringierte Probleme anwenden zu können, werden die Gleichungen (6.97) über die LAGRANGE-Funktion⁴⁵

$$\mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathcal{F}^D(\mathbf{w}) + \boldsymbol{\lambda}^T \mathbf{h}(\mathbf{w}) + \boldsymbol{\mu}^T \mathbf{g}(\mathbf{w}) \quad (6.98)$$

formuliert. Dabei sind $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_{n_{GL}}]^T$ und $\boldsymbol{\mu} = [\mu_1, \dots, \mu_{n_{UG}}]^T$ die den Komponenten von $\mathbf{h} = [h_1, \dots, h_{n_{GL}}]^T$ und $\mathbf{g} = [g_1, \dots, g_{n_{UG}}]^T$ zugeordneten LAGRANGE-Multiplikatoren, die als zusätzliche Freiheitsgrade in das Optimierungsproblem eingehen. Weiter lässt sich der Gradient der LAGRANGE-Funktion

$$\nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \nabla_{\mathbf{w}} \mathcal{F}^D(\mathbf{w}) + \nabla_{\mathbf{w}} \mathbf{h}(\mathbf{w})^T \boldsymbol{\lambda} + \nabla_{\mathbf{w}} \mathbf{g}(\mathbf{w})^T \boldsymbol{\mu} \quad (6.99)$$

bezüglich der Gewichte definieren. Falls eine lokale Lösung $\tilde{\mathbf{w}}$ für das Problem (6.97) existiert, dann müssen an den Punkten $(\tilde{\mathbf{w}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\mu}})$ die KARUSH-KUHN-TUCKER-Bedingungen (KKT-Bedingungen) [74]

$$\begin{aligned} \nabla_{\mathbf{w}} \mathcal{L}(\tilde{\mathbf{w}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\mu}}) &= \mathbf{0}, \\ \mathbf{g}(\tilde{\mathbf{w}}) &= \mathbf{0} \quad \text{und} \\ \tilde{\boldsymbol{\lambda}} &\geq \mathbf{0}, \quad \mathbf{h}(\tilde{\mathbf{w}}) \leq \mathbf{0}, \quad \tilde{\boldsymbol{\lambda}}^T \mathbf{h}(\tilde{\mathbf{w}}) = 0 \end{aligned} \quad (6.100)$$

erfüllt sein. Diese werden auch als Optimalitätskriterien erster Ordnung bezeichnet, da sie lediglich Anforderungen an die Gradienten stellen. Ohne die Formulierung von Nebenbedingungen ist anhand der ersten Gleichung von (6.100) wieder die Form der unrestringierten Minimierungsaufgabe zu erkennen, s. Gleichung (6.73). Für die Diskussion der Existenz eines KKT-Punktes $(\tilde{\mathbf{w}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\mu}})$, wird auf die zuvor genannte Literatur verwiesen. Angemerkt sei, dass ein KKT-Punkt nur ein notwendiges und kein hinreichendes Kriterium für die Beschreibung eines optimalen Punktes des restringierten Minimierungsproblems (6.97) darstellt.

Die direkte Konstruktion einer LAGRANGE-Funktion gemäß Gleichung (6.99) und die anschließende Suche nach einem Optimum ist im Allgemeinen keine geeignete

⁴⁴Nebenbedingungen in Form von Ungleichungen werden im Rahmen dieser Arbeit nicht betrachtet. Aus Gründen der Vollständigkeit sollen sie jedoch hier nicht unerwähnt bleiben.

⁴⁵Siehe hierzu auch Gleichung (2.79), bei der durch die HU-WASHIZU-Formulierung das Gleichgewicht unter der Nebenbedingung der Einhaltung der Kinematik und des Stoffgesetzes eingefordert wird. Die Struktur findet sich auch bei der Einforderung der Randbedingungen in (5.19) auf dem Schalen-RVE wieder.

Grundlage für das KNN-Training. Zum einen ist die Fehlerfläche nicht konvex, sodass die Existenz einer exakten Lösung des restringierten Problems nicht garantiert werden kann. Zum anderen führen die LAGRANGE-Multiplikatoren zusätzliche Freiheitsgrade ein, die – analog zu den Gewichten – mit den in Kapitel 6.2 beschriebenen Verfahren iterativ bestimmt werden müssten. Zielführender ist es daher, die Nebenbedingungen nur näherungsweise zu berücksichtigen. Dies geschieht im Rahmen sogenannter Penalty-Verfahren, deren Struktur sich an der bereits eingeführten Funktion $\mathcal{F}^{L2}(\mathbf{w})$ erkennen lässt.

Im Folgenden wird die Behandlung von Gleichheitsnebenbedingungen mit dem Penalty-Verfahren betrachtet. Die Grundidee besteht darin, das beschränkte Problem (6.97) in eine Folge unbeschränkter Probleme zu überführen. Zu diesem Zweck wird die quadratische Penalty-Funktion eingeführt:

$$\mathcal{F}^P(\mathbf{w}) = \mathcal{F}^D(\mathbf{w}) + \frac{\rho}{2} \sum_{i=1}^{n_{\text{GL}}} h_i(\mathbf{w})^2, \quad (6.101)$$

wobei der bereits aus der L2-Regularisierung bekannte Faktor ρ nun als Penalty-Parameter bezeichnet wird. Auf der zulässigen Menge, d. h. dort, wo die Nebenbedingungen erfüllt sind, stimmt die Penalty-Funktion mit der Zielfunktion $\mathcal{F}^D(\mathbf{w})$ des restringierten Problems (6.97) überein. Außerhalb dieses Bereichs wird die Penalty-Funktion durch den Strafterm erhöht, sodass der Minimierungsalgorithmus wieder in Richtung der zulässigen Menge „zurückgedrängt“ wird. Theoretisch erfolgt dies durch die sequenzielle Lösung einer Folge unrestringierter Probleme, bei denen für die Penalty-Parameter $\{\rho_k\}$ mit $k \rightarrow \infty$ auch $\rho_k \rightarrow \infty$ gilt. Im besten Fall konvergiert diese Folge gegen die Lösung des ursprünglichen Problems (6.97). In der Praxis wird ρ jedoch oft konstant gewählt, wenn eine Approximation der Nebenbedingungen ausreichend ist.

Anschaulich lässt der Einfluss des Penalty-Parameters an dem Minimierungsproblem

$$\min x + y \quad \text{u.d.N} \quad h(x, y) = x^2 + y^2 - 2 = 0, \quad (6.102)$$

in Anlehnung an NOCEDAL und WRIGHT [100], zeigen. Die zugehörige Penalty-Funktion lautet

$$F^P(x, y) = x + y + \frac{\rho}{2}(x^2 + y^2 - 2)^2. \quad (6.103)$$

Die Nebenbedingung wird für $(x, y) = (-1, -1)$ exakt erfüllt. In Abbildung 6.5 ist der Graph der Funktion (6.103) für zwei Penalty-Parameter $\rho \in [1; 10]$ dargestellt. Die „Wanne“ der Minima ist für $\rho = 1$ erkennbar flacher und somit führt die Verletzung der Nebenbedingung zu kleineren Werten von $\mathcal{F}^P(x, y)$ im Vergleich zu $\rho = 10$. Im letzteren Fall ist dahingehend das Minimum von $\mathcal{F}^P(x, y)$ näher an der *exakten* Lösung $(x, y) = (-1, -1)$.

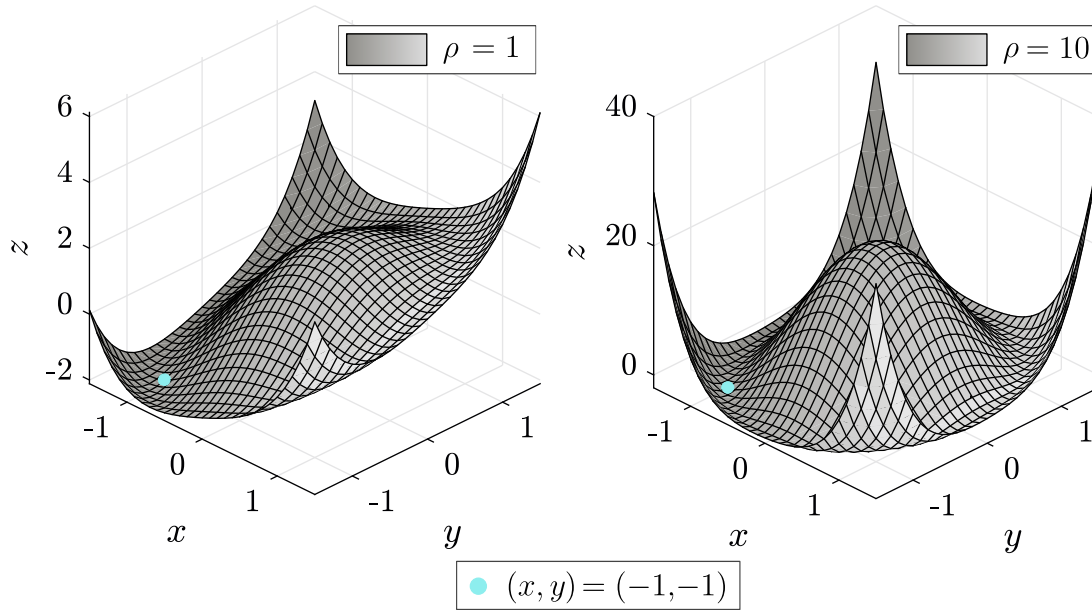


Abbildung 6.5: Darstellung der zu minimierenden Funktion $F^P(x, y)$ aus Gleichung (6.103) für zwei unterschiedliche Penalty-Faktoren und des exakten Minimums unter der Nebenbedingung

Ein weiterer Vorteil der hier verwendeten Penalty-Funktion ist der quadratische Charakter. Damit lässt sich das in Kapitel 6.2.3 vorgestellte LEVENBERG-MARQUARDT-Verfahren direkt für das erweiterte KNN-Training unter Nebenbedingungen anwenden.

6.3.2 Sobolev-Training: Die Tangenten-Nebenbedingung

Das vorgestellte theoretische Korsett zur Behandlung von Nebenbedingungen im Rahmen nichtlinearer Optimierungsprobleme dient im Weiteren zur Einbindung von KNN-Tangenteninformationen in das Training. Dies ist möglich, falls eine Menge an Tangenten-Daten

$$\mathcal{D}^S = \left\{ \left(\mathbf{x}_k^S, \frac{\partial t_{kj}}{\partial x_{ki}^S} \right) \mid \mathbf{x}_k^S \in \mathcal{X}^S \subset \mathbb{R}^{n_i}, k = 1, 2, \dots, P^S \right\} \quad \text{mit} \quad \begin{cases} j = 1, \dots, J \\ i = 1, \dots, I \end{cases} \quad (6.104)$$

vorliegt. Die Einbindung dieser Daten wird auch als SOBOLEV-Training bezeichnet, vgl. CZARNECKI ET AL. [26]. Die im Datensatz $\mathcal{D}^S$ enthaltenen partiellen Ableitungen müssen dabei explizit nicht für alle Ausgangsvariablen nach allen Eingangsvariablen vorliegen, weshalb im Allgemeinen die Unterscheidung $J \neq n_o$ und $I \neq n_i$ vorzunehmen ist. Dies wird im Rahmen der Verwendung des KNNs als Materialmodell in Kapitel 7 deutlich.

Offensichtlich ist ein SOBOLEV-Training nur in Sonderfällen möglich, nämlich dann, wenn eine differenzierbare Referenzlösung $\tilde{f}(\mathbf{x})$ vorliegt, aus der die Tangenteninformationen extrahiert werden können. Dies ist z.B. in der klassischen Materialmodellierung, bei der ein Modell auf Basis von experimentellen Daten konstruiert werden muss, nicht gegeben. Die Formulierung von Nebenbedingungen muss hier über physikalisch motivierte Restriktionen erfolgen. Diese können beispielsweise die Symmetrie der Materialtangente sein (vgl. WEBER ET AL. [146]) oder die Forderung nach materieller Stabilität (vgl. WEBER ET AL. [147]). Im Rahmen der Arbeit werden jedoch synthetische Daten auf einem RVE erzeugt, womit neben den Ausgangsvariablen – den Spannungen, vgl. Gleichung (5.35) – auch die Materialtangente aus Gleichung (5.36) bekannt ist. Falls partielle Ableitungen der KNN-Tangente bekannt sind, lautet das zu lösende Optimierungsproblem

$$\tilde{\mathbf{w}} = \arg \min_{\mathbf{w} \in \mathcal{W}} \mathcal{F}^D(\mathbf{w}) \quad \text{u.d.N} \quad h_{kji}(\mathbf{w}) = 0, \quad k = 1, \dots, P^S \quad (6.105)$$

mit der zugehörigen Nebenbedingung

$$h_{kji} = \frac{(t_{j,\max} - t_{j,\min})}{(x_{i,\max} - x_{i,\min})} \frac{\partial \hat{z}_j(\mathbf{x}_k^S; \mathbf{w})}{\partial \hat{x}_i} - \frac{\partial t_j(\mathbf{x}_k^S)}{\partial x_i} = 0 \quad \text{und} \quad \begin{cases} j = 1, \dots, J \\ i = 1, \dots, I \end{cases}, \quad (6.106)$$

wobei die Transformationsbeziehung (6.52) angewendet wurde. Mit der Beziehung

$$\frac{\partial \hat{t}_j(\mathbf{x}_k^S)}{\partial \hat{x}_i} = \frac{(x_{i,\max} - x_{i,\min})}{(t_{j,\max} - t_{j,\min})} \frac{\partial t_j(\mathbf{x}_k^S)}{\partial x_i} \quad (6.107)$$

lässt sich die Nebenbedingung auch direkt bezüglich der normalisierten partiellen Ableitungen der Zielgrößen formulieren:

$$h_{kji} = \frac{\partial \hat{z}_j(\mathbf{x}_k^S; \mathbf{w})}{\partial \hat{x}_i} - \frac{\partial \hat{t}_j(\mathbf{x}_k^S)}{\partial \hat{x}_i}. \quad (6.108)$$

Diese zunächst unscheinbar wirkende Umformung der ursprünglichen Formulierung von (6.106) hat im Rahmen der Optimierung den entscheidenden Vorteil, dass durch die Einforderung der Nebenbedingung im normalisierten Trainingsraum a priori gewährleistet ist, dass sich die Fehlerterme bezüglich der Daten und der partiellen Ableitungen in ähnlichen Größenordnungen befinden. Dies ist beispielsweise nicht gegeben, wenn die Symmetrie der KNN-Tangente eingefordert wird und sich die für die Transformation benötigten Vorfaktoren nicht kürzen lassen, vgl. hierzu WEBER ET AL. [146]. In diesem Fall empfiehlt es sich, die Nebenbedingung durch Einführung zusätzlich Faktoren zu skalieren⁴⁶. Die zu dem restringierten Problem (6.105) zugehörige Penalty-Funktion lautet damit

$$\mathcal{F}^P(\mathbf{w}) = \mathcal{F}^D(\mathbf{w}) + \mathcal{F}^S(\mathbf{w}), \quad (6.109)$$

⁴⁶Theoretisch kann dies auch direkt über den Penalty-Parameter erfolgen, was allerdings dessen Wahl zusätzlich erschweren würde. Es empfiehlt sich daher, die Nebenbedingung erst zu skalieren und anschließend den Einfluss des Penalty-Parameters zu untersuchen.

wobei der SOBOLEV-Fehlerterm

$$\mathcal{F}^S(\mathbf{w}) := \frac{\rho}{2P^S} \sum_{k=1}^{P^S} \sum_{j=1}^J \sum_{i=1}^I \left(\frac{\partial \hat{z}_j(\mathbf{x}_k^S; \mathbf{w})}{\partial \hat{x}_i} - \frac{\partial \hat{t}_j(\mathbf{x}_k^S)}{\partial \hat{x}_i} \right)^2 = \frac{\rho}{2P^S} \sum_{k=1}^{P^S} \sum_{j=1}^J \sum_{i=1}^I h_{kji}^2 \quad (6.110)$$

definiert wurde. Damit lassen sich die in Kapitel 6.2 diskutierten Verfahren zur Minimierung des Fehlers (6.109) anwenden. Hierfür wird entweder der Gradient

$$\nabla_{\mathbf{w}} \mathcal{F}^P(\mathbf{w}) = \nabla \mathcal{F}^D(\mathbf{w}) + \nabla \mathcal{F}^S(\mathbf{w}) \in \mathbb{R}^{n_w} \quad (6.111)$$

oder die Approximation der Hesse-Matrix

$$\nabla_{\mathbf{w}}^2 \mathcal{F}^P(\mathbf{w}) \approx \mathbf{J}^T \mathbf{J} = [\mathbf{J}^D, \mathbf{J}^S] \begin{bmatrix} \mathbf{J}^D \\ \mathbf{J}^S \end{bmatrix} \in \mathbb{R}^{n_w \times n_w}, \quad (6.112)$$

mit der SOBOLEV-Jacobi-Matrix $\mathbf{J}^S$, benötigt. Da im weiteren Verlauf der Arbeit gradientenbasierte Verfahren keine Rolle spielen, wird auf eine ausführliche Diskussion verzichtet und im Folgenden die Einbindung der Nebenbedingung in das LMV behandelt. Durch entsprechende Modifikation können jedoch auch leicht die Einträge des Gradienten $\nabla_{\mathbf{w}} \mathcal{F}^P(\mathbf{w})$ berechnet werden.

6.3.3 Sobolev-Training mit dem Levenberg-Marquardt-Verfahren

Für das SOBOLEV-LEVENBERG-MARQUARDT-Verfahren (SLMV) ist die Berechnung des Fehlervektors der Tangenten-Nebenbedingung

$$\mathbf{h} = [h_{111}, \dots, h_{kji}, \dots, h_{PSJI}]^T \in \mathbb{R}^{(P^S \cdot J \cdot I)} \quad (6.113)$$

sowie die zugehörige SOBOLEV-Jacobi-Matrix

$$\mathbf{J}^S := \nabla_{\mathbf{w}} \mathbf{h} = \begin{bmatrix} \frac{\partial h_{111}}{\partial w_{10}^{(1)}} & \frac{\partial h_{111}}{\partial w_{20}^{(1)}} & \cdots & \frac{\partial h_{111}}{\partial w_{n_o n_{n_h}}^{(n_h+1)}} \\ \frac{\partial h_{112}}{\partial w_{10}^{(1)}} & \frac{\partial h_{112}}{\partial w_{20}^{(1)}} & \cdots & \frac{\partial h_{112}}{\partial w_{n_o n_{n_h}}^{(n_h+1)}} \\ \vdots & \vdots & \frac{\partial h_{kji}}{\partial w_{ml}^{(L)}} & \vdots \\ \frac{\partial h_{PSJI}}{\partial w_{10}^{(1)}} & \frac{\partial h_{PSJI}}{\partial w_{20}^{(1)}} & \cdots & \frac{\partial h_{PSJI}}{\partial w_{n_o n_{n_h}}^{(n_h+1)}} \end{bmatrix} \in \mathbb{R}^{(P^S \cdot J \cdot I) \times n_w} \quad (6.114)$$

erforderlich. Die Komponenten von $\mathbf{h}$ können nach Gleichung (6.108) unter Anwendung des Algorithmus 6.3 berechnet werden. Für die Einträge der Jacobi-Matrix ist die Berechnung der Terme

$$\frac{\partial h_{kji}}{\partial w_{ml}^{(L)}} = \frac{\partial}{\partial w_{ml}^{(L)}} \left(\frac{\partial \hat{z}_j(\mathbf{x}_k^S; \mathbf{w})}{\partial \hat{x}_i} \right) \quad (6.115)$$

notwendig. Die Herleitung orientiert sich dabei an den Ausführungen in WEBER [146], in der Modifikationen an den von BISHOP [17] vorgeschlagenen Vorschriften vorgenommen wurden. Die Abhängigkeit der Neuronenausgänge z_j von $(\mathbf{x}_k, \mathbf{w})$ wird aus Gründen der Lesbarkeit vernachlässigt. Da die Variablen $\mathbf{x}_k$ und $\mathbf{w}$ unabhängig voneinander sind, lässt sich (6.115) umformen zu

$$\frac{\partial}{\partial w_{ml}^{(L)}} \left(\frac{\partial \hat{z}_j(\mathbf{x}_k^S; \mathbf{w})}{\partial \hat{x}_i} \right) = \frac{\partial}{\partial \hat{x}_i} \left(\frac{\partial \hat{z}_j}{\partial w_{ml}^{(L)}} \right) \quad (6.116)$$

$$= \frac{\partial}{\partial \hat{x}_i} \left(\frac{\partial \hat{z}_j}{\partial s_m^{(L)}} \frac{\partial s_m^{(L)}}{\partial w_{ml}^{(L)}} \right) = \frac{\partial}{\partial \hat{x}_i} \left(\frac{\partial \hat{z}_j}{\partial s_m^{(L)}} \cdot y_l^{(L-1)} \right). \quad (6.117)$$

Analog zu den Ausführungen in Kapitel 6.1.3 werden die generalisierten Delta- und Gamma-Werte

$$\delta_{jm}^{(L)} := \frac{\partial \hat{z}_j}{\partial s_m^{(L)}} \quad \text{und} \quad \gamma_{jim}^{(L)} := \frac{\partial}{\partial \hat{x}_i} \left(\frac{\partial \hat{z}_j}{\partial s_m^{(L)}} \right) \quad (6.118)$$

eingeführt, womit sich unter Anwendung der Kettenregel die Beziehung (6.117) auch schreiben lässt als

$$\frac{\partial}{\partial \hat{x}_i} \left(\frac{\partial \hat{z}_j}{\partial w_{ml}^{(L)}} \right) = \delta_{jm}^{(L)} \left(\frac{\partial y_l^{(L-1)}}{\partial \hat{x}_i} \right) + \gamma_{jim}^{(L)} y_l^{(L-1)}. \quad (6.119)$$

Für die Berechnung der notwendigen Terme ist zunächst eine Vorwärtsrechnung gemäß Algorithmus (6.1) für die partiellen Ableitungen $\partial y^{(L)} / \partial \hat{x}_i$ durchzuführen. Im Zuge dessen ist es sinnvoll, auch die partiellen Ableitungen der gewichteten Summen $\partial s^{(L)} / \partial \hat{x}_i$ abzuspeichern, die später noch benötigt werden. Komponentenweise gilt dabei für die erste verdeckte Schicht

$$\frac{\partial y_k^{(1)}}{\partial \hat{x}_i} = g'(s_k^{(1)}) \frac{\partial s_k^{(1)}}{\partial \hat{x}_i} \quad \text{mit} \quad \frac{\partial s_k^{(1)}}{\partial \hat{x}_i} = w_{ki}^{(1)} \quad (6.120)$$

sowie für die folgenden Schichten

$$\frac{\partial y_m^{(L)}}{\partial \hat{x}_i} = g'(s_m^{(L)}) \frac{\partial s_m^{(L)}}{\partial \hat{x}_i} \quad \text{mit} \quad \frac{\partial s_m^{(L)}}{\partial \hat{x}_i} = \sum_{l=1}^{n_{L-1}} w_{ml} \frac{\partial y_l^{(L-1)}}{\partial \hat{x}_i} \quad (6.121)$$

und in der Ausgangsschicht

$$\frac{\partial z_j}{\partial \hat{x}_i} = g'(s_j^{(n_h+1)}) \frac{\partial s_j^{(n_h+1)}}{\partial \hat{x}_i} = 1 \cdot \frac{\partial s_j^{(n_h+1)}}{\partial \hat{x}_i} \quad \text{mit} \quad \frac{\partial s_j^{(n_h+1)}}{\partial \hat{x}_i} = \sum_{o=1}^{n_{n_h}} w_{jo} \frac{\partial y_o^{(n_h)}}{\partial \hat{x}_i}. \quad (6.122)$$

Die allgemeine Rekursionsformel lässt sich für die Implementierung kompakt in Vektor-Matrix Notation wie folgt angeben:

$$\frac{\partial \mathbf{y}^{(L)}}{\partial \hat{\mathbf{x}}_i} = g'(\mathbf{s}^{(L)}) \odot \frac{\partial \mathbf{s}^{(L)}}{\partial \hat{\mathbf{x}}_i} \in \mathbb{R}^{n_L} \quad \text{mit} \quad \frac{\partial \mathbf{s}^{(L)}}{\partial \hat{\mathbf{x}}_i} = \bar{\mathbf{w}}^{(L)} \left(\frac{\partial \mathbf{y}^{(L-1)}}{\partial \hat{\mathbf{x}}_i} \right) \in \mathbb{R}^{n_L}. \quad (6.123)$$

Die Delta-Werte lassen sich mit vergleichbaren Überlegungen wie in Kapitel 6.1.3 herleiten. Für die Ausgangsschicht folgt

$$\delta^{(n_h+1)} = \delta_{jm}^{(n_h+1)} \mathbf{1}_{n_0} \quad \text{mit} \quad \delta_{jm}^{(n_h+1)} = \frac{\partial \hat{z}_j}{\partial s_m^{(n_h+1)}} = \begin{cases} g'(s_m^{(n_h+1)}) = 1, & j = m \\ 0, & \text{sonst.} \end{cases} \quad (6.124)$$

Für alle weitere Schichten lassen sich diese ebenfalls rekursiv über

$$\delta^{(L)} = g'(\mathbf{s}^{(L)}) \odot (\bar{\mathbf{w}}^T \delta^{(L+1)}) \quad (6.125)$$

bestimmen. Abschließend verbleibt die Berechnung der Gamma-Werte, für welche auf eine komponentenweise Herleitung verzichtet wird. Mit den Definitionen (6.118) sowie Gleichung (6.125) folgt unter Anwendung der Produktregel

$$\begin{aligned} \gamma^{(L)} &= \frac{\partial}{\partial \hat{x}_i} [g'(\mathbf{s}^{(L)}) \odot (\bar{\mathbf{w}}^T \delta^{(L+1)})] \\ &= g''(\mathbf{s}^{(L)}) \odot \frac{\partial \mathbf{s}^{(L)}}{\partial \hat{x}_i} \odot (\bar{\mathbf{w}}^T \delta^{(L+1)}) + g'(\mathbf{s}^{(L)}) \odot \left(\bar{\mathbf{w}}^T \frac{\partial}{\partial \hat{x}_i} \delta^{(L+1)} \right) \end{aligned} \quad (6.126)$$

wobei mit $g''(\mathbf{s})$ der Vektor der zweiten Ableitungen der Aktivierungsfunktion eingeführt wurde. Mit Identifikation von

$$\gamma^{(L+1)} = \frac{\partial}{\partial \hat{x}_i} \delta^{(L+1)} \quad (6.127)$$

wird somit

$$\gamma^{(L)} = g''(\mathbf{s}^{(L)}) \odot \frac{\partial \mathbf{s}^{(L)}}{\partial \hat{x}_i} \odot (\bar{\mathbf{w}}^T \delta^{(L+1)}) + g'(\mathbf{s}^{(L)}) \odot (\bar{\mathbf{w}}^T \gamma^{(L+1)}) . \quad (6.128)$$

erhalten. In analoger Vorgehensweise zu Gleichung (6.124) folgt für die Initialisierung in der Ausgangsschicht

$$\begin{aligned} \gamma^{(n_h+1)} &= \gamma_{jim}^{(n_h+1)} \mathbf{1}_{n_0} \\ \text{mit} \quad \gamma_{jim}^{(n_h+1)} &= \frac{\partial}{\partial \hat{x}_i} \frac{\partial \hat{z}_j}{\partial s_m^{(n_h+1)}} = \begin{cases} g''(s_m^{(n_h+1)}) \frac{\partial s_j^{(n_h+1)}}{\partial \hat{x}_i}, & j = m \\ 0, & \text{sonst.} \end{cases} \end{aligned} \quad (6.129)$$

Damit sind alle notwendigen Beziehungen hergeleitet worden, womit für die Ableitungen des Fehlers der Tangenten-Nebenbedingungen nach den Gewichten der Schicht L folgt:

$$\frac{\partial}{\partial \hat{x}_i} \left(\frac{\partial \hat{z}_j}{\partial \mathbf{w}^{(L)}} \right) = \delta^{(L)} \left(\frac{\partial \mathbf{y}^{(L-1)}}{\partial \hat{x}_i} \right)^T + \gamma^{(L)} \mathbf{y}^{(L-1)T} . \quad (6.130)$$

Auch hier werden die aus der Berechnung resultieren Matrizen in die entsprechenden Spalten der Jacobi-Matrix eingeordnet. Die vorzunehmenden Schritte für eine

Berechne (Alg. 6.1):	$\hat{\mathbf{x}} \mapsto \hat{\mathbf{z}}$
Speicherung:	$\hat{\mathbf{z}}, \mathbf{y}^{(L)}, \mathbf{s}^{(L)}$
Global-Schleife:	Für alle $i \in I$ und $j \in J$
Erste Schicht:	$\frac{\partial \mathbf{s}^{(1)}}{\partial \hat{x}_i} = \bar{\mathbf{w}}_{(:,i)}, \frac{\partial \mathbf{y}^{(1)}}{\partial \hat{x}_i} = g'(\mathbf{s}^{(1)}) \odot \frac{\partial \mathbf{s}^{(1)}}{\partial \hat{x}_i}$
	for $L = 2, n_h$
	$\frac{\partial \mathbf{s}^{(L)}}{\partial \hat{x}_i} = \bar{\mathbf{w}}^{(L)} \left(\frac{\partial \mathbf{y}^{(L-1)}}{\partial \hat{x}_i} \right)$
	$\frac{\partial \mathbf{y}^{(L)}}{\partial \hat{x}_i} = g'(\mathbf{s}^{(L)}) \odot \frac{\partial \mathbf{s}^{(L)}}{\partial \hat{x}_i}$
	end
Ausgangsschicht:	$\frac{\partial z_j}{\partial \hat{x}_i} = \frac{\partial s_j^{(n_h+1)}}{\partial \hat{x}_i} = \bar{\mathbf{w}}_{(j,:)}^{(n_h+1)} \left(\frac{\partial \mathbf{y}^{(n_h)}}{\partial \hat{x}_i} \right)$
Initialisierung Delta:	$\boldsymbol{\delta}^{(n_h+1)} \in \mathbb{R}^{n_o}$, mit $\delta_{(j)}^{(n_h+1)} = 1$, 0 sonst
Initialisierung Gamma:	$\boldsymbol{\gamma}^{(n_h+1)} \in \mathbb{R}^{n_o}$,
	mit $\gamma_{(j)}^{(n_h+1)} = g''(s_j^{(n_h+1)}) \frac{\partial s_j^{(n_h+1)}}{\partial \hat{x}_i}$, 0 sonst
	$\nabla_{\mathbf{w}^{(n_h+1)}} \mathbf{h} = \boldsymbol{\delta}^{(n_h+1)} \left(\frac{\partial \mathbf{y}^{(n_h)}}{\partial \hat{x}_i} \right)^T + \boldsymbol{\gamma}^{(L)} \mathbf{y}^{(n_h)T}$
	for $L = n_h, \dots, 1$
	$\boldsymbol{\delta}^{(L)} = g'(\mathbf{s}^{(L)}) \odot \left(\bar{\mathbf{w}}^T \boldsymbol{\delta}^{(L+1)} \right)$
	$\boldsymbol{\gamma}^{(L)} = g''(\mathbf{s}^{(L)}) \odot \frac{\partial \mathbf{s}^{(L)}}{\partial \hat{x}_i} \odot \left(\bar{\mathbf{w}}^T \boldsymbol{\delta}^{(L+1)} \right) \dots$
	$+ g'(\mathbf{s}^{(L)}) \odot \left(\bar{\mathbf{w}}^T \boldsymbol{\gamma}^{(L+1)} \right)$
	$\nabla_{\mathbf{w}^{(L)}} \mathbf{h} = \boldsymbol{\delta}^{(L)} \left(\frac{\partial \mathbf{y}^{(L-1)}}{\partial \hat{x}_i} \right)^T \dots$
	$+ \boldsymbol{\gamma}^{(L)} \mathbf{y}^{(L-1)T}$
	end
Sortierung:	$\nabla_{\mathbf{w}^{(L)}} \mathbf{h} \rightarrow \mathbf{J}^S, L = 1, \dots, n_h + 1$

Algorithmus 6.8: Berechnungsvorschrift für die Einträge der Jacobi-Matrix $\nabla_{\mathbf{w}} \mathbf{h}$ für das SLMV

Implementierung sind Algorithmus 6.8 zusammengefasst. Die weiteren Schritte im Rahmen der Aktualisierung folgen dem Iterationsschema in Algorithmus 6.7.

Abschließend sei angemerkt, dass durch das SOBOLEV-Training kein zusätzlicher Berechnungsaufwand bei der *Aktualisierung* der Gewichte entsteht. Die Dimensionen des um die Nebenbedingung erweiterten Gradienten und der Hesse-Matrix – s. Gleichungen (6.111) und (6.112) – hängen nur von der Anzahl der Gewichte

n_w ab. Der „Bottleneck“ des LMV – die Invertierung der Hesse-Matrix – bleibt somit unbeeinflusst. Die in Algorithmus 6.8 gelisteten Berechnungsschritte fallen durch kleinere Vektor-Matrix-Operationen weniger ins Gewicht.

6.3.4 Beispiel zur Funktionsapproximation mit Sobolev-Training

Um die Vorteile des SOBOLEV-Trainings zu demonstrieren, wird das bereits diskutierte Beispiel aus Kapitel 6.2.4 aufgegriffen. Die partiellen Ableitungen der Sinc-Funktion sind gegeben durch

$$\nabla_{\mathbf{x}} \tilde{f}(\mathbf{x}) = \begin{bmatrix} \frac{\partial z}{\partial x} \\ \frac{\partial z}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{r \cos(r) - \sin(r)}{r^3} x \\ \frac{r \cos(r) - \sin(r)}{r^3} y \end{bmatrix}. \quad (6.131)$$

Die Berücksichtigung dieser Tangenten-Informationen der Referenzfunktion erfolgt über den zusätzlichen Fehlerterm

$$\mathcal{F}^S(\mathbf{w}) = \frac{\rho}{2P^S} \sum_{k=1}^{P^S} \sum_{i=1}^2 \left(\frac{\partial \hat{z}_k}{\partial \hat{x}_i} - \frac{\partial \hat{t}_k}{\partial \hat{x}_i} \right)^2. \quad (6.132)$$

In diesem Fall gilt $I = 2$ und $J = 1$ bezüglich der einzufordernden Ableitungskomponenten. Durch die Einforderung der Nebenbedingungen im normalisierten Trainingsraum wird der Penalty-Parameter zu $\rho = 1$ gesetzt, um die Fehlerterme $\mathcal{F}^S(\mathbf{w})$ und $\mathcal{F}^D(\mathbf{w})$ gleich zu gewichten. Hierzu kann angemerkt werden, dass keine allzu große Sensitivität bezüglich des Parameters besteht und für $\rho \in [10^{-3}, 10^2]$ ähnliche Effekte verzeichnet werden. Für kleinere Werte ist kein Einfluss der Nebenbedingung erkennbar und für größere Werte bricht das Training aufgrund der schlechten Konditionierung ab.

Für die Untersuchung des Einflusses der Anzahl der Punkte, an denen die Nebenbedingungen eingefordert wird, ist in Abbildung 6.6 die Approximation für $P^S = 10, 50$ und 100 dargestellt. Dabei werden die gleichen KNN-Architekturen $2 - 40 - 1$ verwendet, sowie der gleiche Satz mit 100 Trainingsdaten aus Kapitel 6.2.4. Es ist erkennbar, dass sich die Approximationsgüte mit zunehmender Anzahl an Punkten verbessert, an denen die Nebenbedingung eingefordert wird. Bereits für $P^S = 50$ ist visuell keine Unterscheidung mehr zur Referenzlösung (vgl. Abbildung 6.3) möglich. Dabei gilt stets, dass $\mathcal{D}^S \subseteq \mathcal{D}^{\text{train}}$, d.h., die Einforderung der Nebenbedingung erfolgt immer auf einer – ebenfalls mit der LHS-Methode generierten – Teilmenge der Datenpunkte. Dies ist sinnvoll, wenn die Referenzlösung bekannt ist und entsprechend unendlich viele Datenpunkte zur Verfügung stehen. Im Falle begrenzter Datensätze ist es notwendig, die Nebenbedingung an

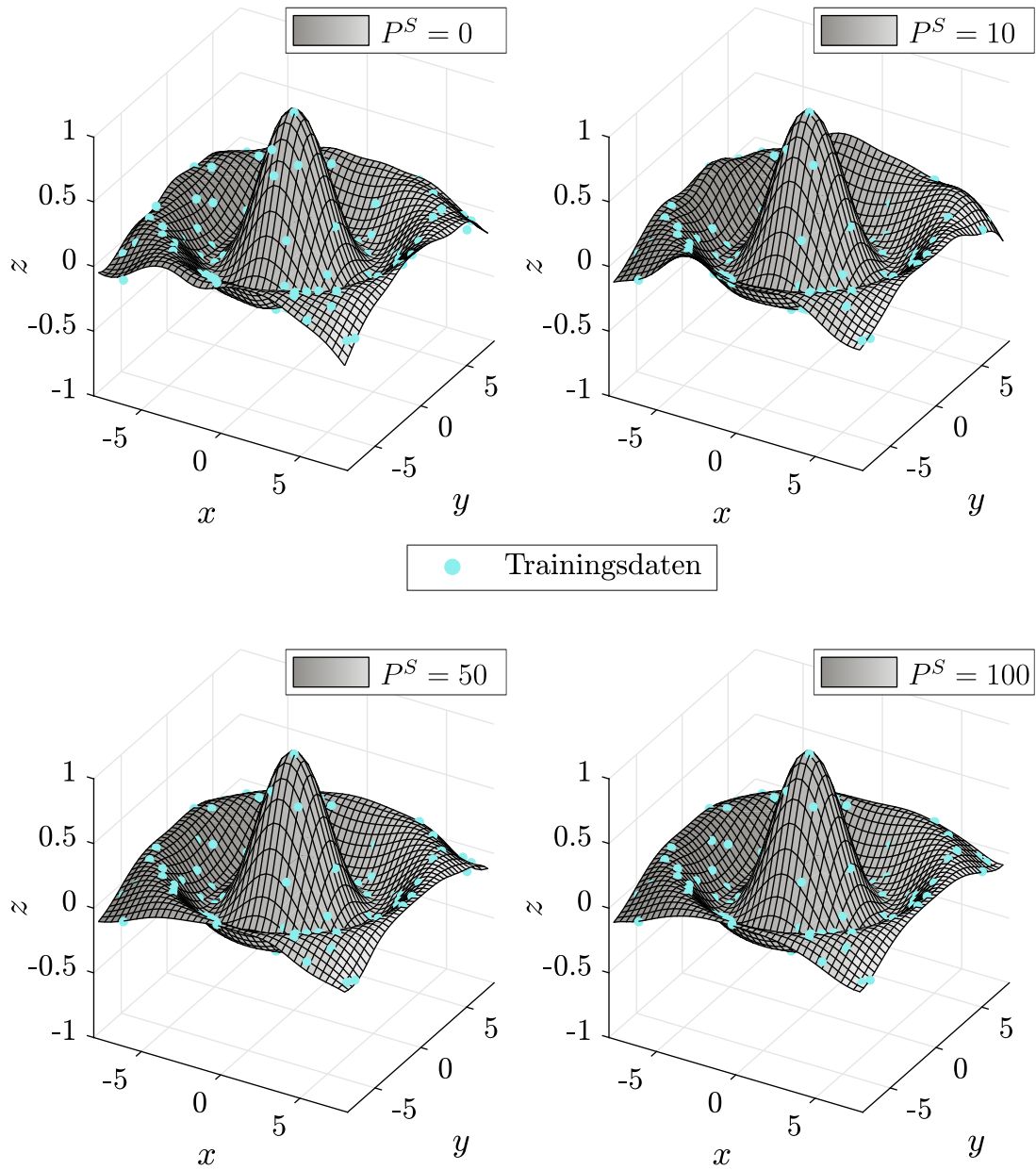


Abbildung 6.6: Darstellung der Approximation der Sinc-Funktion mit 100 LHS generierten Trainingsdaten ohne Nebenbedingung ($P^S = 0$), als auch mit steigender Anzahl an Nebenbedingungspunkten P^S

anderen Punkten innerhalb des Trainingsraumes einzufordern. Ein Beispiel hierfür wäre die Symmetrieeigenschaft der Sinc-Funktion auszunutzen, die lediglich als Anforderung in den partiellen Ableitungen des KNN formuliert werden kann und somit keine Referenzlösung benötigt wird. Für eine Diskussion hierzu wird auf WEBER [145] verwiesen.

Mit Hinzunahme der Information über die Tangente kann damit als Fazit bei

gleichbleibender Anzahl an Datenpunkte eine signifikante Verbesserung der Approximationsgüte erreicht werden. Diese Tatsache wird in den folgenden Kapiteln genutzt, um die Konstruktion von viskoelastischen KNN-Materialmodellen effizienter zu gestalten, in dem die Materialtangente in das KNN-Training integriert wird.

7 Konstruktion viskoelastischer Materialmodelle mit KNN

Im vorherigen Kapitel wurde demonstriert, wie ein KNN sehr erfolgreich zur Approximation einer analytischen Referenzlösung - der Sinc-Funktion - genutzt werden kann. Der Erfolg der Approximation für solch ein Trivialbeispiel ist der Tatsache geschuldet, dass zum einen eine klare Definition der Abbildungsvorschrift $(x, y) \mapsto z$ durch die Referenzlösung vorgegeben wird. Zum anderen handelt es sich um einen niedrigdimensionalen Zusammenhang, für den mittels LHS-Methode ohne weitere Überlegungen sehr einfach Trainingsdaten erzeugt werden konnten. Die Konstruktion von KNN-Materialmodellen erfordert bei den genannten Aspekten deutlich tiefergehende Überlegungen.

Im Folgenden werden dahingehend systematisch die getroffenen Überlegungen zur Konstruktion des im Rahmen der Arbeit verwendeten viskoelastischen KNN-Materialmodells erläutert. Beginnend mit der Definition einer geeigneten KNN-Architektur zur Erfassung aller Einflüsse auf das Materialverhalten, wird eine Methode zur Generierung von Trainingsdaten vorgeschlagen, die auf Grundlage des zuvor erwähnten Prinzips (vgl. Kapitel 6.2.4) „so viel wie nötig, so wenig wie möglich“ basiert. Weiterhin wird der Ansatz verfolgt, das KNN grundsätzlich auch auf Basis experimenteller Daten konstruieren zu können. Letztlich wird die Einbindung von Nebenbedingungen für viskoelastische KNN-Materialmodelle diskutiert. Dabei greifen stets die mechanischen Grundlagen der Kapitel 2-5 Hand in Hand mit denen der KNN aus Kapitel 6, was die notwendige Symbiose beider theoretischer Rahmen zur erfolgreichen Konstruktion hervorhebt.

7.1 KNN-Architektur für Viskoelastizität

Wie in Kapitel 3 beschrieben, zeichnet sich viskoelastisches Materialverhalten durch eine Abhängigkeit von den Dehnungen, deren Raten sowie weiteren Zustandsgrößen aus, welche die Belastungsgeschichte des Materials repräsentieren. Dies lässt sich in Abbildung 7.1 anhand eines eindimensionalen Beispiels verdeutlichen. Auf Materialpunktebene werden dazu zwei Dehnungsverläufe betrachtet, die für ein hier nicht näher spezifiziertes Material jeweils eine vollständige Be- und Entlastung durchlaufen. Beide Pfade erreichen die gleiche maximale Dehnung $\varepsilon = 0,01$, jedoch zu unterschiedlichen Zeitpunkten. In der Abbildung sind zusätzlich die resultierenden Spannungs-Dehnungs-Hysteresen dargestellt.

Für ein rein elastisches Material existieren solche Hysteresen nicht, da die Belastungs- und Entlastungspfade zusammen fallen, sodass ein eindeutiger funktionaler Zusammenhang $\sigma = \sigma(\varepsilon)$ vorliegt. Im Falle der Viskoelastizität treten

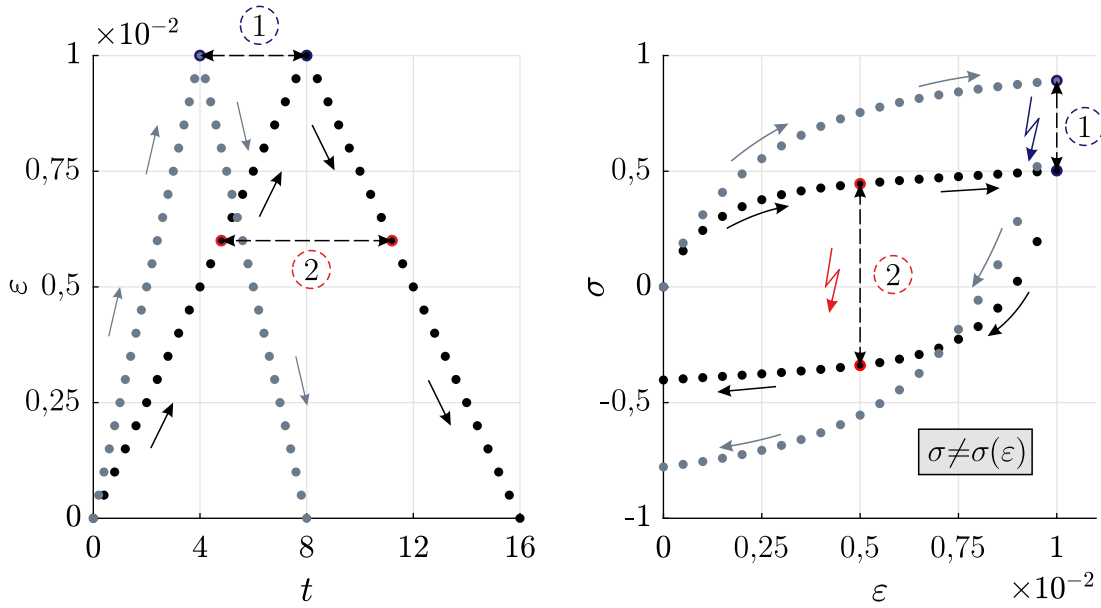


Abbildung 7.1: Darstellung zweier viskoelastischer Hysteresen (rechts) für zwei unterschiedliche Dehnungsverläufe (links) bei vollständiger Be- und Entlastung.

jedoch zwei Eigenschaften auf, die im Widerspruch zu rein elastischen Materialien stehen und die diesen Zusammenhang aufheben:

1. **Ratenabhängigkeit:** Bei gleicher Dehnung können unterschiedliche Spannungen auftreten (vgl. Punkt 1 in Abbildung 7.1), da die viskosen Anteile der Materialantwort von der *Belastungsgeschwindigkeit* abhängen.
2. **Pfadabhängigkeit:** Innerhalb eines Pfades können identische Dehnungen zu unterschiedlichen Spannungsantworten führen, abhängig von der *Belastungsgeschichte* (vgl. Punkt 2 in Abbildung 7.1).

Diese Beobachtungen verdeutlichen, dass für eine korrekte Abbildung des Materialverhaltens nicht nur die aktuelle Dehnung, sondern auch die Dehnrates und geeignete Zustandsvariablen, welche die Historie erfassen, berücksichtigt werden müssen. Entsprechend motiviert dies die Wahl der Eingangsvariablen für das KNN, was im Folgenden unter Verwendung des in Kapitel 6.1 diskutierten Feedforward-Netzes erfolgt.

7.1.1 Modellierung mit Feedforward-Netzen

Eine mögliche Definition der Eingangs- und Ausgangsvariablen für die KNN-Materialfunktion $\mathbf{x} \mapsto \mathbf{z}$, welche die zuvor genannten Effekte berücksichtigt, ist

$$\mathbf{x} := [\Delta \mathbf{E}_{n+1}, \mathbf{E}_n, \mathbf{S}_n, \Delta t_{n+1}, \mathbf{p}]^T \in \mathbb{R}^{(3n_{\text{str}}+1+n_p)} \quad \text{und} \quad \mathbf{z} := \mathbf{S}_{n+1} \in \mathbb{R}^{n_{\text{str}}}. \quad (7.1)$$

Dabei werden die in Kapitel 3.3 eingeführten Indizes n und $n+1$ zur Bezeichnung des letzten sowie aktuellen Gleichgewichtspunktes verwendet. Entsprechend sind $\mathbf{E}_n$ und $\mathbf{S}_n$ die Verzerrungen und Spannungen des letzten Gleichgewichtspunktes und $\Delta\mathbf{E}_{n+1}$ das aktuelle Verzerrungsinkrement. Diese führen auf die ersten $3n_{\text{str}}$ Eingangsvariablen des KNN, wobei mit n_{str} die von der Problemstellung abhängige Anzahl an Verzerrungskomponenten eingeführt wurde. Im Rahmen dieser Arbeit gilt:

$$n_{\text{str}} = \begin{cases} 1, & \text{im Eindimensionalen} \\ 6, & \text{für dreidimensionale Kontinua} \\ 8, & \text{für Schalen} \end{cases} \quad (7.2)$$

an. Neben dem aktuellen Zeitinkrement Δt_{n+1} stellen diese die notwendigen Eingangsvariablen dar, womit der Eingangsvektor immer aus mindestens $3n_{\text{str}} + 1$ Variablen besteht. Ein weiterer Bestandteil ist der Vektor $\mathbf{p} \in \mathbb{R}^{n_p}$, der zusätzliche Variablen, wie beispielsweise geometrische oder materielle Eigenschaften einer Mikrostruktur, beinhalten kann. Dieser ist optional und wird erst in Kapitel 8 genauer diskutiert. Der Ausgangsvektor enthält nur die gesuchten Spannungen $\mathbf{S}_{n+1}$ des aktuellen Gleichgewichtspunktes womit das KNN bezüglich der Eingangsvariablen inkrementell und bezüglich der Ausgangsvariablen explizit definiert wird⁴⁷. Mit den Definitionen (7.1) folgt die KNN-Materialtangente

$$\mathbf{C}_T^{\text{KNN}} = \frac{\partial \mathbf{S}_{n+1}}{\partial \mathbf{E}_{n+1}} = \frac{\partial \mathbf{S}_{n+1}}{\partial \Delta \mathbf{E}_{n+1}} = \bar{\mathbf{C}}_{T(1:n_{\text{str}}, 1:n_{\text{str}})}^{\text{KNN}} \in \mathbb{R}^{n_{\text{str}} \times n_{\text{str}}}, \quad (7.3)$$

deren Einträge mit Algorithmus 6.3 berechnet werden können. Wichtig ist dabei anzumerken, dass zunächst keine Symmetrieanforderungen an die KNN-Materialtangente gestellt werden und diese somit per Konstruktion nicht symmetrisch ist. Damit folgt

$$\mathbf{C}_T^{\text{KNN}} \neq (\mathbf{C}_T^{\text{KNN}})^T, \quad (7.4)$$

was im Widerspruch zu den beschriebenen analytischen viskoelastischen Materialmodellen – z.B. dem verallgemeinerten MAXWELL-Modell – steht. Dies stellt zunächst keine Einschränkung bezüglich des Einsatzes dar, jedoch lassen sich hier bereits gewisse Inkonsistenzen zu physikalischen Rahmenbedingungen erkennen. Dies wird im nächsten Unterkapitel diskutiert.

Grundsätzlich lassen sich an der Struktur des KNN-Materialmodells Ähnlichkeiten zu der konstitutiven Beziehung des verallgemeinerten MAXWELL-Modells aus Gleichung (3.75) erkennen. Insbesondere hängt die aktuelle Spannung nicht nur vom aktuellen Verzerrungszustand, sondern zusätzlich von datenbasierten Geschichtsvariablen ab. Weiterhin wird die Ratenabhängigkeit des Materialverhaltens über die Zeitschrittweite explizit berücksichtigt. Im Gegensatz zum verallgemeinerten MAXWELL-Modell erfolgt beim hier verwendeten KNN jedoch keine

⁴⁷Eine inkrementelle Darstellung der Ausgangsvariablen $\mathbf{z} := \Delta \mathbf{S}_{n+1}$, die beispielsweise in WEBER [145] für elastoplastischen Materialverhalten verwendet wird, wäre ebenfalls denkbar.

explizite Trennung in volumetrische und deviatorische Spannungsanteile. Dies ist beim hier verwendeten KNN nicht der Fall. Auch dies stellt keine Einschränkung dar, das KNN muss allerdings selbstständig aus den zur Verfügung gestellten Daten erlernen, dass sich gegebenenfalls nur die deviatorischen Verzerrungskomponenten viskos verhalten. Des Weiteren werden beim analytischen Modell offensichtlich andere Zustandvariablen (Geschichtsvariablen) verwendet. Dabei handelt es jedoch um eben jene im Allgemeinen nicht messbaren inneren Variablen, die aus reinen Spannungs-Dehnungs-Hyteresen nicht zu extrahieren sind, weshalb beim KNN die Spannung des letzten Gleichgewichtspunkts als Geschichtsvariable verwendet wird.

Abschließende Bemerkung zur Ratenabhängigkeit

Im vorliegenden Fall wurde die Ratenabhängigkeit nicht explizit über den Quotienten $\dot{\mathbf{E}}_{n+1} = \Delta \mathbf{E}_{n+1} / \Delta t_{n+1}$ vorgegeben, sondern durch die getrennte Angabe von Verzerrungsincrementen und Zeitinkrement als eigenständige Eingangsvariablen. Grundsätzlich könnte auch direkt die Dehnrates verwendet werden. Die Trennung erleichtert jedoch einerseits die Definition geeigneter Grenzen des Trainingsraums. Des Weiteren führte dies in der Praxis zu einer Verbesserung des KNN-Trainings. Der zusätzliche Aufwand ist gering, da lediglich eine weitere Eingangsgröße im Vergleich zu den ohnehin vorhandenen $3 n_{\text{str}}$ Variablen hinzukommt.

7.1.2 Konsistenz des Modells mit Grundsätzen der Materialmodellierung

Wie in der Einleitung zu Kapitel 3 und insbesondere im Abschnitt 3.2 diskutiert, existieren verschiedene Anforderungen, um Materialmodelle als physikalisch konsistent einzustufen. Eine kritische Einordnung des hier verwendeten viskoelastischen KNN-Materialmodells ist daher notwendig.

Die Grundanforderungen zur Einordnung in die Theorie der *einfachen Stoffe* sind erfüllt, da ausschließlich objektive, vom lokalen Punkt $\mathbf{X}$ abhängige Größen $\mathbf{E}$ und $\mathbf{S}$ verwendet werden. Das KNN-Materialmodell ist somit konsistent bezüglich der Prinzipien *Determinismus*, *lokale Wirkung* und *materielle Objektivität*.

Die Forderung einer spannungsfreien Referenzkonfiguration (vgl. Gleichung (3.30)) ist im vorliegenden Ansatz nicht streng enthalten. Für ein rein elastisches, nur in $\mathbf{E}$ und $\mathbf{S}$ formuliertes KNN könnte diese Bedingung beispielsweise durch Elimination der Bias-Neurone sichergestellt werden sowie der Verwendung einer Aktivierungsfunktion mit $g(0) = 0$, sodass $f(\mathbf{0}; \mathbf{w}) = \mathbf{0}$ gilt. Mit Einführung des Zeitinkrements als zusätzlicher Eingangsvariable ist dies nicht unmittelbar gewährleistet. Die Bedingung kann jedoch über eine Nebenbedingung schwach er-

zwungen werden (vgl. Abschnitt 7.3.2).

Ein weiterer Aspekt betrifft die Erfüllung der Clausius–Planck-Ungleichung (3.31), die sich inkrementell als

$$\mathcal{D}_{int_{n+1}} \approx \mathbf{S}_{n+1} \cdot \Delta \mathbf{E}_{n+1} - (\Psi_{n+1} - \Psi_n) \geq 0 \quad (7.5)$$

darstellen lässt. Die verwendete Netzarchitektur kann diese Bedingung strukturell nicht garantieren, da weder die Dissipation noch die freie Energie explizit im Modell enthalten sind, bzw. berechnet werden können. Eine thermodynamische Konsistenz ist daher nicht a priori gegeben. Dies bedeutet jedoch nicht, dass das Modell für praktische Anwendungen ungeeignet ist. Es muss jedoch berücksichtigt werden, dass im hochdimensionalen Approximationsraum Punkte existieren können, an denen das Modell potenziell unphysikalisches Verhalten zeigt.

7.1.3 Diskussion weiterer KNN Architekturen

Als Abschluss zu der zuvor geführten Diskussion werden an dieser Stelle einige weitere KNN-Modelle erwähnt, die zur Modellierung von viskoelastischen Materialien verwendet werden. Dies soll nicht den wissenschaftlichen Stand der Einleitung wiederholen, sondern das hier verwendete Modell einordnen und einen Ausblick hinsichtlich möglicher Alternativen vermitteln.

Nach besten Wissen des Autors gehen die erste Ansätze viskoelastischer KNN-Modelle auf die Arbeiten von JUNG und GHABOUSSI [62] zurück, die die Verwendung einer impliziten KNN-Architektur vorschlagen. Die Ein- und Ausgangsvektoren sind dabei wie folgt definiert:

$$\mathbf{x} := [\mathbf{E}_{n+1}, \mathbf{E}_n, \mathbf{S}_{n+1}, \mathbf{S}_n, \dot{\mathbf{E}}_{n+1}, \dot{\mathbf{E}}_n, \dot{\mathbf{S}}_n] \quad \text{und} \quad \mathbf{z} := \dot{\mathbf{S}}_{n+1}. \quad (7.6)$$

Damit ergibt sich die implizite Gleichung in den gesuchten Spannungen

$$\mathbf{S}_{n+1} = \mathbf{S}_n + \Delta t_{n+1} \dot{\mathbf{S}}_{n+1}, \quad (7.7)$$

die iterativ gelöst werden muss, da die aktuelle Spannungsrate von der gesuchten Spannung im Eingangsvektor des KNNs abhängt. Der Vorteil der Methode ist die Unabhängigkeit von der Zeitschrittweite, für die im Rahmen des Trainings keine expliziten Grenzen gewählt werden müssen. Das Modell ist dahingehend in der Theorie für beliebige Zeitschrittweiten einsetzbar. Neben der iterativen Lösung des Gleichungssystems - die bei der Verwendung in der FEM an jedem GAUSS-Punkt erfolgen muss - führt der Aufbau des Eingangsvektors bei der Verwendung als Konstitutivmodell der Schale auf 56 Eingangsvariablen, was zu tiefen und rechenintensiven KNN-Architekturen führt.

Als Alternative zu klassischen Feedforward-Netzen ist auch der Einsatz von rekurrenten neuronalen Netzen (RNN) denkbar. Ansätze hierzu sind u.a. in OESER und FREITAG [102] oder CHEN [20] zu finden. RNN zeichnen sich durch eine interne Rückkopplung aus, womit Neuronenausgaben vergangener Zustände $[\mathbf{y}_n^{(L)}, \mathbf{y}_{n-1}^{(L)}, \mathbf{y}_{n-2}^{(L)}, \dots]$ als zusätzliche Eingänge in den aktuellen Zustand $\mathbf{y}_{n+1}^{(L)}$ eingehen. Damit entfällt die explizite Wahl von Geschichtsvariablen. Das Training dieser Netz-Architekturen benötigt jedoch erfahrungsgemäß große Datenmengen und das Phänomen des *verschwindenden Gradienten* das in Kapitel 6.2.1 angesprochen wurde ist hier besonders ausgeprägt. Es hat sich außerdem gezeigt, dass sich diese in der Numerik instabil verhalten, weshalb auch nach Wissens des Autos derzeit keine Arbeiten zu deren erfolgreichen Einsatz für viskoelastische Problemstellung im Rahmen der FEM existieren.

Die letzte Kategorie von derzeit untersuchten KNN-Architekturen zielt auf die direkte Einbindung von physikalischen Rahmenbedingungen bei deren Konstruktion ab. Häufig wird in diesem Zusammenhang der Begriff *Physics Augmented Neural Networks* verwendet. Deren Funktionsweise entspringt dem Gedanken zur Lösung partieller Differentialgleichungen und entsprechend lassen sich diese auch für die Materialmodellierung einsetzen. Hierbei werden sogenannte Pseudo-Potentiale für die freie Energie Ψ und die Dissipation Γ eingeführt, aus denen sich wiederum durch Bilden der partiellen Ableitung nach den Eingangsvariablen – s. Gleichung (3.34) – die Spannungen über

$$\mathbf{S}_{n+1} = \frac{\partial \Psi}{\partial \mathbf{E}_{n+1}} \quad (7.8)$$

berechnen lassen. Durch geschickte Konstruktion der KNN-Architektur lassen sich mathematische Zwangsbedingungen, wie u.a. die Konvexitätsanforderungen an die Potentiale, einbauen. Dieses Vorgehen wird beispielsweise in HUANG ET AL. [59] verfolgt. Ein darauf aufbauender vielversprechender Ansatz, der auch ohne die explizite Definition von internen Variablen auskommt, ist in ROSENKRANZ ET. AL. [114] zu finden. Eine Integration des Modells in die FEM ist hierbei noch ausstehend.

7.2 Generierung von viskoelastischen Trainingsdaten

Ein zentraler Schritt bei der Konstruktion des KNN-Materialmodells ist die Bereitstellung geeigneter Trainingsdaten. Deren Qualität und Struktur bestimmen maßgeblich die Leistungsfähigkeit des Modells. Ziel ist daher die Entwicklung einer effizienten Strategie zur Datenerzeugung, die mit möglichst wenigen Datenpunkten einen maximalen Informationsgehalt sicherstellt. Im Mittelpunkt steht die Frage, welche Komponenten des Eingangsvektors in welcher Form über die

Zeit variiert werden. Ergänzend müssen geeignete Grenzen des Trainingsraums definiert werden, insbesondere im Hinblick auf die Wahl der Zeitschrittweite Δt . Grundlage dieser Überlegungen bilden die in Kapitel 3.4 diskutierten Methoden viskoelastischer Materialversuche, womit sich das hier vorgeschlagene Vorgehen grundsätzlich auch auf experimentelle Spannungs-Verzerrungs-Daten anwenden lässt. Das übergeordnete Ziel stellt jedoch die Datengenerierung unter Verwendung numerischer Homogenisierung dar.

7.2.1 Definition expliziter Zeitpfade auf Basis von Materialversuchen

Wegen der Pfadabhängigkeit des Materialverhaltens lassen sich die für das Training der KNN-Architektur benötigten Daten – anders als im Einführungsbeispiel aus Kapitel 6.2.4 – nicht einfach mithilfe der LHS-Methode in den Eingangsvariablen erzeugen. Während die Eingangsgrößen des Modells inkrementell definiert sind, basieren sie zunächst auf expliziten Zeitpfaden, innerhalb derer die Inkremente berechnet werden. Ein solcher Pfad $\mathcal{P}_\alpha$ mit $\alpha = 1, \dots, N_{\text{Pfade}}$ hat die Form

$$\mathcal{P}_\alpha = \left\{ (\mathbf{E}_n, \mathbf{S}_n, t_n, \mathbf{p}, \mathbf{C}_{Tn}) \right\}_{n=0}^N, \quad (7.9)$$

und kann anschließend in die inkrementelle Darstellung (7.1) überführt werden. Dabei ist zu beachten, dass in der inkrementellen Darstellung die explizite Pfadinformation nicht mehr enthalten ist. Die Daten können daher prinzipiell unabhängig vom ursprünglichen Pfad angeordnet oder reduziert werden.

Die Größen $\mathbf{E}$, $\mathbf{S}$, t_n und $\mathbf{p}$ müssen dabei innerhalb geeigneter Grenzen vorgegeben werden (vgl. Abschnitt 7.2.4). Grenzen sind auch für die Ausgangsgröße $\mathbf{S}$ erforderlich, da synthetische Materialversuche sowohl spannungs- als auch verzerrungsgesteuert durchgeführt werden können. Die Materialtangente $\mathbf{C}_T$ fällt hingegen als Nebenprodukt der Datengenerierung im Rahmen der Homogenisierung – oder unter Verwendung eines analytischen Referenzmodells – an und kann im KNN-Training über zusätzliche Nebenbedingungen berücksichtigt werden.

Die zentrale Frage ist, wie Verzerrungen und Spannungen auf Materialpunktebene über die Zeit vorgegeben werden. Eine Möglichkeit wäre, zufällige Verzerrungspfade (engl.: *random strain walks*) über ein definiertes Zeitintervall $t \in [0, T]$ zu erzeugen. Eine solche Strategie wird beispielsweise von BOZKURZT und TAGARELLI [18] genutzt⁴⁸. In dieser Arbeit wird hingegen ein anderer Ansatz verfolgt. Aufbauend auf den in Kapitel 3.4 erläuterten experimentellen Methoden werden Pfade konstruiert, die das Relaxations-, Kriech- und zyklische Verhalten des

⁴⁸In der genannten Arbeit wird nichtlineares viskoelastisches Materialverhalten untersucht. Aufgrund der dort berücksichtigten großen Verzerrungen unterscheidet sich der Eingangsraum deutlich, sodass die Größe der verwendeten Datensätze im weiteren Verlauf der Arbeit nicht direkt mit diesen vergleichbar ist.

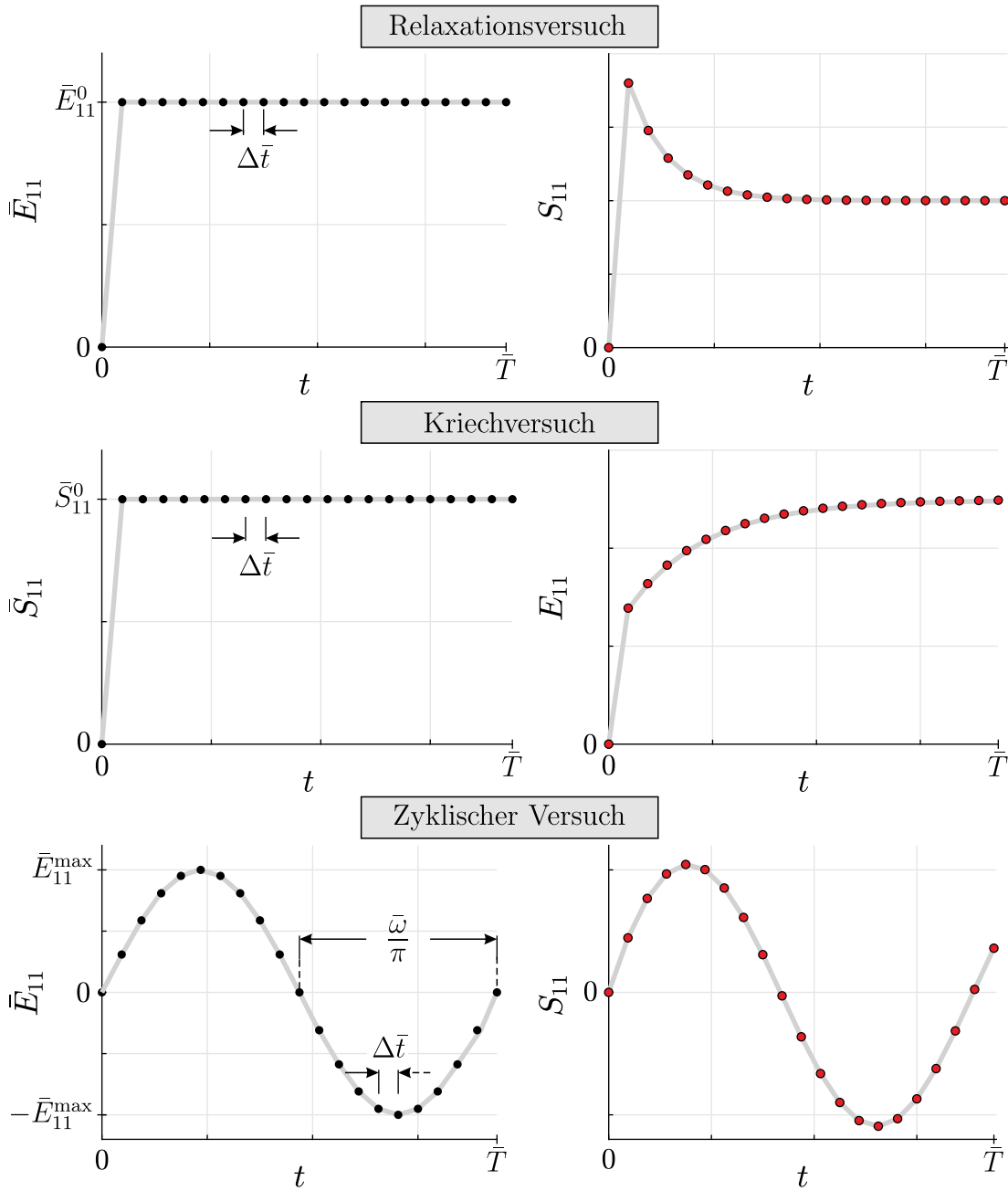


Abbildung 7.2: Darstellung der drei verwendeten Trainingspfade mit den zu wählenden Pfadparametern (links) sowie die resultierende Materialantworten (rechts)

Materials widerspiegeln. Der entscheidende Vorteil dieses Vorgehens besteht darin, dass das KNN-Modell prinzipiell auch unmittelbar auf Basis experimenteller Versuchsdaten trainiert werden kann.

Dahingehend werden im Folgenden für das *Training* des KNN drei Pfade verwendet, die wie folgt parametrisiert sind:

Versuchstyp	Parameter
Relaxationsversuch	$\Delta\bar{t}, \bar{T}, \bar{E}_{11}^0$
Kriechversuch	$\Delta\bar{t}, \bar{T}, \bar{S}_{11}^0$
Zyklischer Versuch	$\Delta\bar{t}, \bar{T}, \bar{E}_{11}^{\max}, 2\bar{\omega}/\pi$

Tabelle 7.1: Parametrisierung zur Definition einzelner Trainingspfade

Relaxationsversuch:	$\mathbf{E}(t) = \bar{\mathbf{E}}^0$	(konstante Verzerrung)
Kriechversuch:	$\mathbf{S}(t) = \bar{\mathbf{S}}^0$	(konstante Spannung)
Zyklischer Versuch:	$\mathbf{E}(t) = \bar{\mathbf{E}}^{\max} \sin(\bar{\omega}t)$	(sinusförmige Verzerrung)

Die Zeitdiskretisierung erfolgt durch

$$t_n = n \Delta\bar{t}, \quad n = 0, \dots, N, \quad N\Delta\bar{t} = \bar{T}. \quad (7.10)$$

Alle mit $(\bar{\cdot})$ gekennzeichneten Größen sind dabei geeignet zu wählen. Hierzu zählen die minimalen und maximalen Verzerrungen bzw. Spannungen sowie die Zeitschrittweite, die zugleich als Eingangsvariablen in das KNN eingehen und deren Grenzen in Kapitel 7.2.4 diskutiert werden. Zusätzlich müssen die Versuchsdauer $\bar{T}$ und im Falle zyklischer Pfade die Periode $2\bar{\omega}/\pi$ festgelegt werden. Für Relaxations- und Kriechversuche ist es dabei erforderlich, dass die Trainingsdaten auch die rein elastische Materialantwort abbilden, was die Festlegung von $\bar{T}$ steuert. Bei zyklischen Versuchen wird die Periode so gewählt, dass mindestens eine, jedoch nicht mehr als zwei Schwingungen durchlaufen werden, um eine übermäßige Dopplung von Daten zu vermeiden. Entsprechend lassen sich die drei Pfade im Eindimensionalen wie in Tabelle 7.1 parametrisieren. Die verwendeten Trainingspfade sind in Abbildung 7.2 qualitativ für den eindimensionalen Fall dargestellt.

Im Rahmen der Implementierung kann für jeden der gezeigten Parameter ein Intervall definiert werden, innerhalb dessen mit der LHS-Methode eine Schar von n_{Pfade} Pfaden erzeugt wird. Die entsprechende Parametrisierung ist aus Tabelle 7.2 zu entnehmen. Die Gesamtanzahl der verwendeten Trainingspfade N_{Pfade} ergibt sich somit aus der Summe der generierten Pfadscharen n_{Pfade} . Für die zyklischen Versuche wird die Eigenkreisfrequenz $\bar{\omega}$ stets so gewählt, dass gilt $P = 2\bar{\omega}/\pi \in [1, 2]$.

Da das KNN die Fähigkeit zur Generalisierung besitzen muss, werden neben den drei Typen von Trainingspfaden zusätzliche Testpfade vorgegeben, die an den entsprechenden Stellen individuell diskutiert werden. Dabei handelt es sich beispielsweise um Zugversuche mit Entlastung, die dem KNN aus dem Training unbekannt

sind. Die Pfade lassen sich außerdem sinngemäß auf den mehrdimensionalen Fall erweitern. Die Parametrisierung der Pfade muss in dem Fall dahingehend erweitert werden, dass neben der Größe der Verzerrungs- bzw. Spannungskomponente auch die Richtung vorgegeben werden muss. Für Bi- oder Triaxialversuche müssen entsprechend alle zuvor Komponenten und entsprechende Richtungen angegeben werden. Die finale Erweiterung ist die Einbindung der zusätzlichen Variablen $\mathbf{p} \in \mathbb{R}^{n_p}$, die es ermöglicht Pfade unter zusätzlicher Vorgabe wie Materialeigenschaften oder geometrischen Größen zu parametrisieren.

7.2.2 Eine adaptive Zeitschrittweiten-Strategie

Innerhalb eines Pfades wurde die Zeitschrittweite bislang konstant gewählt. Hier stellt sich die Frage, ob die Verwendung einer konstanten Zeitschrittweite effizient ist, welcher im Rahmen des Kapitels 8 nachgegangen wird. Die Frage lässt sich am Relaxationsversuch in Abbildung 7.2 (oben rechts) motivieren; Der anfängliche Bereich mit starkem Spannungsabfall wird dabei genauso fein diskretisiert wie der spätere Bereich, in dem sich die Materialantwort nur noch langsam dem elastischen Gleichgewicht annähert. Um diesem entgegenzuwirken, wird eine adaptive zeitliche Diskretisierung vorgeschlagen. Der aktuelle Zeitpunkt t_n ergibt sich dabei aus einem Exponentialansatz

$$t_n = C \cdot (e^{\alpha n} - 1), \quad n = 0, \dots, N, \quad (7.11)$$

wobei C ein Skalierungsfaktor und α die Krümmung der Exponentialfunktion steuert. Die zugehörige Parametrisierung für die Relaxations- und Kriechversuche sind in Tabelle 7.3 gegeben.

Unter Vorgabe der minimalen Zeitschrittweite zu Beginn, der maximalen am Ende sowie der Gesamtdauer $\bar{T}$ können die Parameter des Ansatzes (7.11) bestimmt werden. Mit

$$t_1 = C \cdot (e^\alpha - 1) \stackrel{!}{=} \Delta \bar{t}_{\min} \quad (7.12)$$

und

$$t_N - t_{N-1} = C \cdot (e^{\alpha N} - e^{\alpha(N-1)}) \stackrel{!}{=} \Delta \bar{t}_{\max} \quad (7.13)$$

Versuchstyp	Parametrisierung
Relaxationsversuch	$n_{\text{Pfade}}, [\Delta \bar{t}_{\min}, \Delta \bar{t}_{\max}], [\bar{T}_{\min}, \bar{T}_{\max}], [\bar{E}_{11\min}^0, \bar{E}_{11\max}^0]$
Kriechversuch	$n_{\text{Pfade}}, [\Delta \bar{t}_{\min}, \Delta \bar{t}_{\max}], [\bar{T}_{\min}, \bar{T}_{\max}], [\bar{S}_{11\min}^0, \bar{S}_{11\max}^0]$
Zyklischer Versuch	$n_{\text{Pfade}}, [\Delta \bar{t}_{\min}, \Delta \bar{t}_{\max}], [\bar{T}_{\min}, \bar{T}_{\max}], [\bar{E}_{11}^{\min}, \bar{E}_{11}^{\max}], \dots$ $[2\bar{\omega}/\pi] \in [1, 2]$

Tabelle 7.2: Parametrisierung zur Generierung einer Schar von Trainingspfaden

ergibt sich schließlich die nichtlineare Gleichung zur Bestimmung von α :

$$\frac{e^{\alpha N} - e^{\alpha(N-1)}}{e^{\alpha} - 1} = \frac{\Delta \bar{t}_{\max}}{\Delta \bar{t}_{\min}}. \quad (7.14)$$

Die Nullstelle der Gleichung kann beispielsweise mit der MATLAB-Funktion **fzero** bestimmt werden. Dazu wird zunächst ein Intervall für die Anzahl der Zeitschritte festgelegt, hier $N \in [20, 80]$. Beginnend mit $N = 20$ wird zunächst der Dämpfungsparameter α ermittelt. Daraus ergibt sich in der ersten Iteration der Skalierungsfaktor

$$C = \frac{\Delta \bar{t}_{\min}}{e^{\alpha} - 1}. \quad (7.15)$$

Mit diesen Parametern kann anschließend die Gesamtdauer des Versuchs berechnet werden:

$$t_N = C \cdot (e^{\alpha N} - 1). \quad (7.16)$$

Im allgemeinen Fall gilt $t_N \neq \bar{T}$. Daher wird eine Toleranz der angestrebten Versuchsdauer zugelassen, $\bar{T} \in [T_{\min}, T_{\max}]$. Liegt die berechnete Dauer außerhalb dieser Schranken, so wird die Anzahl der Zeitschritte N iterativ erhöht, bis eine passende Lösung gefunden ist. Die Implementierungsvorschrift ist in Algorithmus 7.4 für ein festgelegtes Intervall $N \in [N_{\min}, N_{\max}]$ zusammengefasst. Sollte innerhalb dieses Bereichs keine Lösung gefunden werden, sind die Parameter (z. B. Toleranzgrenzen oder Wahl von $N_{\min}$) entsprechend anzupassen.

In Abbildung 7.3 ist exemplarisch ein Relaxationsversuch mit adaptiver Steuerung der Zeitschrittweite dargestellt. Es ist erkennbar, dass im Vergleich zu Abbildung 7.2 durch eine feinere Zeitdiskretisierung zu Beginn des Versuchs zu einer verbesserten Auflösung der Spannungsantwort führt. Den Einfluss der adaptiven Zeitdiskretisierung auf das KNN-Training sowie auf die Approximationsfähigkeit des KNN wird in Kapitel 8 untersucht.

7.2.3 Explizite Zeitpfade: Zufällige Datengenerierung

Der in Kapitel 7.2.1 vorgestellte Ansatz, explizite Pfade auf Basis experimenteller Materialversuche zur Datengenerierung zu nutzen, wird in Kapitel 8 einem Ver-

Versuchstyp	Parameter	Parameterintervall
Relaxationsversuch	$\Delta \bar{t}_{\min}, \Delta \bar{t}_{\max}, \bar{T}, \bar{E}_{11}^0$	$[\bar{T}_{\min}, \bar{T}_{\max}], [\bar{E}_{11,\min}^0, \bar{E}_{11,\max}^0]$
Kriechversuch	$\Delta \bar{t}_{\min}, \Delta \bar{t}_{\max}, \bar{T}, \bar{S}_{11}^0$	$[\bar{T}_{\min}, \bar{T}_{\max}], [\bar{S}_{11,\min}^0, \bar{S}_{11,\max}^0]$

Tabelle 7.3: Parametrisierung und zugehörige Parameterintervalle für die Definition einzelner Trainingspfade mit adaptiver Zeitschrittweite sowie Parameterintervalle zur Erzeugung einer Schar von Pfaden

```

Initialisierung:  $\Delta \bar{t}_{\min}, \Delta \bar{t}_{\max}, N_{\min}, N_{\max}, \bar{T}_{\min}, \bar{T}_{\max}$ 
Schleife: for  $N = N_{\min}, \dots, N_{\max}$ 
    Löse:  $\frac{e^{\alpha \cdot N} - e^{\alpha \cdot (N-1)}}{(e^{\alpha} - 1)} = \frac{\Delta \bar{t}_{\max}}{\Delta \bar{t}_{\min}}$  für  $\alpha$ 
     $C = \frac{\Delta \bar{t}_{\min}}{e^{\alpha} - 1}$ 
     $t_N = C \cdot (e^{\alpha \cdot N} - 1)$ 
    if  $\bar{T}_{\min} \leq t_N \leq \bar{T}_{\max}$ 
        Berechne:  $t_n = C \cdot (e^{\alpha \cdot n} - 1), n = 0, \dots, N$ 
        return
    else
         $N = N + 1$ 
        if  $N = N_{\max}$ 
             $\rightarrow$  Fehler, kein  $N$  gefunden
        end
    end
end
    
```

Algorithmus 7.4: Implementierungsvorschrift für eine adaptive zeitliche Diskretisierung mittels eines Exponentialansatzes

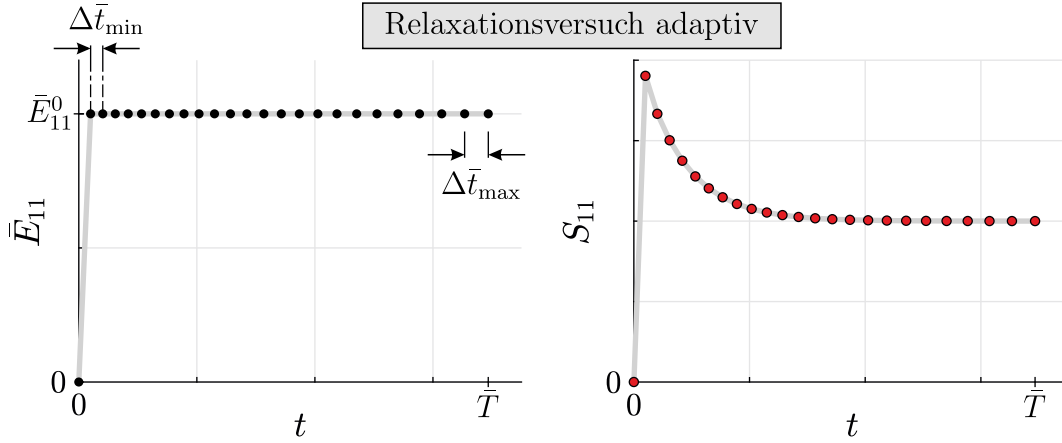


Abbildung 7.3: Relaxationsversuch mit adaptiver Zeitschrittweite

fahren gegenübergestellt, das vollständig auf zufällig erzeugten Verzerrungspfaden basiert. Ein solcher Pfad lässt sich durch die Parameter

$$\Delta \bar{t}, \bar{T}, \bar{E}_{11}^{\min}, \bar{E}_{11}^{\max} \quad (7.17)$$

beschreiben. Für einen Pfad $\mathcal{P}_\alpha$ wird mit konstanten Zeitschrittweiten $\Delta \bar{t}$ über die Versuchsdauer $\bar{T}$ an jedem Zeitschritt n eine Verzerrung aus einer Gleichverteilung im Intervall $[\bar{E}_{11}^{\min}, \bar{E}_{11}^{\max}]$ gezogen. Zur Erzeugung ganzer Pfadscharen

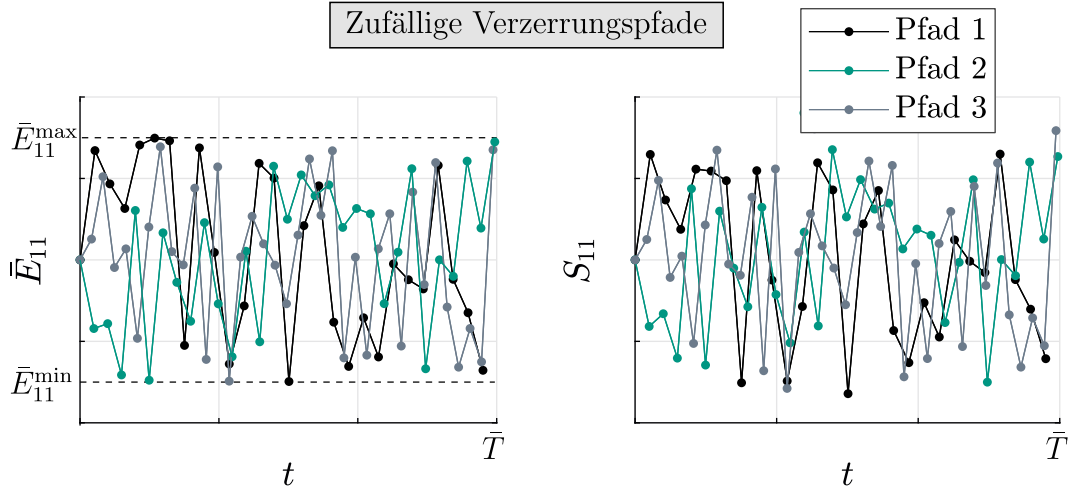


Abbildung 7.4: Darstellung von drei zufälligen Verzerrungspfaden

werden zusätzliche Parameter

$$n_{\text{Pfade}}, \Delta \bar{t}_{\min}, \Delta \bar{t}_{\max}, \bar{T}_{\min}, \bar{T}_{\max}, \bar{E}_{11}^{\min}, \bar{E}_{11}^{\max}, \quad (7.18)$$

festgelegt. Für jeden Pfad $n = 1, \dots, n_{\text{Pfade}}$ werden $\Delta \bar{t}$ und $\bar{T}$ aus den vorgegebenen Intervallen $[\Delta \bar{t}_{\min}, \Delta \bar{t}_{\max}]$ bzw. $[\bar{T}_{\min}, \bar{T}_{\max}]$ gewählt. Abbildung 7.4 zeigt exemplarisch drei solcher zufälligen Verzerrungspfade für den Fall $n_{\text{Pfade}} = 3$.

7.2.4 Grenzen des viskoelastischen Trainingsraums

Selbst für ein gut trainiertes KNN ist die Generalisierungsfähigkeit auf den Bereich des Trainingsraums beschränkt. Eine Überschreitung dieser Grenzen würde Extrapolation erfordern, die gewöhnliche Feedforward-Netze nicht leisten können. Daher ist die Festlegung des Trainingsraums ein zentraler Aspekt, da das Modell in der praktischen Anwendung auf diesen Bereich begrenzt bleibt.

Begrenzungen müssen ausschließlich für die tatsächlichen Eingangsgrößen des Netzes festgelegt werden – also für die Verzerrungen $\mathbf{E}$, die Spannungen $\mathbf{S}$ und das Zeitinkrement Δt . Größen, die nur abgeleitet oder vorverarbeitet werden, wie etwa das Verzerrungsinkrement, benötigen dagegen keine eigenen Schranken. Umgekehrt können auch Größen, die nicht direkt ins KNN eingehen (z. B. die Gesamtdauer eines Versuchs), dennoch mit Intervallen vorgegeben werden, etwa um die Datengenerierung zu steuern.

Eine Besonderheit stellen die Spannungen dar: Sie sind sowohl Eingangs- als auch Ausgangsgrößen. Theoretisch müssten sie daher nicht separat beschränkt werden. Da jedoch in dieser Arbeit auch spannungsgesteuerte Kriechversuche betrachtet werden, ist sicherzustellen, dass die gewählten Spannungsbereiche so gewählt sind,

dass die daraus resultierenden Verzerrungen innerhalb der vorgegebenen Verzerrungsgrenzen bleiben.

Damit werden zunächst die Verzerrungen diskutiert. Da das Modell auf lineare Viskoelastizität beschränkt ist, wird die Annahme großer Verformungen bei gleichzeitig kleinen Verzerrungen getroffen. Der Begriff „klein“ wird dabei in der Literatur unterschiedlich gehandhabt und ist weiterhin häufig auch abhängig vom jeweiligen Beispiel. Im Rahmen dieser Arbeit wird daher das Kriterium

$$||\mathbf{E}|| \leq 0,03 \quad (7.19)$$

gewählt. Dieses Kriterium lässt sich ohne weitere Überlegungen sowohl auf ein- als auch auf dreidimensionale Materialmodelle anwenden.

Eine wichtige Ergänzung ergibt sich bei der Kinematik von Strukturelementen, insbesondere bei Schalenelementen. Durch die Reduktion auf ein zweidimensionales Kontinuum treten neben den Membran- und Schubverzerrungen auch Krümmungsanteile auf (vgl. Gleichung (2.73)). Unter Berücksichtigung der Beziehung zwischen Kontinuums- und Schalenverzerrungen in Gleichung (5.10) ergibt sich die Bedingung

$$E_{ij} = \varepsilon_{ij} + z_{\max} \kappa_{ij} \leq 0,03, \quad i, j = 1, 2. \quad (7.20)$$

Exemplarisch reduziert sich diese Bedingung in Abwesenheit der Membrananteile ε_{ij} bei uniaxialen Versuchen, sowie mit $z_{\max} = h/2$, zu

$$\kappa_{ij} \leq 2 \cdot \frac{0,03}{h}. \quad (7.21)$$

Damit wird deutlich, dass die Krümmungsanteile dimensionsbehaftet sind und direkt von der Schalenhöhe h abhängen. Dies stellt jedoch keine Einschränkung des Modells dar, sondern ist lediglich ein Aspekt, der bei der Generierung der Trainingsdaten berücksichtigt werden muss.

Letztlich verbleibt die Wahl der Grenzen für die Zeitschrittweiten. Im Gegensatz zu den Verzerrungen lassen sich hierfür keine materialunabhängigen Grenzwerte angeben⁴⁹. In dieser Arbeit wird daher ein pragmatischer Ansatz gewählt, der sich am Verhalten des Materials in Relaxations- und Kriechversuchen orientiert. Da die Versuchsdauer $\bar{T}$ – vgl. Kapitel 7.2.1 – so festgelegt wird, dass die elastische Gleichgewichtsantwort approximativ erreicht ist, kann hieraus eine minimale und maximale Anzahl an Zeitschritten zur Diskretisierung der Zeitachse abgeleitet werden. Analog zur adaptiven Strategie aus Kapitel 7.2.2 werden hierfür in der Regel mindestens 20 und maximal 80 Inkremente verwendet.

⁴⁹ Ähnliche Überlegungen wie für die Verzerrungen lassen sich grundsätzlich auch auf deren Raten übertragen. Im Wesentlichen gilt es dabei, quasi-statische von dynamischen Prozessen zu unterscheiden, vgl. Kapitel 2.2.1. Einen Überblick hierzu geben ZHANG und ZHAO [154].

Die Wahl der Zeitschrittweite Δt als zusätzliche Eingangsvariable stellt einen wesentlichen Nachteil des KNN-Materialmodells gegenüber dem analytischen verallgemeinerten MAXWELL-Modell dar, das prinzipiell für beliebige Zeitschrittweiten⁵⁰ einsetzbar ist. Dem Vorteil, dass das hier entwickelte explizite KNN-Modell im Gegensatz zu den in Kapitel 7.1.3 behandelten impliziten Verfahren ohne Iteration auskommt, steht somit der Nachteil der expliziten Zeitschrittweitenabhängigkeit gegenüber⁵¹.

7.2.5 Synthetische Materialversuche

Wäre die Arbeit auf eindimensionales Materialverhalten beschränkt, würden die Ausführungen der Kapitel 7.2.1-7.2.4 bereits den Rahmen zur Erzeugung der Trainingsdaten abstecken. Da nur E_{11} und S_{11} existieren, entfällt jede weitere Festlegung für zusätzliche Komponenten von $\mathbf{E}$ und $\mathbf{S}$. Die Versuche sind hier zwangsläufig *uniaxial*. Im mehrdimensionalen Fall hingegen muss jede Komponente von $\mathbf{E}$ und $\mathbf{S}$ explizit vorgegeben werden. Dadurch wächst die Zahl der möglichen Kombinationen exponentiell mit der Zahl der Spannungs- bzw. Verzerrungskomponenten n_{str} an. Bereits im dreidimensionalen Fall ergeben sich so $2^{n_{\text{str}}} = 64$ Varianten, bei Schalenmodellen sogar 256.

Neben den bereits motivierten Versuchsarten (Relaxation, Kriechen, zyklische Versuche) wird auch bei der Vorgabe der Komponenten auf experimentelle Herangehensweisen zurückgegriffen, was den Kreis der Definition des Begriffs *synthetische Materialversuche* schließt. Aufbauend auf den Ausführungen in Kapitel 3.4, werden entsprechende Verzerrungskomponenten vorgegeben und die verbleibenden Komponenten über die Annahme berechnet, dass nur die Spannungskomponenten in der vorgegebenen Richtung ungleich Null sind. Dies können beispielsweise uniaxiale oder biaxiale Versuche sein. Um die Wahl der Pfade weiter zu vereinfachen, werden für das *Training* nur uniaxiale Versuche verwendet. Auch hier greift die Motivation, dass das Materialmodell erneut auf experimentelle Daten - unter Kenntnis gemessener Querkontraktionskomponenten - anwendbar ist, die häufig nur uniaxial durchgeführt werden.

Die Durchführung eines synthetischen Materialversuchs setzt die Existenz eines bekannten Materialmodells voraus, welches entweder analytischer Natur sein

⁵⁰Für sehr kleine Zeitschrittweiten $\Delta t \rightarrow 0$ können numerische Probleme aufgrund der Division durch Null auftreten. Diese lassen sich jedoch zum einen durch eine TAYLOR-Reihenapproximation umgehen und sind zum anderen der numerischen Umsetzung, nicht dem analytischen Modell selbst geschuldet (Details s. WRIGGERS [152]).

⁵¹Dadurch, dass die Spannungs- und Verzerrungsraten ebenfalls Teil des Eingangsvektors sind, wird die Zeitschrittweite in dem Modell von JUNG und GHABOUSSI [62] zumindest auch implizit berücksichtigt. Indirekt ist dieses Modell damit ebenfalls an Grenzen für die Zeitschrittweite gekoppelt.

kann (vgl. Kapitel 3) oder auf numerischen Homogenisierungsmethoden beruht (vgl. Kapitel 5.3). In beiden Fällen kann dann für einen Zeitpunkt t und einen gegebenen Verzerrungszustand $\mathbf{E}(t) = [\mathbf{E}^G, \mathbf{E}^F]^T$ eine Referenzlösung für die Spannungen $\mathbf{S}^{\text{ref}}(\mathbf{E}, t) = [\mathbf{S}^G, \mathbf{S}^F]^T$ und die Materialtangente berechnet werden $\mathbf{C}_T^{\text{ref}}(\mathbf{E}, t)$. Die Anteile $(\dots)^G$ bezeichnen dabei die vorgegebenen und $(\dots)^F$ die freien - noch unbekannten - Anteile. Auf Basis dessen wird ein Residuum

$$\mathbf{R}^V = \bar{\mathbf{S}}^{\text{ref}}(\mathbf{E}, t) - \mathbf{S}^{\text{ref}}(\mathbf{E}, t) = \mathbf{0} \quad (7.22)$$

definiert, wobei der hochgestellte Index V die Verzerrungssteuerung notiert. Weiterhin gilt $\bar{\mathbf{S}}^{\text{ref}}(\mathbf{E}, t) = [\mathbf{S}^G, \mathbf{0}]^T$, was gerade der Bedingungen entspricht, dass die Spannungen lediglich in Richtung der vorgegebenen Verzerrungskomponenten ungleich Null sind. Es verbleiben die unbekannten freien Komponenten $\mathbf{E}^F$, die unter Verwendung des NEWTON-Verfahrens berechnet werden können. Dafür wird die Gleichung (7.22) zunächst linearisiert

$$L[\mathbf{R}^V] = \mathbf{R}^V + \Delta \mathbf{R}^V = [\bar{\mathbf{S}}^{\text{ref}}(\mathbf{E}, t) - \mathbf{S}^{\text{ref}}(\mathbf{E}, t)] + \mathbf{C}_T^{\text{ref}}(\mathbf{E}, t) \Delta \mathbf{E} = \mathbf{0}, \quad (7.23)$$

was auf das zugehörige Gleichungssystem

$$\begin{bmatrix} \mathbf{S}^G \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{S}^G \\ \mathbf{S}^F \end{bmatrix} - \begin{bmatrix} \mathbf{C}_T^{GG} & \mathbf{C}_T^{GF} \\ \mathbf{C}_T^{FG} & \mathbf{C}_T^{FF} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{E}^G \\ \Delta \mathbf{E}^F \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad (7.24)$$

führt. Da die vorgegeben Verschiebungen fest sind, gilt $\Delta \mathbf{E}^G = \mathbf{0}$ womit das Gleichungssystem direkt nach den gesuchten Verschiebungsinkrementen

$$\Delta \mathbf{E}^F = -(\mathbf{C}_T^{FF})^{-1} \mathbf{S}^F(\mathbf{E}, t) \quad (7.25)$$

aufgelöst und die Aktualisierung der Verzerrungen erfolgen kann. Anschließend werden die Verzerrungen zum Zeitpunkt t aktualisiert:

$$\mathbf{E}(t) = \begin{bmatrix} \mathbf{E}^G \\ \mathbf{E}^F + \Delta \mathbf{E}^F \end{bmatrix}. \quad (7.26)$$

Das Verfahren konvergiert im linearen Fall in einem Schritt, d.h., immer im Rahmen der Datenerzeugung, da die verwendeten analytischen Referenzlösungen rein lineares Materialverhalten abbilden. Dies ist äquivalent zu der Aussage, dass die Tangente unabhängig von den Verzerrungen ist, $\mathbf{C}_T^{\text{ref}} \neq \mathbf{C}_T^{\text{ref}}(\mathbf{E})$. Bei Verwendung einer konstanten Zeitschrittweite innerhalb eines Pfades gilt zusätzlich $\mathbf{C}_T^{\text{ref}} \neq \mathbf{C}_T^{\text{ref}}(\mathbf{E}, t)$, vgl. Kapitel 3.3.

Alternativ können auch spannungsgesteuerte Kriechversuche gefahren werden. In diesem Fall sind alle Verzerrungskomponenten frei, d.h. $\mathbf{E} = \mathbf{E}^F$. Das zugehörige Residuum auf Basis des vorgegebenen Spannungsvektors $\bar{\mathbf{S}}^{\text{ref}} \neq \bar{\mathbf{S}}^{\text{ref}}(\mathbf{E}, t)$ lautet dann

$$\mathbf{R}^S = \bar{\mathbf{S}}^{\text{ref}} - \mathbf{S}^{\text{ref}}(\mathbf{E}, t) = \mathbf{0} \quad (7.27)$$

sowie die entsprechenden Linearisierung

$$L[\mathbf{R}^S] = \mathbf{R}^S + \Delta \mathbf{R}^S = \mathbf{R}^S + \mathbf{C}_T^{\text{ref}}(\mathbf{E}, t) \Delta \mathbf{E} = \mathbf{0}. \quad (7.28)$$

Die Gleichung kann direkt nach den gesuchten Verzerrungssinkrementen aufgelöst werden:

$$\Delta \mathbf{E} = -(\mathbf{C}_T^{\text{ref}}(\mathbf{E}, t))^{-1} \mathbf{R}^S. \quad (7.29)$$

Im Rahmen der Iteration folgt im linearen Fall direkt $\mathbf{E}(t) = \mathbf{E}(t-1) + \Delta \mathbf{E}$.

Anmerkungen zu Materialversuchen mit KNN

Neben der Erzeugung von Trainingsdaten mittels eines bestehenden Materialmodells können die Versuche auch als Materialtests für das KNN verwendet werden. Hierfür müssen lediglich die Spannungen $\mathbf{S}^{\text{ref}}(\mathbf{E}, t)$ sowie die Materialtangente $\mathbf{C}_T^{\text{ref}}(\mathbf{E}, t)$ durch die entsprechenden KNN Größen $\mathbf{S}^{\text{KNN}}(\mathbf{E}, t)$ und $\mathbf{C}_T^{\text{KNN}}(\mathbf{E}, t)$ ausgetauscht werden. Zu beachten ist hierbei der entscheidende Aspekt, dass die KNN-Tangente immer von der aktuellen Verzerrung abhängig ist. Dies ist der Tatsache geschuldet, dass das KNN durch die Verwendung der gewählten Aktivierungsfunktionen eine nichtlineare Abbildungsvorschrift in allen Eingangsgrößen darstellt. Entsprechend ist im Gegensatz zu der *Datenerzeugung* - unabhängig ob mittels eines analytischen Stoffgesetzes oder numerischer Homogenisierung - eine vollständige Newton-Iteration notwendig. Folglich muss ein Konvergenzkriterium definiert werden, was hier zu $\text{TOL} = 10^{-5} \|\mathbf{R}_0\|$ gewählt wird, wobei $\mathbf{R}_0$ dem initialen Residuum im Rahmen einer Iteration entspricht.

7.2.6 Synthetische Materialversuche mit Schalen-RVEs

Wie bereits angesprochen, lassen sich die synthetischen Materialversuche sowohl mit analytischen Materialmodellen als auch mit numerischen Homogenisierungsverfahren durchführen. Im Rahmen der Arbeit steht vor allem letzteres im Fokus. Des Weiteren werden die Versuche auf Basis von Schalen-RVEs durchgeführt, womit auf die Ausführung zur Berechnung der effektiven Größen in Kapitel 5 zurückgegriffen wird. Die zuvor beschriebenen synthetischen Materialversuche lassen sich in gleicher Weise anwenden, jedoch müssen die Schalenverzerrungen $\boldsymbol{\epsilon}$ sowie die Schalenschnittgrößen $\boldsymbol{\sigma}$ (vgl. Kapitel 2.4.1) verwendet werden.

Durch die Schalenkinematik ist die Übertragung der experimentellen Begrifflichkeiten, wie bspw. uniaxial oder biaxial, sinngemäß zu verstehen. Im Rahmen der Arbeit wird damit lediglich impliziert, dass beispielsweise im uniaxialen eine Verzerrungskomponente gesteuert wird, während die restlichen frei sind. Die vorgebene Komponente kann dabei sowohl Membran-, Krümmungs- oder Schubanteile

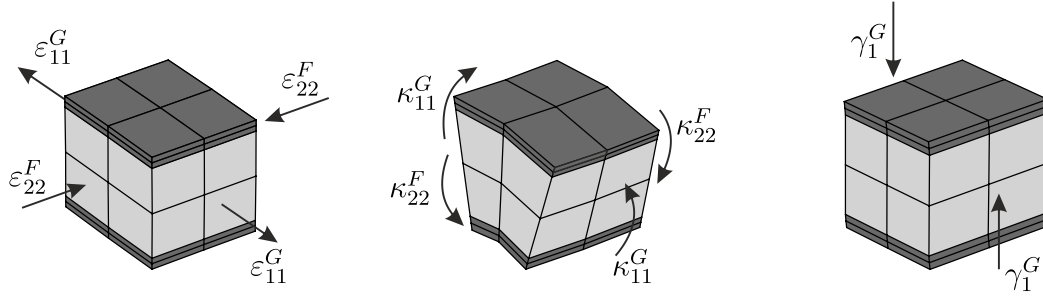


Abbildung 7.5: Exemplarische Darstellung dreier uniaxialer synthetischer Materialversuche mit einem Schalen-RVE. Eine Verzerrungskomponente (Index G) wird dabei vorgegeben und die verbleibenden (Index F) mit dem Vorgehen aus Kapitel 7.2.5 berechnet.

betreffen. Die Kopplung zu den freien Komponenten, die im Rahmen des NEWTON-Verfahrens berechnet werden, hängt von der Struktur der Materialmatrix $\mathbf{D}$ ab. Für die hier verwendeten Schalen-RVEs wird die Materialmatrix immer in der Form

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_m & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_b & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_s \end{bmatrix}, \quad (..) _m = \text{Membran}, (..) _b = \text{Biegung}, (..) _s = \text{Schub}, \quad (7.30)$$

mit den zugehörigen Submatrizen

$$\mathbf{D}_m = \begin{bmatrix} D_{11} & D_{12} & 0 \\ D_{21} & D_{22} & 0 \\ 0 & 0 & D_{33} \end{bmatrix}, \quad \mathbf{D}_b = \begin{bmatrix} D_{44} & D_{45} & 0 \\ D_{54} & D_{55} & 0 \\ 0 & 0 & D_{66} \end{bmatrix}, \quad \mathbf{D}_s = \begin{bmatrix} D_{77} & 0 \\ 0 & D_{88} \end{bmatrix}. \quad (7.31)$$

vorliegen, was eine Entkopplung der Membran-, Biegungs- und Schubanteile mit sich bringt⁵². Dies ist dann der Fall, wenn die Schalenmittelfläche als Referenzfläche gewählt wird und der Aufbau der Mikrostruktur entsprechende Symmetrien aufweist. Aus der Besetzung der Submatrizen in Gleichung (7.31) lassen sich weitere Entkopplungen erkennen, womit beispielsweise die Vorgabe einer Schubverzerrungskomponente γ_1^G für einen uniaxialen Versuch keine weiteren Verzerrungen hervorruft. In Abbildung 7.5 sind zusätzlich zu dem zuvor diskutierten Fall exemplarisch zwei Versuche unter Vorgabe einer Membrankomponente ε_{11}^G sowie einer Krümmung κ_{11}^G und dem daraus resultierenden Verformungszustand qualitativ dargestellt.

Für die Generierung der Trainingsdaten auf dem Schalen-RVE wird auf das in

⁵²Dies stellt keine grundsätzliche Einschränkung der vorgestellten Methode dar, womit sich das Vorgehen zur Datengenerierung auch auf beliebig anisotrope Mikrostrukturen übertragen lässt.

eine erweiterte Version von FEAP [128] implementierte numerische Homogenisierungsschema aus Kapitel 5.3 zurückgegriffen. Entsprechend werden die im Rahmen der Materialversuche benötigten Streckenschnittgrößen $\boldsymbol{\sigma}^{\text{ref}}(\boldsymbol{\varepsilon}, t)$ und die effektive Materialmatrix $\mathbf{D}^{\text{ref}}(\boldsymbol{\varepsilon}, t)$ durch Aufbringen der Schalenverzerrungen $\boldsymbol{\varepsilon}(t) = [\boldsymbol{\varepsilon}^G, \boldsymbol{\varepsilon}^F]^T$ in Form der periodischen Randbedingungen (vgl. Kapitel 5.2) berechnet. Dies erfolgt über ein Eingabefile, das für jeden Zeitschritt t_n sowie theoretisch im allgemein nichtlinearen Fall auch in jeder NEWTON-Iteration modifiziert werden muss. Da hier nur ein Iterationsschritt notwendig ist, können direkt die Verzerrungen aktualisiert sowie die Geschichtsvariablen abgespeichert werden, welche für den folgenden Zeitschritt t_{n+1} benötigt werden.

7.3 Nebenbedingungen für viskoelastisches Materialverhalten

Zum Abschluss des Kapitels sollen noch Möglichkeiten zur Anreicherung der aus den synthetischen Materialversuchen resultierenden Spannungs-Verzerrungs-Daten im Rahmen des KNN-Trainings diskutiert werden. Die Verfügbarkeit der Materialtangente - bzw. Materialmatrix - stellt die naheliegende Grundlage des SOBOLEV-Trainings [26] für viskoelastische KNN-Materialmodelle. Formal werden einige weitere Nebenbedingungen diskutiert, die in erster Linie aus Gründen der Vollständigkeit genannt werden, jedoch auch einen möglichen Ausblick für zukünftige Arbeiten bieten.

7.3.1 Sobolev Training - Einbindung der Materialtangente

Aus den Gleichungen (7.23) und (7.28) wird deutlich, dass bei synthetischen Materialversuchen automatisch auch Tangenteninformationen vorliegen⁵³. Bei analytischen Referenzmodellen entstehen die Tangentenwerte als Nebenprodukt des NEWTON-Verfahrens, während sie im Rahmen der numerischen Homogenisierung ohnehin berechnet werden, da sie für die FEM auf makroskopischer Skala benötigt werden.

Diese zusätzlichen Informationen werden im Folgenden mit den Ausführungen aus Kapitel 6.3.2 für das KNN-Training berücksichtigt. Die allgemeine Darstellung des SOBOLEV-Fehlerterms aus Gleichung (6.110) konkretisiert sich in diesem Fall zu

$$\mathcal{F}^S(\mathbf{w}) = \frac{\rho}{2PS} \sum_{k=1}^{PS} \sum_{j=1}^{n_{\text{str}}} \sum_{i=1}^{n_{\text{str}}} \left(\frac{\partial \hat{z}_j(\mathbf{x}_k; \mathbf{w})}{\partial \hat{x}_i} - \hat{C}_{T_{ji}}(\mathbf{x}_k^C) \right)^2. \quad (7.32)$$

⁵³Die Berechnung der Materialtangente verursacht in der Regel keinen nennenswerten Mehraufwand. Unabhängig von der Art der Datengenerierung ist es daher sinnvoll, diese Information zu nutzen.

An jedem Punkt $\mathbf{x}_k^C$, an dem die Materialtangente eingefordert wird, stehen somit $n_{\text{str}} \times n_{\text{str}}$ Nebenbedingungen zur Verfügung. Die Einbindung im Rahmen des KNN-Trainings erfolgt mit den Berechnungsvorschriften aus Kapitel 6.3.3.

Physikalisch betrachtet fordern die Tangenten-Nebenbedingungen im rein elastischen Fall schwach die Energieerhaltung des Materialmodells ein (s. z.B. WEBER ET AL. [146]). Wie bereits in Kapitel 7.1.3 diskutiert, muss bei viskoelastischen Materialien die Dissipationsungleichung erfüllt sein, um thermodynamische Konsistenz zu garantieren. In diesem Fall kann dahingehend keine weitere physikalische Interpretation vorgenommen werden. Unabhängig davon stellt die Nebenbedingung aber einen wesentlichen zusätzlichen Informationsgehalt dar, der zumindest die Approximationseigenschaften des KNNs signifikant verbessert (vgl. Kapitel 6.3.4). Auch wenn dahingehend keine strukturelle Garantie besteht, wird die Wahrscheinlichkeit für unphysikalisches Verhalten damit reduziert. Dieser Aspekt ist auch Teil der Untersuchungen in Kapitel 8.

7.3.2 Diskussion weiterer Nebenbedingungen

Abschließend werden weitere Nebenbedingungen vorgestellt, welche beim KNN-Training mit viskoelastischen Daten nach Gleichung (7.1) berücksichtigt werden können. Sie sind insbesondere dann nützlich, wenn nur begrenzte Datenmengen vorliegen, wie beispielsweise wenn das Materialmodell auf experimenteller Basis aufgebaut werden soll.

Normalisierungs-Nebenbedingung

Die Forderung nach einer spannungsfreien Referenzkonfiguration in Gleichung (3.30) lässt sich für das KNN durch die Nebenbedingungen

$$h_{kj}^0 = z_j(\mathbf{x}_k^0; \mathbf{w}) = 0, \quad j = 1, \dots, n_{\text{str}} \quad (7.33)$$

mit

$$\mathbf{x}_k^0 = [\mathbf{0}, \mathbf{0}, \mathbf{0}, \Delta t, \mathbf{p}] \quad \text{und} \quad k = 1, \dots, P^0 \quad (7.34)$$

berücksichtigen. Die Forderung führt somit auf n_{str} Nebenbedingungen für jeden Datenpunkt $\mathbf{x}_k^0$. Mit der Transformationsbeziehung (6.5) folgt die Darstellung in der normalisierten Form

$$\hat{h}_{kj}^0 = \hat{z}_j(\mathbf{x}_k^0; \mathbf{w}) + 2 \left(\frac{t_{j,\min}}{t_{j,\max} - t_{j,\min}} \right) + 1 = 0 \quad (7.35)$$

die durch den zusätzlichen Fehlerterm

$$\mathcal{F}^0 = \frac{\rho}{2P^0} \sum_{k=1}^{P^0} \sum_{j=1}^{n_{\text{str}}} (\hat{h}_{kj}^0)^2 = \frac{\rho}{2P^0} \sum_{k=1}^{P^0} \|\hat{\mathbf{h}}_k^0\|^2 \quad (7.36)$$

im Rahmen des KNN-Trainings berücksichtigt werden können. Der dazugehörige Gradient $\nabla_{\mathbf{w}} \mathcal{F}^0$, bzw. die für das LMV benötigte JACOBI-Matrix $\nabla_{\mathbf{w}} \mathbf{h}^0$ lassen sich mit den Herleitungen der Kapitel 6.1.3 und 6.2.3 bestimmen. Darüber hinaus erzwingt die Nebenbedingung in schwacher Form eine zeitliche Invarianz in der Form, dass die initiale Antwort des Materials im spannungsfreien Zustand unabhängig von der Zeitschrittweite bleibt.

Einforderung der nicht-negativen Energiedissipation

Die Forderung der Dissipationsungleichung $\mathcal{D} \geq 0$ kann physikalisch auch so interpretiert werden, dass keine Energieproduktion innerhalb von zwei sequentiellen Belastungsschritten gleicher Größe aber unterschiedlicher Vorzeichen stattfinden darf. Dies lässt sich inkrementell auch schreiben als

$$\Delta D = \frac{1}{2} \Delta \mathbf{E}_{n+1}(\mathbf{S}_n + \mathbf{S}_{n+1}) - \frac{1}{2} \Delta \mathbf{E}_{n+1}(\mathbf{S}_{n+1} + \mathbf{S}_{n+2}) \geq 0. \quad (7.37)$$

Die Form der Nebenbedingung wurde erstmals in WEBER ET AL. [147] für die Modellierung ratenunabhängiger Plastizität verwendet und von DETTMER ET AL. zur Validierung des Modells eingesetzt⁵⁴. Da es sich hierbei um eine Ungleichheits-Nebenbedingung handelt, sind weitere Ausführung zur Behandlung dieser im Kontext des KNN-Trainings notwendig womit auf die Herleitungen in [147] verwiesen wird. Auch hier wird der thermodynamische Kontext nur schwach berücksichtigt, sodass keine strukturelle Garantie für die Einhaltung der Forderung besteht.

Abschließender Kommentar

Neben den zuvor diskutierten Nebenbedingungen sind grundsätzlich Weitere denkbar. Dabei ist zu berücksichtigen, dass solche Bedingungen vor allem dann sinnvoll sind, wenn sie einen klaren Mehrwert an Informationsgehalt für das Training liefern. So stellt etwa die Einforderung der Materialtangente zusätzlichen Informationsgehalt bereit, der in den primären Spannungs-Verzerrungsdaten nicht enthalten ist. Stehen dagegen ausreichend viele Daten zur Verfügung, ist nicht jede Nebenbedingung zwingend erforderlich. Jeder zusätzliche Fehlerterm beeinflusst das KNN-Training und kann dazu führen, dass die Nebenbedingungen besser approximiert werden als der eigentlich zugrunde liegende funktionale Zusammenhang. Die Entscheidung über die Notwendigkeit zusätzlicher Nebenbedingungen sollte daher stets im Kontext der jeweiligen Problemstellung erfolgen.

⁵⁴Die Autoren berücksichtigen die Forderung nicht im Rahmen des Trainings sondern werten lediglich die Gleichung bei KNN-Materialtests aus, um die thermodynamische Konsistenz des Modells zu prüfen.

8 KNN-Untersuchungen auf Materialpunkt- und Strukturebene

In diesem Kapitel wird der Einsatz eines KNN als viskoelastisches Materialmodell systematisch untersucht. Zunächst wird das Verhalten des Modells auf Materialpunktebene analysiert, um anschließend dessen Eignung für die Strukturanalyse im Rahmen der Finite-Elemente-Methode (FEM) anhand numerischer Beispiele zu bewerten. Um die grundlegenden Herausforderungen und das Vorgehen bei der Modellentwicklung nachvollziehbar darzustellen, beginnen die Untersuchungen mit eindimensionalen Problemstellungen. Darauf aufbauend werden die erforderlichen Erweiterungen für ein dreidimensionales Materialmodell beschrieben. Abschließend richtet sich der Fokus auf die Strukturebene: Auf Basis von Schalen-RVEs wird das Materialmodell konstruiert und anschließend für baupraktische numerische Anwendungsfälle eingesetzt.

8.1 Eindimensionale Untersuchungen

Viele Phänomene, die im mehrdimensionalen Fall auftreten, entfallen in einer eindimensionalen Betrachtung. Da nur eine Belastungsrichtung vorliegt, reduziert sich das Problem auf den uniaxialen Fall. Dies ermöglicht es, zunächst ein „ingenieurmäßiges Gefühl“ für die benötigte Datenmenge, geeignete KNN-Architekturen sowie den Einfluss der Datengenerierung zu entwickeln – Erkenntnisse, die später für die mehrdimensionalen Modelle von Bedeutung sind. Die vorgenommenen Schritte zur Entwicklung des Modells orientieren sich dabei stets an den Ausführungen in Kapitel 7.2.

8.1.1 Einfluss der Datengenerierung

Ein zentraler Aspekt dieser Arbeit ist die vorgeschlagene Methodik zur Erzeugung synthetischer Trainingsdaten in Anlehnung an klassische, viskoelastische Experimente. Zunächst wird untersucht, wie viele Versuche nötig sind, um eine Generalisierungsfähigkeit des KNN für unbekannte Pfade zu erreichen und die Methodik zur Bestimmung der Grenzen des Eingangsraums erläutert.

Dazu wird ein analytisches Materialmodell definiert, mit dem sich synthetische Spannungs-Verzerrungs-Daten erzeugen lassen. Verwendet wird ein verallgemeinertes MAXWELL-Modell mit einem MAXWELL-Term und den in Tabelle 8.1 angegebenen Materialparametern. Diese sind rein akademischer Natur, was die Methodik und die Aussagekraft der Untersuchungen jedoch nicht beeinflusst. Verzerrungsgesteuerte Versuche erfolgen bis zu einem maximalen Wert von $\bar{E}_{11} = 0,03$.

Initialer E-Modul E_0 [N/mm ²]	MAXWELL-Element i	Modulanteil $\mu_i = E_i/E_0$ [-]	Relaxationszeit τ_i [s]
2600	-	-	-
	1	0,50	33,33

Tabelle 8.1: Verwendete Materialparameter für das eindimensionale verallgemeinerte MAXWELL-Modell

Die Definition der Versuchsdauer $\bar{T}$ und der daraus abgeleiteten minimalen bzw. maximalen Zeitschrittweiten $\Delta\bar{t}_{\min}$ und $\Delta\bar{t}_{\max}$ erfordert eine Voruntersuchung. Hierzu wird bei maximaler zulässiger Verzerrung ein Relaxationsversuch durchgeführt (Abbildung 8.1, links). Da nur die Zeit bis zum Erreichen der Gleichgewichtsspannung von Interesse ist, ist die genaue Größe der aufgebrachten Verzerrung unerheblich. In diesem Beispiel wird das rein elastische Gleichgewicht nach $T \approx 300$ s erreicht. Für die Generierung der Datensätze wird eine Diskretisierung von mindestens 20 und höchstens 60 Zeitschritten gewählt, was $\Delta\bar{t}_{\min} = 5$ s und $\Delta\bar{t}_{\max} = 15$ s entspricht. Wie in Kapitel 7.2.4 diskutiert, erfolgt diese Wahl rein auf Basis einer „ingenieurmäßig sinnvollen“ Diskretisierung der Zeitachse und kann grundsätzlich beliebig gewählt werden.

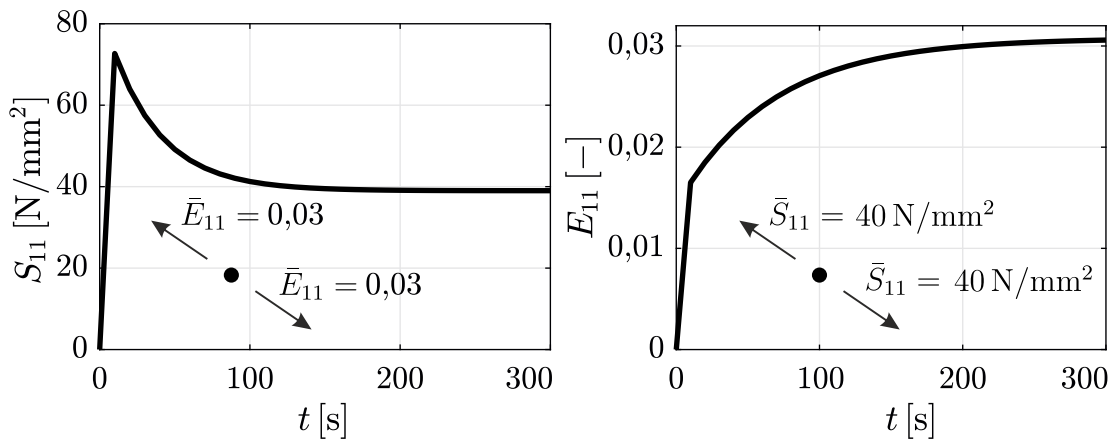


Abbildung 8.1: Relaxationsversuch (links) und Kriechversuch (rechts) im Rahmen der Voruntersuchung zur Bestimmung der Grenzen des Trainingsraums

Mit $E_\infty = (1 - \mu_1)E_0 = 1300$ N/mm² lässt sich die maximal zulässige Spannung $\bar{S}_{11}$ über $E_{11\infty} = \bar{S}_{11}/E_\infty \leq 0,03$ abschätzen, was zu einem oberen Grenzwert von $\bar{S}_{11} \approx 40$ N/mm² für spannungsgesteuerte Versuche führt. Bei verzerrungsgesteuerten Versuchen kann diese Grenzen jedoch überschritten werden. Die Materialantwort im Rahmen eines Kriechversuchs ist ebenfalls in Abbildung 8.1 (rechts)

Verzerrung $\bar{E}_{11}$ [–]	Spannung $\bar{S}_{11}$ [N/mm ²]	Versuchsdauer $\bar{T}$ [s]	Zeitschrittweite $\Delta\bar{t}$ [s]
[-0,03, 0,03]	[-40, 40]	[300, 300]	[5, 15]

Tabelle 8.2: Grenzen der Parameter zur zufälligen Erzeugung expliziter Zeit-Verzerrungspfade

dargestellt. Auch hier ist für $T \approx 300$ s ein Konvergenzverhalten der Verzerrung gegen $E_{11\infty} \approx 0,03$ erkennbar, womit die Dauer der Versuche gleich gewählt wird. Die so bestimmten Grenzen für die explizite Pfaderzeugung sind in Tabelle 8.2 zusammengefasst.

Für die Erzeugung der Trainingsdaten werden Kriech- und Relaxationsversuche sowie zyklische Versuche verwendet, bei denen die Werte aus Tabelle 8.2 für die Parameterintervalle der Pfade in Tabelle 7.2 genutzt werden. Diese Vorgehensweise wird verglichen mit der Datengenierung basierend auf zufälligen Verzerrungspaden, wie in Kapitel 7.2.3 beschrieben. Zur Validierung des Modells werden vier Testpfade verwendet, deren Parameter in Tabelle 8.3 angegeben sind. Dabei werden jeweils ein Kriechversuch, Relaxationsversuch und ein zyklischer Versuch gefahren. Zusätzlich wird ein Zugversuch bei konstanter Dehnrate, unter Berücksichtigung von Be- und Entlastung vorgegeben, was einen aus dem Training – zumindest explizit – unbekannten Pfad darstellt. Die Testpfade werden zudem mit zwei verschiedenen Zeitschrittweiten, $\Delta t = 8$ s und $\Delta t = 12$ s, diskretisiert.

Angemerkt sei dabei, dass die Kriechversuche mit dem KNN nicht spannungsgesteuert gefahren werden. Diese Methode wird nur zur Erzeugung der zugehörigen Kriechverzerrungen auf Basis des analytischen Materialmodells verwendet. Im Rahmen der Validierung werden dem KNN explizit die so erhaltenen Verzerrungen vorgegebenen, womit das KNN eine konstante Spannungsantwort liefern muss. Dies findet auch im mehrdimensionalen Fall Anwendung, womit kein NEWTON-Verfahren gemäß Kapitel 7.2.5 notwendig ist, da keine implizite Spannungssteuerung vorliegt. Diese Vorgehensweise ermöglicht, dass kein weiterer Testfehler in

Pfad Nr.	Versuchsart	Komponente	Wert	Δt [s]	T [s]
1	Zyklisch	E_{11}	-0,01, 0,01 [–]	8	300
2	Relaxation	E_{11}	0,01 [–]	8	300
3	Kriechen	S_{11}	20 [N/mm ²]	12	300
4	Zugversuch	E_{11}	0,01 [–] [1/s]	12	300

Tabelle 8.3: Verwendete Validierungspfade unter Angabe der gewählten Parameter

den Verzerrungen definiert werden muss, sondern im Rahmen des Trainings alle Fehlermaße auf Basis der Ausgangsvariablen des KNN berechnet werden können. Grundsätzlich stellt es jedoch keine Problematik dar, spannungsgesteuerte Kriechversuche mit dem KNN durchzuführen. In dieser Arbeit wird das spannungsgesteuerte Kriechverhalten direkt auf Strukturebene im Rahmen der numerischen Beispiele analysiert.

Im eindimensionalen Fall und ohne die Berücksichtigung von zusätzlichen Parametern besitzt das KNN vier Eingangs- und eine Ausgangsvariable, vgl. Gleichung (7.1). Im vorliegenden Fall hat sich eine verdeckte Schicht mit zwanzig Neuronen im Rahmen von Voruntersuchungen als geeignet erwiesen, womit $4 - 20 - 1$ die verwendete KNN-Topologie definiert. Im Folgenden werden fünf Trainings-Durchläufe mit einem auf den Materialversuchen basierenden $\text{KNN}_{\text{Versuche}}$ und mit einem auf den zufälligen Verzerrungspfaden KNN_{Rand} durchgeführt. Dabei ergeben sich durch die zufällige Pfad-Erzeugung zum einen unterschiedliche Datenmengen und zum anderen unterscheiden sich die Gewichte zum Beginn des Trainings durch die zufällige Initialisierung. Trainiert wurden maximal 100 Epochen mit dem LMV, ohne die Verwendung von Nebenbedingungen.

Die KNN-Antworten auf die vier Validierungspfade sind in Abbildung 8.2 dargestellt. Es wurden jeweils die KNN mit dem geringsten Test-Fehler aus den fünf Trainings gewählt. Beide Ansätze liefern sehr gute Approximationen der Referenzlösungen. Für das $\text{KNN}_{\text{Versuche}}$ waren pro Training drei zufällig gewählte Kriech-, Relaxations- und zyklische Versuche erforderlich, was aufgrund der variierenden zeitlichen Diskretisierung 270 bis 359 Datenpunkte ergab. Das KNN_{Rand} benötigte fünf zufällige Pfade wobei 149 bis 180 Datenpunkte erzeugt wurden. Die geringere Datenmenge erklärt sich durch die bessere Abdeckung des Eingangsraums bei zufälliger Pfaderzeugung. Es sei jedoch angemerkt, dass die im eindimensionalen Fall vorteilhafte zufälligen Verzerrungspfade in höheren Dimensionen nicht automatisch eine ausreichende Abdeckung des Eingangsraums gewährleistet – dieser Aspekt wird in den folgenden mehrdimensionalen Untersuchungen noch detailliert betrachtet. Gleichzeitig zeigt sich, dass auch mit wenigen experimentell abgeleiteten Versuchen ein zuverlässiges viskoelastisches KNN-Materialmodell erstellt werden kann. Dabei muss berücksichtigt werden, dass für die Anwendung auf echte Versuchsdaten der Einfluss der Streuung untersucht werden muss. Da im Rahmen der Arbeit das Ziel die Konstruktion von KNN-Materialmodellen auf Basis synthetischer Daten von RVEs verfolgt wird, wird dieser Aspekt im Weiteren hier nicht betrachtet.

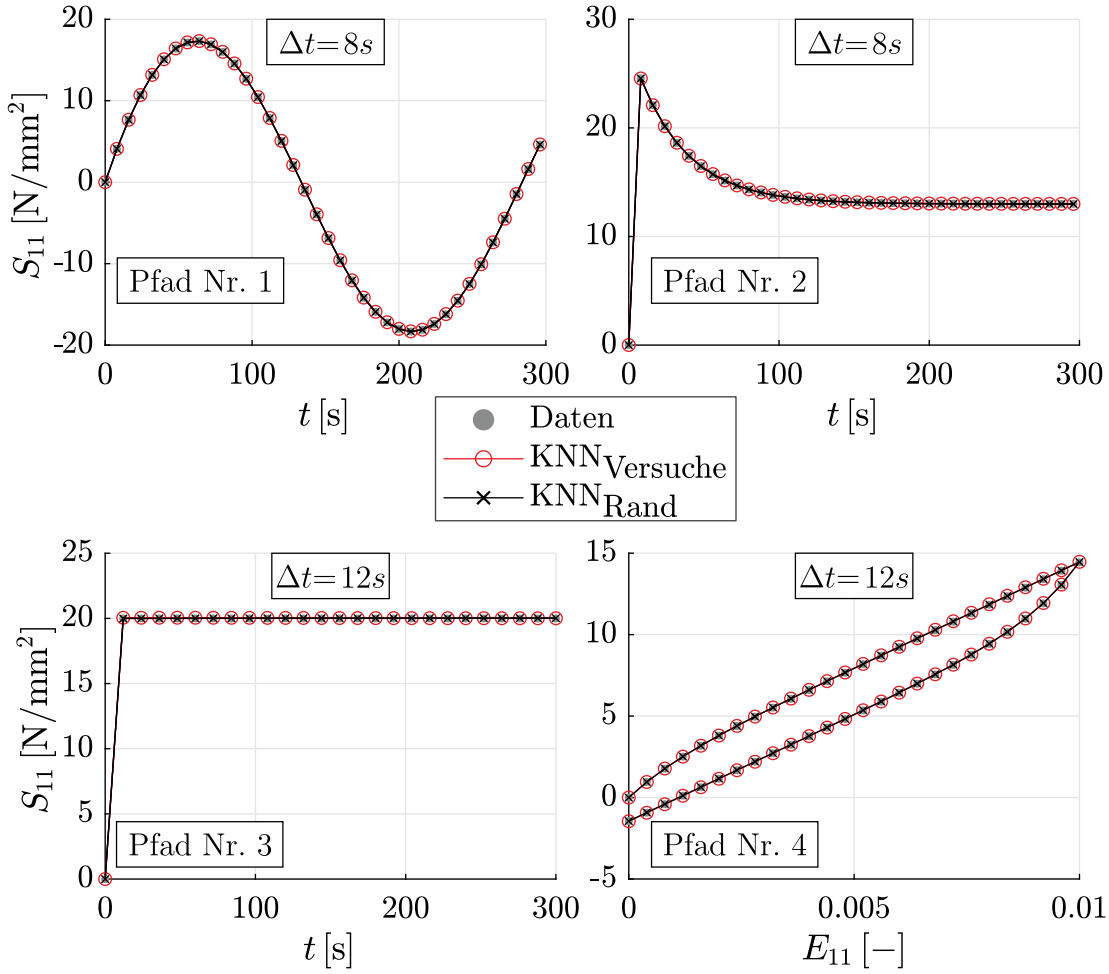


Abbildung 8.2: Darstellung der besten Approximation der in Tabelle 8.3 gegebenen Pfade mit zwei KNN, welche mit experimenteller und rein zufälliger Datenerzeugung trainiert wurden

8.1.2 Einfluss der adaptiven Zeitschrittweitenstrategie

Anstelle einer konstanten Zeitschrittweite kann die in Kapitel 7.2.2 beschriebene adaptive Strategie eingesetzt werden. Dies hat den Vorteil, dass die Zeitschrittweite $\Delta \bar{t}$ bei der Erzeugung von Pfadscharen nicht zufällig aus dem Intervall $[\Delta \bar{t}_{\min}, \Delta \bar{t}_{\max}]$ gewählt wird. Die Parameterintervalle sind dadurch nur für Verzerrungen, Spannungen und Versuchsdauern festzulegen (s. auch Tabelle 7.3). Zusätzlich zu den zuvor eingeführten KNN_Versuche und KNN_Rand wird ein weiteres KNN – das KNN_Adap – hinzugezogen, bei dem die Kriech- und Relaxationsversuche mittels adaptiver Zeitschrittweite im Intervall $[5, 15]$ s diskretisiert werden. Alle drei KNN, mit einer Topologie 4–20–1, werden erneut fünf mal trainiert und die beste Approximation für die zwei in Tabelle 8.4 dargestellten Pfade untersucht. Dabei handelt es sich um einen Relaxations- und einen Kriechversuch,

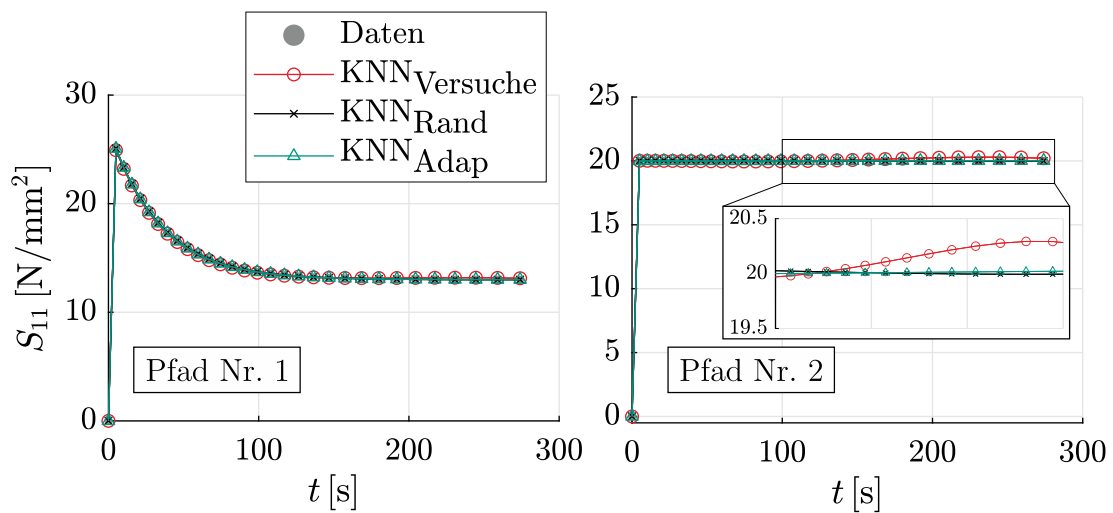
Pfad Nr.	Versuchsart	Komponente	Wert	$\Delta t_{\min}$ [s]	$\Delta t_{\max}$ [s]	T [s]
1	Relaxation adaptiv	E_{11}	0,01 [–]	5	15	300
2	Kriechen adaptiv	S_{11}	20 [N/mm ²]	5	15	300

Tabelle 8.4: Verwendete Validierungspfade unter Angabe der gewählten Parameter

die ebenfalls mit adaptiver Zeitschrittweite diskretisiert werden.

Pro Trainingsdurchlauf des KNN_{Adap} wurden zusätzlich zu zwei Kriech- und Relaxationsversuchen die zwei im vorherigen Kapitel definierten zyklischen Versuche verwendet, was in 167-207 Datenpunkte resultierte. Für das $\text{KNN}_{\text{Versuche}}$ wurden im Gegensatz zum vorherigen Kapitel ebenfalls nur zwei statt drei der beschriebenen Pfade zufällig erzeugt, was zu einem Intervall [197, 237] in ähnlicher Größenordnung führte. Die erneute Generierung von fünf zufälligen Verzerrungspfaden ergab beim KNN_{Rand} Datenpunkte im Intervall [137, 190].

Durch die adaptive Diskretisierung liegen die Zeitschrittweiten der Validierungspfade an den Rändern des Trainingsraums. Beim KNN_{Adap} sind diese per Konstruktion abgedeckt, während beim KNN_{Rand} und $\text{KNN}_{\text{Versuche}}$ aufgrund der zufälligen Wahl nicht immer eine ausreichende Abdeckung gewährleistet ist. Selbstverständlich könnten die Versuche auch an den Grenzen des Trainingsraums gefahren werden in dem die Versuche manuell korrigiert werden. Eine solche Strategie wird hier jedoch nicht verfolgt und in höheren Dimensionen wird sich zeigen, dass durch die größere Datenmenge die Ränder automatisch gut erfasst werden.

**Abbildung 8.3:** Darstellung der besten Approximation zweier Pfade mit adaptiver Zeitschrittweite und drei verschiedenen trainierten KNN

Die Ergebnisse der drei KNN sind in Abbildung 8.3 dargestellt. Sowohl das KNN_{Adap} als auch das KNN_{Rand} liefern sehr gute Approximationen der Referenzlösung. Beim $\text{KNN}_{\text{Versuche}}$ deckt die Wahl einer konstanten Zeitschrittweite den Eingangsraum nicht ausreichend ab, womit eine etwas schlechtere Approximation des Kriechversuchs festzustellen ist. Grundsätzlich sei dazu angemerkt, dass unterschiedliche gute Approximationen in den fünf Durchläufen zu verzeichnen sind und sich die anderen beiden KNN robuster verhalten. Im Rahmen der Datengenerierung auf Basis von Kriech-, Relaxations- und zyklischen Versuchen wäre in diesem Fall – und bei geringen Datenmengen – eine adaptive zeitliche Diskretisierung einer konstanten vorzuziehen.

8.1.3 Einfluss unterschiedlichen Verhaltens bei Kriechen und Relaxation

Für das zuvor eingeführte analytische Materialmodell wurde nur ein MAXWELL-Element mit der Relaxationszeit τ_1 gewählt. In diesem Fall ist ein ähnliches Abklingverhalten für Kriechen und Relaxation zu verzeichnen. Durch Hinzufügen weiterer Elemente mit jeweiligen Relaxationszeiten τ_i können sich dagegen scheinbar unterschiedliche Zeitskalen für Kriech- und Relaxationsverhalten ergeben⁵⁵. Im Folgenden wird daher ein analytischen Modell mit drei MAXWELL-Elementen zur Datengenerierung verwendet, mit den zugehörigen Materialparameter aus Tabelle 8.5. Wie in Kapitel 8.1.1 wird ein Relaxationversuch mit $\bar{E}_{11} = 0,03$ am Materialpunkt gefahren, der in Abbildung 8.4 dargestellt ist.

Durch den kleineren Gleichgewichts-Modul $E_\infty = (1 - \sum_{i=1}^3 \mu_i) E_0 = 650 \text{ N/mm}^2$ ist ein deutlich viskoseres Verhalten zu erkennen, charakterisiert durch Erreichen der Gleichgewichtsspannung $S_{11\infty} = E_\infty \cdot \bar{E}_{11} \approx 20 \text{ N/mm}^2$ nach $T \approx 1000 \text{ s}$. Ebenfalls abgebildet ist ein Kriechversuch für eine vorgegebene Spannung $\bar{S}_{11} = 20 \text{ N/mm}^2$. Die Gleichgewichts-Verzerrung $E_{11\infty}$ wird hier nach $T \approx 3000 \text{ s}$ näherungsweise erreicht. Für dieses Material ist es dahingehend sinnvoll, die Versuchsdauern für die Generierung der synthetischen Daten entsprechend des Abklingverhaltens individuell zu wählen.

Dafür werden erneut die drei Netze – das $\text{KNN}_{\text{Versuche}}$, KNN_{Adap} und KNN_{Rand} – aus Kapitel 8.1.2 bezüglich ihres Approximationsverhalten untersucht. Das $\text{KNN}_{\text{Versuche}}$ wird mit Relaxations-, Kriech- und zyklische Versuchen mit den in Tabelle 8.6 aufgeführten Parameterintervallen und Pfadzahlen trainiert. Pro Versuch werden jeweils 15 Pfade innerhalb der Parameterintervalle generiert, was bei

⁵⁵Die zugrunde liegenden Zeitkonstanten sind für beide Phänomene identisch, dahingehend wird hier der Begriff „scheinbar“ verwendet. Die zusätzliche Überlagerung mehrerer MAXWELL-Terme führt jedoch zu unterschiedlichen Amplituden und damit zu einer veränderten Wahrnehmung der dominanten Zeitskala.

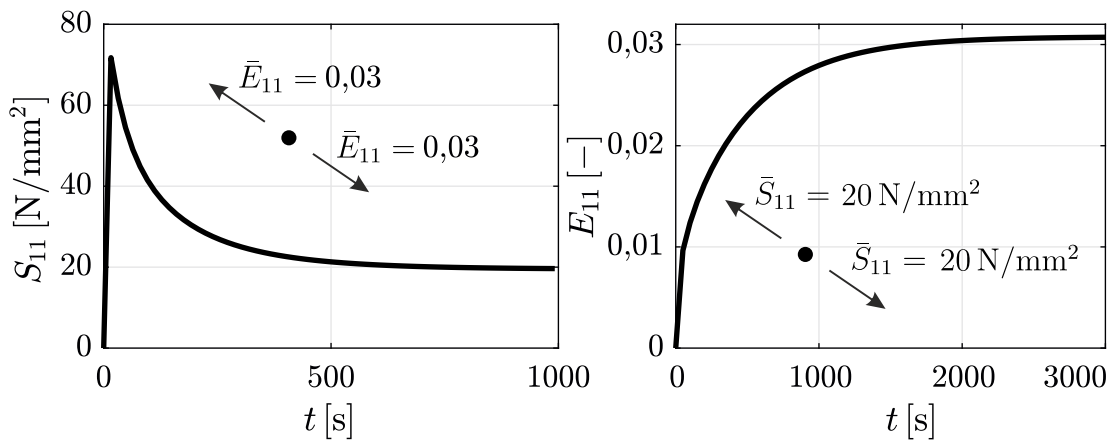


Abbildung 8.4: Relaxationsversuch (links) und Kriechversuch (rechts) im Rahmen der Voruntersuchung zur Bestimmung der Grenzen des Trainingsraums und Parameterintervalle für die Versuche

Initialer E-Modul E_0 [N/mm^2]	MAXWELL-Element i	Modulanteil $\mu_i = E_i/E_0$ [-]	Relaxationszeit τ_i [s]
2600	-	-	-
	1	0,25	33,33
	2	0,25	100,00
	3	0,25	200,00

Tabelle 8.5: Verwendete Materialparameter für das eindimensionale verallgemeinerte MAXWELL-Modell mit drei MAXWELL-Elementen

fünf Trainingsläufen insgesamt 1440 – 1593 Datenpunkte ergibt. Das KNN_{Adap} nutzt dieselben Parameterbereiche, jedoch mit adaptiver zeitlicher Diskretisierung für Relaxations- und Kriechversuche. Dadurch entstehen 1342 – 1807 Datenpunkte. Für das KNN_{Rand} werden die in Tabelle 8.7 angegebenen Bereiche und $n_{\text{Pfade}} = 41$ gewählt, was zu 1193 – 1697 Datenpunkten führt und damit eine mit den anderen Netzen vergleichbare Datenmenge liefert.

Die zur Untersuchung des Approximationsverhaltens verwendeten Validierungspfade sind in Tabelle 8.8 angegeben. Neben den bereits eingesetzten Relaxations-, Kriech- und zyklischen Versuchen wird diesmal zusätzlich ein zufälliger Verzerungspfad vorgegeben. Alle Pfade werden mit konstanter Zeitschrittweite diskretisiert. Für alle drei Netze wird dieselbe KNN-Architektur 4 – 30 – 1 verwendet. Beim Training mit dem LMV-Algorithmus waren jeweils höchstens 100 Epochen bis zur Konvergenz erforderlich.

Die Approximation der drei KNN, die über fünf Trainingsdurchläufe die besten

Versuchstyp	n_{Pfade}	Parameterintervalle
Relaxationsversuch	15	$\Delta \bar{t} \in [15, 50] \text{ s}, \bar{T} \in [1000, 1000] \text{ s},$ $\bar{E}_{11} \in [-0,03, 0, 03]$
Kriechversuch	15	$\Delta \bar{t} \in [50, 150] \text{ s}, \bar{T} \in [3000, 3000] \text{ s},$ $\bar{S}_{11} \in [-20, 20] \text{ N/mm}^2$
Zyklischer Versuch	15	$\Delta \bar{t} \in [15, 150] \text{ s}, \bar{T} \in [1000, 3000] \text{ s},$ $\bar{E}_{11} \in [-0,03, 0, 03], 2\bar{\omega}/\pi \in [1, 2]$

Tabelle 8.6: Pfadparameter der Datengenerierung für das KNN_{Versuche}

Versuchstyp	n_{Pfade}	Parameterintervalle
Zufällig	41	$\Delta \bar{t} \in [15, 150] \text{ s}, \bar{T} \in [1000, 3000] \text{ s},$ $\bar{E}_{11} \in [-0,03, 0, 03]$

Tabelle 8.7: Pfadparameter der Datengenerierung für das KNN_{Rand}

Ergebnissen erzielen, sind in Abbildung 8.5 dargestellt. Sowohl das KNN_{Versuche} als auch das KNN_{Adap} zeigen mit der vorgeschlagenen Datengenerierungsstrategie eine sehr gute Übereinstimmung mit den Referenzdaten. Für den Relaxationsversuch lässt sich sogar eine etwas bessere Nachbildung des anfänglichen Abklingverhaltens durch das KNN_{Versuche} beobachten. Beide Netze können zudem auch zufällige – im Training nicht enthaltene – Verzerrungspfade zufriedenstellend abbilden.

Das KNN_{Rand} liefert hingegen nur für den zufälligen Verzerrungspfad (Pfad 4) eine akzeptable Approximation. Weder das Abklingverhalten des Relaxations- noch die Abbildung des Kriechversuchs wird hinreichend erfasst. Selbst eine Verdopplung der Trainingsdatenmenge brachte hier keine spürbare Verbesserung. Dies unterstreicht, dass eine rein zufällige Pfaderzeugung ohne materialbezogene

Pfad Nr.	Versuch	Komponente	Wert		Δt [s]	T [s]
1	Relaxation	E_{11}	0,01	[-]	40	1000
2	Kriechen	S_{11}	10	[N/mm ²]	100	3000
3	Zugversuch	E_{11}	0,01	[-]	100	1000
4	Zufällig	E_{11}	±0,01	[-]	75	1500

Tabelle 8.8: Verwendete Validierungspfade unter Angabe der gewählten Parameter

nes Wissen für komplexere viskoelastische Vorgänge nicht ausreicht, da hier die relevanten Zeitbereiche der kritischen Phasen nicht systematisch abgedeckt werden.

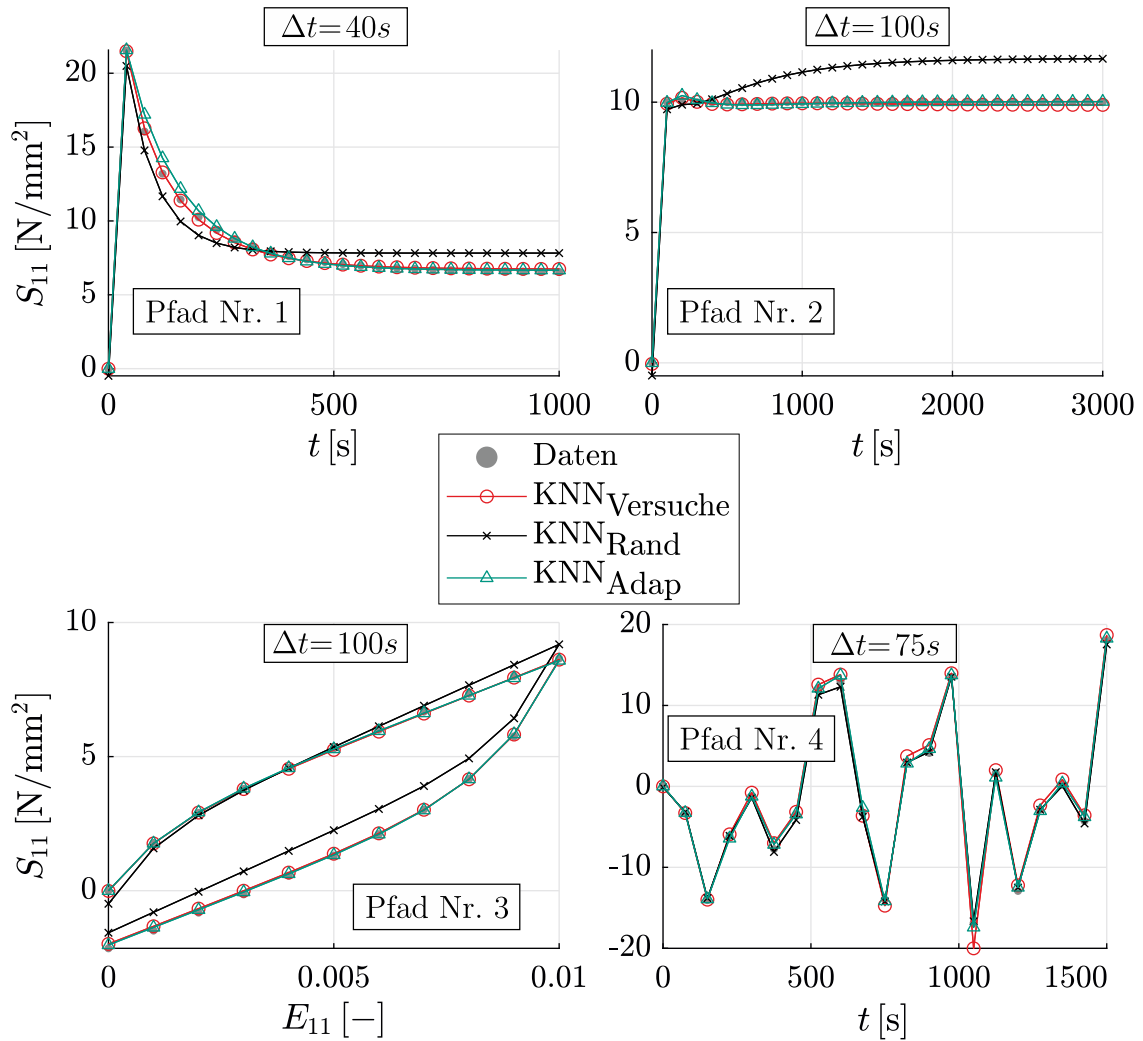


Abbildung 8.5: Darstellung der besten Approximation der in Tabelle 8.8 gegebenen Pfade mit drei KNN, welche mit experimenteller, experimentell adaptiv und rein zufälliger Datenerzeugung trainiert wurden

8.1.4 Berücksichtigung von Materialparametern als Eingangsvariablen

Zum Abschluss der eindimensionalen Untersuchungen am Materialpunkt wird geprüft, inwieweit sich Materialparameter direkt als Eingangsgrößen des KNN

Initialer E-Modul E_0 [N/mm ²]	MAXWELL-Element i	Modulanteil $\mu_i = E_i/E_0$ [-]	Relaxationszeit τ_i [s]
[100, 500]	-	-	-
	1	0,5	25,0

Tabelle 8.9: Verwendete Materialparameter für das eindimensionale verallgemeinerte MAXWELL-Modell mit variierendem initialen E-Modul E_0

berücksichtigen lassen. Der Vorteil liegt darin, dass das Verhalten mehrerer verallgemeinerter MAXWELL-Modelle durch *ein einziges* KNN-Materialmodell abgebildet werden kann. Für rein eindimensionale Anwendungen mit analytisch erzeugten Trainingsdaten ist dieser Ansatz allerdings nur von begrenztem praktischem Nutzen, da sich keine nennenswerten Einsparungen bei den Rechenzeiten ergeben, bzw. die Auswertung eines einfachen analytischen Modells am GAUSS-Punkt in der Regel schneller ist, als die KNN-Auswertung. Auf Basis von synthetischen Daten, die mittels eines parametrisierbaren RVEs gewonnen werden, stellt diese Vorgehensweise jedoch einen immensen Vorteil dar. Die folgenden Untersuchungen dienen daher vor allem als Machbarkeitsstudie und zur Erläuterung der Datengenerierung unter Einbezug von Materialparametern, die in Kapitel 8.3.5 für komplexere Beispiele benötigt wird.

Im Folgenden wird ein verallgemeinertes MAXWELL-Modell mit einem MAXWELL-Element als Referenzlösung zur Erzeugung synthetischer Daten verwendet.

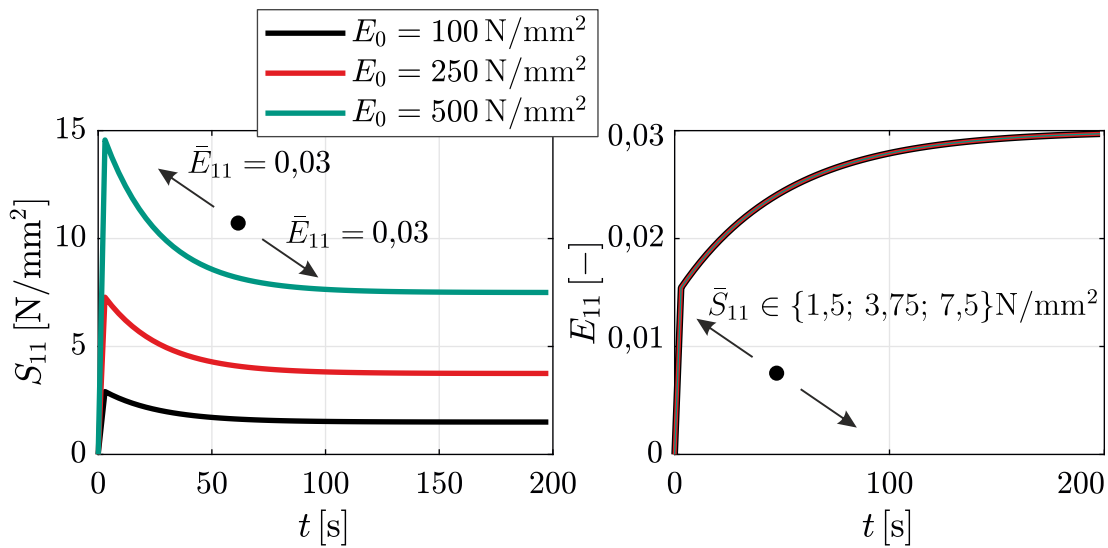


Abbildung 8.6: Exemplarische Relaxations- und Kriechversuche für das in Tabelle 8.9 gegebene analytische Modell unter Verwendung dreier initialer E-Moduln E_0

Versuchstyp	n_{Pfade}	Parameterintervalle
Relaxationsversuch	30	$\Delta \bar{t} \in [2,5, 10] \text{ s}$, $\bar{T} \in [200, 200] \text{ s}$, $\bar{E}_{11} \in [-0,03, 0, 03]$
Kriechversuch	30	$\Delta \bar{t} \in [2,5, 10] \text{ s}$, $\bar{T} \in [200, 200] \text{ s}$, $\bar{S}_{11} \in [-5, 5] \text{ N/mm}^2$
Zyklischer Versuch	30	$\Delta \bar{t} \in [2,5, 10] \text{ s}$, $\bar{T} \in [200, 200] \text{ s}$, $\bar{E}_{11} \in [-0,03, 0, 03]$, $2\bar{\omega}/\pi \in [1, 2]$
Für alle Pfade		$p \in [100, 500] \text{ N/mm}^2$

Tabelle 8.10: Pfadparameter der Datengenerierung für das $\text{KNN}_{\text{Versuche}}$ unter zusätzlicher Berücksichtigung variierender Materialparameter

Pfad-Nr.	Versuch	E_{11}	Δt	T	E_0
Nr.		$[-]$	$[\text{s}]$	$[\text{s}]$	$[\text{N/mm}^2]$
1	Relaxation	0,01	4	200	150
2	Relaxation	0,01	4	200	450
3	Zugversuch	0,01	8	400	150
4	Zugversuch	0,01	8	400	450

Tabelle 8.11: Verwendete Validierungspfade mit Angabe der E_{11} -Werte und des Materialparameters E_0

Die zugrunde liegenden Materialparameter sind in Tabelle 8.9 angegeben. Der initiale Elastizitätsmodul E_0 variiert im Intervall $[100, 500] \text{ N/mm}^2$, während der Modulanteil μ_1 und die Relaxationszeit τ_1 für alle drei Varianten konstant bleiben. Das Relaxationsverhalten für eine vorgegebene Verzerrung $\bar{E}_{11} = 0,03$ sowie Kriechversuche unter den Spannungen $\bar{S}_{11} \in \{1,5; 3,75; 7,5\} \text{ N/mm}^2$ sind in Abbildung 8.6 für drei unterschiedliche Anfangsmodule $E_0 \in \{100; 250; 500\} \text{ N/mm}^2$ dargestellt. Die Kriechverläufe stimmen überein, da die vorgegebenen Spannungen so gewählt wurden, dass die Gleichgewichtsverzerrung für alle Materialparameter den gleichen Wert von 0,03 nicht überschreitet. Für sämtliche Versuche zeigt sich ein Konvergenzverhalten bei etwa $T \approx 200 \text{ s}$, womit diese Versuchsdauer für die Datengenerierung des KNN übernommen wird.

Zur zeitlichen Diskretisierung werden 20–80 Stützstellen angestrebt, was einer Zeitschrittweite $\Delta \bar{t} \in [2,5, 10] \text{ s}$ entspricht. Die Spannungsgrenzen der Kriechversuche werden zu $S_{11} \in [-5, 5] \text{ N/mm}^2$ festgelegt, sodass bei $E_0 \rightarrow 500 \text{ N/mm}^2$ die resultierenden Verzerrungen unterhalb von $E_{11} = 0,03$ bleiben, während bei

$E_0 \rightarrow 100 \text{ N/mm}^2$ Werte oberhalb dieser Grenze erreicht werden können. Alternative Strategien zur Datengenerierung könnten eingesetzt werden, um die Einhaltung dieser Grenze strenger zu gewährleisten. Die gewählten Pfadparameter sind in Tabelle 8.10 zusammengefasst. Alle Trainingsdaten stammen aus Versuchen mit konstanter Zeitschrittweite, eine adaptive Diskretisierung wird in diesem Beispiel nicht berücksichtigt, da durch die deutlich höhere benötigte Datenmenge der Trainingsraum bezüglich der Zeitschrittweiten ohnehin gut abgedeckt wird.

Für das Training des KNN werden pro Versuch 30 zufällige Pfade erzeugt. Diese Datenmenge führt zu einem robusten Trainingsverhalten, das unabhängig vom jeweiligen Trainingsdurchlauf ist. Insgesamt ergeben sich so für das hier eingesetzte Netz 3399 Datenpunkte. Durch die Aufnahme des Materialparameters E_0 in den Eingangsvektor steigt zum einen der benötigte Datenumfang im Vergleich zu den vorherigen Beispielen, zum anderen erweitert sich die Eingangsdimension des KNN. Die verwendete Topologie lautet daher $5-30-1$. Zur Validierung dienen die in Tabelle 8.11 aufgeführten Pfade, wobei zwei Relaxations- sowie zwei Zugversuche mit unterschiedlichen Zeitschrittweiten und verschiedenen initialen Elastizitätsmodulen definiert werden.

Die Spannungsantworten des KNN für diese vier Versuche sind in Abbildung 8.7 dargestellt. Es zeigt sich eine sehr gute Übereinstimmung zwischen der analytischen Referenzlösung und der KNN-Approximation. Mit ausreichend vielen Trainingsdaten lässt sich somit der zusätzliche Parameter E_0 problemlos in ein viskoelastisches KNN-Materialmodell integrieren. Denkbar ist dieser Ansatz sowohl für die Konstruktion von KNN-Materialmodellen auf Basis experimenteller Versuchsdaten verschiedener viskoelastischer Materialien als auch – wie in Kapitel 8.3.5 gezeigt – für synthetisch erzeugte Datensätze, bei denen RVE-Parameter gezielt variiert werden.

8.1.5 Numerisches Beispiel: Weicher Aktuator mit viskoelastischen Fasern

Die bisher vorgestellten Untersuchungen wurden ausschließlich am Materialpunkt durchgeführt, da sich die Eignung eines Materialmodells zunächst auf dieser Ebene prüfen lässt. Für die praktische Anwendung ist jedoch entscheidend, dass das Modell auch für die Verwendung in Rahmen der Finite-Elemente-Methode brauchbar ist. Rein eindimensionale numerische Beispiele besitzen dabei nur begrenzte Aussagekraft, da sie im Wesentlichen dieselben Zusammenhänge wie auf Materialpunktebene abbilden. Auf einfache Stabbeispiele wird daher verzichtet und das Verhalten des KNN-Materialmodells direkt an einem komplexeren numerischen Beispiel demonstriert.

Im Folgenden wird das in Abschnitt 8.1.4 beschriebene KNN-Materialmodell als

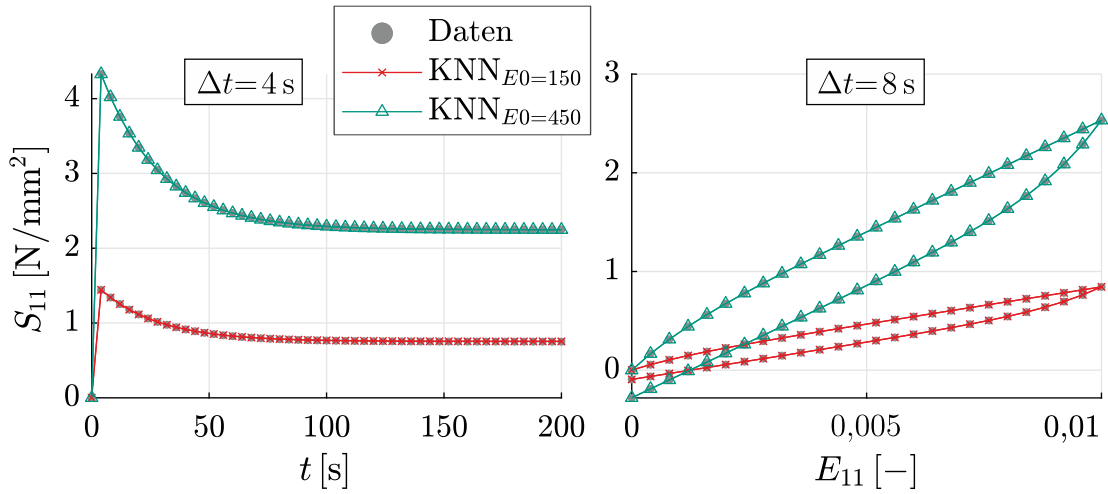


Abbildung 8.7: Darstellung der Approximation der in Tabelle 8.11 gegebenen Pfade mit einem KNN unter Berücksichtigung zusätzlicher Materialparameter als Eingangsvariable

eindimensionales Fasermmodell eingesetzt. Betrachtet wird ein sogenannter weicher Einkammer-Aktuator, dargestellt in Abbildung 8.8. Derartige Aktuatoren kommen in der Softrobotik zum Einsatz und bestehen überwiegend aus hyperelastischen (gummiartigen) Werkstoffen, die im aktivierten Zustand großen Deformationen ausgesetzt sein können. Ausführliche Beschreibungen und Analysen der Modellierungsmöglichkeiten verschiedener Ein- und Mehrkammeraktuatoren finden sich beispielsweise bei XAVIER ET AL. [153] oder PAGOLI ET AL. [104].

Durch Anlegen eines Innendrucks werden Verformungen des Aktuators gezielt gesteuert. Neben Längenänderungen können dabei auch – etwa infolge geometrischer Schwerpunktexzentrizitäten – Biegeverformungen hervorgerufen werden. Um parasitären Verformungen, dem sogenannten Aufblähen (engl.: *ballooning*), der die Kammer begrenzenden Membran entgegenzuwirken, werden steifere Fasern in unterschiedlichen Anordnungen eingebettet, um das Deformationsverhalten gezielt zu steuern. Als Faserwerkstoffe kommen unter anderem Baumwolle oder Glasfasern (PEEL [105]) sowie Holzzellstoffe oder Nylon (GOETTLER und SHEN [41]) in Betracht. Durch die Steifigkeitsdifferenzen der verwendeten Fasern können die unerwünschten Verformungen dabei unterschiedlich effektiv begrenzt werden. Darüber hinaus werden im vorliegenden Beispiel auch viskoelastische Effekte berücksichtigt.

Der hier untersuchte Aktuator besitzt eine Länge von $L_e = 100$ mm sowie eine Höhe und Breite von $h_e = b_e = 20$ mm. Die Kammerwände sind $t = 1$ mm dick. Am Kopfende bei $x = 200$ mm wird der Hohlraum des Aktuators durch eine weitere Membran mit Dicke $t = 1$ mm geschlossen. Die Lagerung wird als

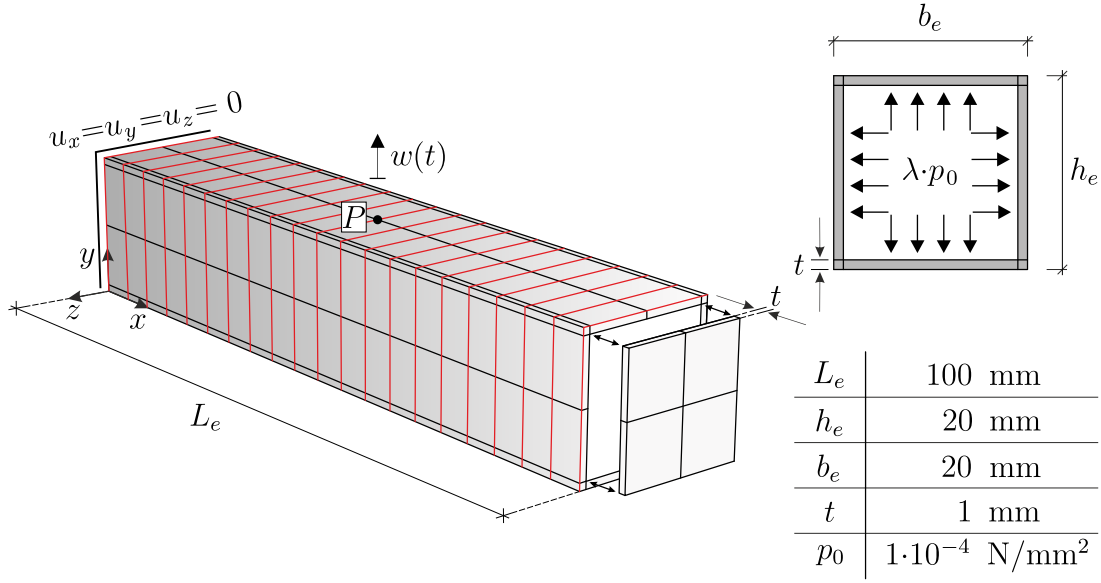


Abbildung 8.8: Geometrie, Belastung und Auflagerbedingungen des weichen Einkammer-Aktuators mit in rot dargestellten Faserumwicklung

vollständige Einspannung modelliert, sodass am Rand $x = 0$ die Verschiebungen $u_x = u_y = u_z = 0$, unter Verwendung von 3D-Volumenelementen, vorgegeben sind. Im Inneren wirkt ein Anfangsdruck von $p_0 = 1 \cdot 10^{-4} \text{ N/mm}^2$, der über einen Lastfaktor λ schrittweise erhöht wird. Da Oberflächendrücke stets senkrecht auf die Fläche wirken, wird die Belastung als Folgelast modelliert, womit durch die Verformungsabhängigkeit auch Lastanteile in die tangentielle Steifigkeitsmatrix eingehen. Von Interesse ist die vertikale Verformung $w(t)$ am Punkt $P(50, 20, -10) \text{ cm}$.

Für die Diskretisierung kommen geometrisch nichtlineare 27-Knoten Volumenelemente zum Einsatz, da sie im Vergleich zu 8-Knoten Elementen bessere Biegeeigenschaften aufweisen. Konvergenzuntersuchungen ergaben, dass die in Abbildung 8.8 dargestellte Netztopologie mit 20 Elementen in Längsrichtung und je zwei Elementen über Höhe und Breite ausreichend ist. Die ECKEelemente sind dabei nicht zwingend für die Erzeugung des FE-Netzes erforderlich.

Die Membran wird als hyperelastisches Material mit einer Verzerrungsenergie-dichtefunktion nach OGDEN beschrieben:

$$\mathcal{W}(\mathbf{C}) = \sum_{r=1}^m \frac{\mu_r}{\alpha_r} (\lambda_1^{\alpha_r} + \lambda_2^{\alpha_r} + \lambda_3^{\alpha_r}) + \mathcal{W}_{\text{vol}}, \quad (8.1)$$

wobei $\lambda_1, \lambda_2, \lambda_3$ die Hauptstreckungen sind, die als Eigenwerte von $\mathbf{F}$ bzw. als Wurzel der Eigenwerte von $\mathbf{C}$ identifiziert werden. Der kompressible Anteil wird

r	μ_r [N/mm ²]	α_r [–]	Λ [N/mm ²]
1	$6,3 \cdot 10^{-1}$	1,3	1000
2	$1,2 \cdot 10^{-3}$	5,0	
3	$-1,0 \cdot 10^{-2}$	-2,0	

Tabelle 8.12: Parameter (μ_r, α_r) und Lamé-Parameter Λ für die Treloar-Daten nach OGDEN

über

$$\mathcal{W}_{\text{vol}} = \frac{\Lambda}{4} (J^2 - 1 - \ln J^2) - \sum_{r=1}^m \mu_r \ln J \quad (8.2)$$

berücksichtigt. Weitere Details finden sich bei KLINKEL [70]. Die Materialparameter stammen aus OGDEN [103] auf Basis der Versuche von TRELOAR [133] und sind in Tabelle 8.12 zusammengefasst⁵⁶. Mit dem gewählten Lamé-Parameter $\Lambda = 1000 \text{ N/mm}^2$, entsprechend einer Poissonzahl von $\nu \approx 0,4998$, wird nahezu inkompressibles Materialverhalten abgebildet. Die Fasern werden als nichtlineare Fachwerkelemente modelliert und zeigen ein viskoelastisches Materialverhalten gemäß den Parametern aus Tabelle 8.9. Während die Membran großen Dehnungen in Längsrichtung unterliegt und deshalb ein hyperelastisches Stoffgesetz erfordert, genügt für die Fasern ein lineares viskoelastisches Gesetz, da die in den Fachwerkelementen auftretenden Verzerrungen im Rahmen des Beispiels klein bleiben. Die initialen Elastizitätsmoduln $E_0 \in [100, 500] \text{ N/mm}^2$ repräsentieren dabei weiche Elastomere (untere Grenze) bis hin zu steifen Holzzellstoffen (obere Grenze). Baumwollfasern ordnen sich dabei zwischen diesen Werten ein. Der Modulanteil μ_1 sowie die Relaxationszeit τ_1 werden angenommen, da hier keine Referenzwerte auf Basis echter Materialparameter vorliegen. Das Materialverhalten der Fasern wird sowohl mit dem analytischen Referenz-Modell als auch mit dem KNN-Materialmodell aus dem vorherigen Abschnitt beschrieben, wobei letzteres den initialen E-Modul als zusätzliche Eingangsvariable berücksichtigt.

Der Aktuator wird zunächst in 70 Lastschritten mit einer Zeitschrittweite von $\Delta t = 3 \text{ s}$ bis zum Zeitpunkt $T = 210 \text{ s}$ belastet, sodass der aufgebrachte Innendruck $p = \lambda \cdot p_0 = 0,021 \text{ N/mm}^2$ beträgt. Anschließend bleibt der Druck konstant, und die vertikale Verformung $w(t)$ an der Oberseite des Aktuators (vgl. Abbildung 8.8) wird bis $T = 362 \text{ s}$ beobachtet, wobei 19 Zeitschritte mit $\Delta t = 8 \text{ s}$ verwendet werden. Aufgrund der viskoelastischen Eigenschaften der Fasern sind dabei Kriechverformungen bei konstantem Innendruck zu verzeichnen.

⁵⁶Die Schubparameter μ_r des OGDEN-Modells sind nicht mit den Modulanteilen μ_i des verallgemeinerten MAXWELL-Modells zu verwechseln. Auf eine weitere Differenzierung der Notation wird hier verzichtet.

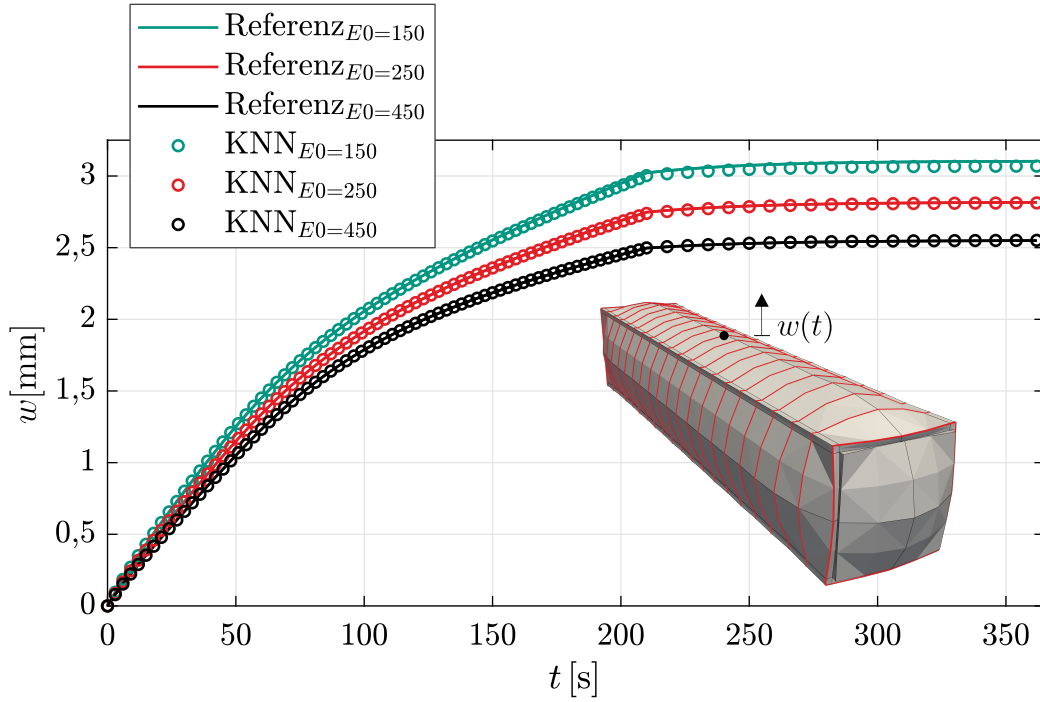


Abbildung 8.9: Vergleich der Belastungs- und Kriechkurven am Punkt P auf der Aktuatoroberfläche für drei Fasern unterschiedlicher Steifigkeit, mit der Lösung des analytischen Referenz- und KNN-Materialmodells

	$E_0 = 150 \text{ N/mm}^2$		$E_0 = 250 \text{ N/mm}^2$		$E_0 = 450 \text{ N/mm}^2$	
Zeit [s]	w_{Ref} [cm]	w_{KNN} [cm]	w_{Ref} [cm]	w_{KNN} [cm]	w_{Ref} [cm]	w_{KNN} [cm]
210	3,02	3,00	2,75	2,74	2,50	2,50
362	3,10	3,07	2,82	2,81	2,55	2,55

Tabelle 8.13: Vertikale Verschiebung in der Mitte des Aktuators zu zwei Zeitpunkten für verschiedene Fasersteifigkeiten und Materialmodelle

Die Ergebnisse der beiden Materialmodelle für drei unterschiedliche initiale Elastizitätsmoduln $E_0 \in \{150; 250; 450\} \text{ N/mm}^2$ sind in Abbildung 8.9 dargestellt. Es zeigt sich zunächst, dass sich abhängig von der Fasersteifigkeit bei beiden Modellen deutliche Unterschiede in den resultierenden Verformungen ergeben. Zugleich ist eine sehr gute Übereinstimmung der KNN-Approximation mit den Ergebnissen der FE-Referenzlösung mit einem analytischen Materialmodell zu erkennen. Die Resultate zu den Zeitpunkten $T = 210 \text{ s}$ und $T = 362 \text{ s}$ sind zusätzlich in Tabelle 8.13 zusammengefasst. Die Abweichungen zwischen den beiden Modellen bleiben dabei marginal. Das Beispiel belegt, dass ein gut trainiertes

KNN-Materialmodell auch für den Einsatz in anspruchsvollen numerischen FE-Simulationen zuverlässig geeignet ist.

8.2 Dreidimensionale Untersuchungen

Nach den zuvor durchgeführten Untersuchungen und den Erläuterungen der vorgenommenen Schritte zur Konstruktion eindimensionaler KNN-basierter Materialmodelle, wird in diesem Kapitel der Fokus auf die wesentlichen Aspekte der dreidimensionalen Konstruktion gelegt. Mit $n_{\text{str}} = 6$ steigt die Eingangsdimension des KNN auf $n_i = 19$ und die Ausgangsdimension auf $n_o = 6$, sofern keine weiteren Parameter berücksichtigt werden, s. Gleichung (7.1). Dies hat die Konsequenz, dass zum einen der Bedarf an Trainingsdaten steigt und zum anderen eine Strategie entwickelt werden muss, welche Verzerrungskomponenten bei der Datenerzeugung variiert werden. Dahingehend wird zunächst ein Methode verfolgt, die rein auf Basis uniaxialer Versuche basiert, in Anlehnung an die Ausführungen in Kapitel 7.2.6. Weiterhin wird der Einfluss der Tangenten-Nebenbedingungen auf das KNN-Training und das Approximationsverhalten untersucht, sowie die Fähigkeit, die rein elastische volumetrische von der viskoelastischen deviatorischen Materialantwort zu unterscheiden. Das Kapitel schließt mit einem numerischen Beispiel. Die Untersuchungen werden an einem einfachen verallgemeinerten MAXWELL-Modell durchgeführt, die Anwendung der Methoden auf unterschiedliche Materialmodelle erfolgt im Rahmen der Schalenbeispiele.

8.2.1 Konstruktion des Materialmodells auf Basis uniaxialer Versuche

Wie in Kapitel 3.4 erwähnt, kommen zur experimentellen Prüfung von Werkstoffen häufig uniaxiale Materialversuche zum Einsatz. Im Weiteren wird daher geprüft, ob sich dieses Vorgehen auch als Strategie zur Datengenerierung für das KNN-Training eignet. Dabei wird auf das Vorgehen der synthetischen Materialpunktversuche aus Kapitel 7.2.5 zurückgegriffen. Da im Folgenden nur isotrope Werkstoffe betrachtet werden, besitzt die Materialmatrix immer die allgemeine Form

$$\mathbf{C}_T = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{21} & C_{22} & C_{23} & 0 & 0 & 0 \\ C_{31} & C_{32} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix}. \quad (8.3)$$

Initialer E-Modul E_0 [N/mm ²]	Poisson- Zahl ν_0 [-]	MAXWELL Element i	Modulanteil $\mu_i = E_i/E_0$ [-]	Relaxationszeit τ_i [s]
2600	0,2	-	-	-
		1	0,50	33,33

Tabelle 8.14: Verwendete Materialparameter für das dreidimensionale verallgemeinerte MAXWELL-Modell

Die Schuberverzerrungen sind somit vollständig entkoppelt von den verbleibenden Dehnungskomponenten.

Für die mehrdimensionalen Untersuchungen wird das verallgemeinerte MAXWELL-Modell aus Kapitel 8.1.1 aufgegriffen und zusätzlich eine initiale Poisson-Zahl von $\nu_0 = 0,2$ angenommen, s. Tabelle 8.14. Dabei wird davon ausgegangen, dass sich nur die deviatorischen Verzerrungsanteile viskoelastisch verhalten. Damit können die Grenzen der eindimensionalen Pfad-Parameter aus Tabelle 8.2 weitestgehend übernommen werden. Es erfolgt lediglich eine Anpassung der Spannungskomponenten für die mehrdimensionalen Kriechversuche, um die entsprechende Annahme kleiner Verzerrungen $\|\mathbf{E}\| \leq 0,03$ einzuhalten. In Analogie zum eindimensionalen Vorgehen (vgl. Kapitel 8.1.1) ergibt sich daraus:

$$\bar{S}_{ii} \in [-40, 40] \text{ N/mm}^2 \quad \text{und} \quad \bar{S}_{ij} \in [-20, 20] \text{ N/mm}^2, \quad i \neq j. \quad (8.4)$$

Das KNN-Training erfolgt rein auf Basis von Daten, die durch uniaxiale Versuche erzeugt werden. Die Anzahl der Pfadscharen beträgt $n_{\text{Pfade}} \in \{5; 10; 20\}$ pro Verzerrungskomponente und für alle drei Versuche (Relaxation, Kriechen, Zyklisch). Die auf den entsprechenden Pfadscharen trainierten KNNs werden mit $\text{KNN}_{n_{\text{Pfade}}}$ bezeichnet. Es werden erneut fünf Trainingsdurchläufe durchgeführt um der Zufälligkeit der gewählten Pfadparameter Rechnung zu tragen. Dies führt für die drei KNN auf die folgenden Datenmengen:

- $n_{\text{Pfade}} = 5$: 2962–3175 Datenpunkte
- $n_{\text{Pfade}} = 10$: 5906–6098 Datenpunkte
- $n_{\text{Pfade}} = 20$: 11786–12304 Datenpunkte

Die zur Validierung der KNN verwendeten Pfade sind in Tabelle 8.15 dargestellt. Hierbei kommen zum einen die drei aus dem Training bekannten Versuchsarten zum Einsatz, ergänzt um einen Zugversuch mit Be- und Entlastung. Darüber hinaus handelt es sich um biaxiale Versuche, die so nicht im Trainingsdatensatz

Pfad Nr.	Versuchsart	Komponenten	Werte	Δt [s]	T [s]
1	Zyklisch	E_{11}, E_{22}	-0,01, 0,01 [–]	8	300
2	Relaxation	E_{22}, E_{33}	0,01, 0,01 [–]	8	300
3	Kriechen	S_{33}, S_{13}	10, 5 [N/mm ²]	12	300
4	Zugversuch	E_{13}, E_{23}	0,01, 0,01 [–]	12	300

Tabelle 8.15: Verwendete Validierungspfade unter Angabe der gewählten Parameter

enthalten sind. Beispielsweise wird durch die Kombination $\bar{E}_{11}$ - E_{22} - E_{33} infolge der Kopplung der Komponenten bereichsweise das Verhalten für Zug–Druck–Druck infolge der uniaxialen Versuche abgebildet, nicht jedoch Zug–Zug–Druck, wie es in Pfad-Nr. 1 vorkommt. Die Diskretisierung erfolgt weiterhin mit unterschiedlichen Zeitschrittweiten für Pfad-Nr. 1–2 sowie 3–4.

Im Rahmen von Voruntersuchungen wurden KNN-Topologien identifiziert, die die geringsten Testfehler in Bezug auf die definierten Datenmengen aufweisen. Für die drei KNN wurden damit die in Tabelle 8.16 angegebenen Topologien gewählt.

KNN ₅	KNN ₁₀	KNN ₂₀
19 – 40 – 6	19 – 60 – 6	19 – 80 – 6

Tabelle 8.16: Verwendete KNN-Architekturen

Die besten Approximationen, auf Basis des kleinsten Test-Fehler über jeweils fünf Trainingsdurchläufe, der Spannungskomponenten S_{11} , S_{33} und S_{12} über die Zeit sowie von S_{13} über die Verzerrung E_{13} sind für die drei KNN in Abbildung 8.10 dargestellt. Es zeigt sich, dass dem KNN₅ nicht genügend Daten zur Verfügung stehen, um biaxiale Spannungs–Verzerrungszustände zufriedenstellend zu approximieren. Mit zunehmender Datenmenge verbessert sich zwar die Approximation, jedoch ist auch bei KNN₂₀ kein robustes Training mit verlässlicher Generalisierungsfähigkeit für biaxiale Belastungszustände (vgl. etwa die Spannungsantwort bei Pfad-Nr. 2) zu erkennen.

Dabei sei angemerkt, dass in den fünf Durchläufen trotz ausreichender Datenmenge teilweise unbrauchbare KNN-Modelle trainiert wurden. Es lässt sich damit schließen, dass sich auf Basis uniaxialer Versuche prinzipiell ein mehrdimensionales KNN-Materialmodell konstruieren lässt. Das Training liefert jedoch nicht zuverlässig robuste Ergebnisse, weshalb im Folgenden die zusätzliche Einbindung der KNN-Materialtangente in das Training untersucht wird. Damit wird das Ziel

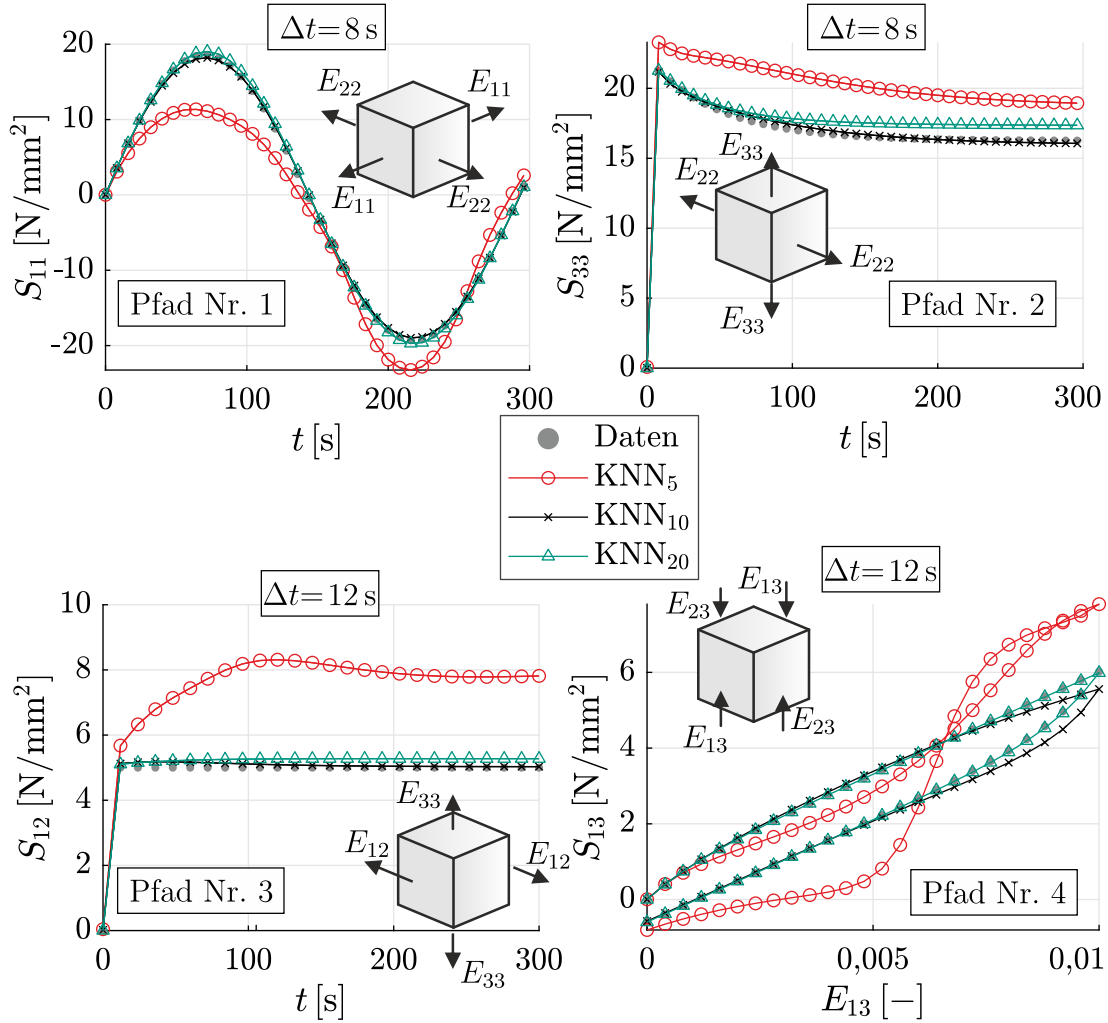


Abbildung 8.10: Darstellung der Approximation der in Tabelle 8.15 gegebenen Pfade mit drei KNN, trainiert auf verschiedenen Datenmengen

der Arbeit – ein gut gestelltes Optimierungsproblem zu formulieren, welches unabhängig von den Initialwerten (Daten, Gewicht) immer zu brauchbaren Minima konvergiert – verfolgt.

8.2.2 Einfluss der Tangenten-Nebenbedingungen

Um eine Verbesserung der Approximationsgüte auf Basis der uniaxialen Daten zu erreichen, wird im Weiteren die Tangenten-Nebenbedingung aus Kapitel 7.3.1 im Rahmen des KNN-Training berücksichtigt. Dafür wird lediglich das mit fünf trainierten Pfaden KNN₅ – vgl. Kapitel 8.2.1 – aus den vorherigen Untersuchungen herangezogen und zusätzlich die Anzahl der Datenpunkte P^S , an welchen die Nebenbedingung eingefordert wird, erhöht. Diese nehmen im folgenden die Werte

$\text{KNN}_5(P^S = 500)$	$\text{KNN}_5(P^S = 1000)$	$\text{KNN}_5(P^S = 2000)$
19 – 60 – 6	19 – 80 – 6	19 – 80 – 6

Tabelle 8.17: Verwendete KNN-Topologien für das SOBOLEV-Training

$P^S \in \{500; 1000; 2000\}$ an. Entsprechend wurden für die drei verwendeten KNN die in Tabelle 8.17 gegebenen Topologien, die im Rahmen von Voruntersuchungen als brauchbar identifiziert wurden, verwendet.

Es zeigt sich, dass eine einzelne verdeckte Schicht mit vielen Neuronen besser geeignet ist als tiefere Netze mit mehreren Schichten. Weiterhin hat sich eine Erhöhung der Anzahl verdeckter Neurone bei einer Erhöhung der Datenpunkte von 1000 auf 2000 als nicht notwendig herausgestellt. Es wurden erneut pro KNN fünf Trainingsdurchläufe durchgeführt und die gleichen Testpfade wie in den vorherigen Untersuchungen verwendet (vgl. Tabelle 8.15). Der Penalty-Parameter wird dabei stets zu $\rho = 1$ gewählt, in Anlehnung an die Ausführungen in Kapitel 6.3.2. Für alle durchgeführten KNN-Trainings konnten Konvergenzen sowohl im Trainings- als auch im Testfehler innerhalb 100 Epochen festgestellt werden.

Die Spannungsantworten der KNN mit den geringsten Test-Fehlern sind in Abbildung 8.11 dargestellt. Angemerkt sei dabei, dass unabhängig vom Trainingsdurchlauf ähnliche Approximationen erzielt werden. Hier zeigt sich bereits für $P^S = 500$ eine zufriedenstellende und vergleichsweise bessere Approximation als die rein auf Primärdaten trainierten KNN aus dem vorherigen Kapitel. Mit $P^S = 1000$ lassen sich kaum Abweichungen von der Referenzlösung des analytischen Materialmodells erkennen. Eine weitere Erhöhung auf $P^S = 2000$ liefert daher keinen signifikanten Mehrwert und beeinflusst das Training auch nicht negativ.

Daraus lässt sich schlussfolgern, dass eine Strategie zur Datengenerierung auf Basis uniaxialer Materialversuche und anschließendem SOBOLEV-Training zu einem robusten KNN-Training und guten Generalisierungseigenschaften für beliebige mehrdimensionale Spannungs-Verzerrungszustände führt. Durch die Information über die Materialtangente werden im Gegensatz zu rein datenbasiertem Training wichtige Informationen über die Kopplungen der entsprechenden Komponenten mitgegeben. Wichtig ist dabei, dass diese Information hier aufgrund der synthetischen Daten immer zugänglich ist, was bei experimentellen Daten nicht der Fall ist. Von Vorteil ist weiterhin, dass keine signifikanten Überlegungen zur Wahl der Anzahl der Datenpunkte und des Penalty-Parameters vorgenommen werden müssen. Bei ersterem führt eine unnötige Erhöhung lediglich zu steigenden Rechenzeiten für das Training und letzterer kann aufgrund der a priori sichergestellten gleichen Wichtung der Daten- und Tangenten-Fehlerterme zu $\rho = 1$ gesetzt werden.

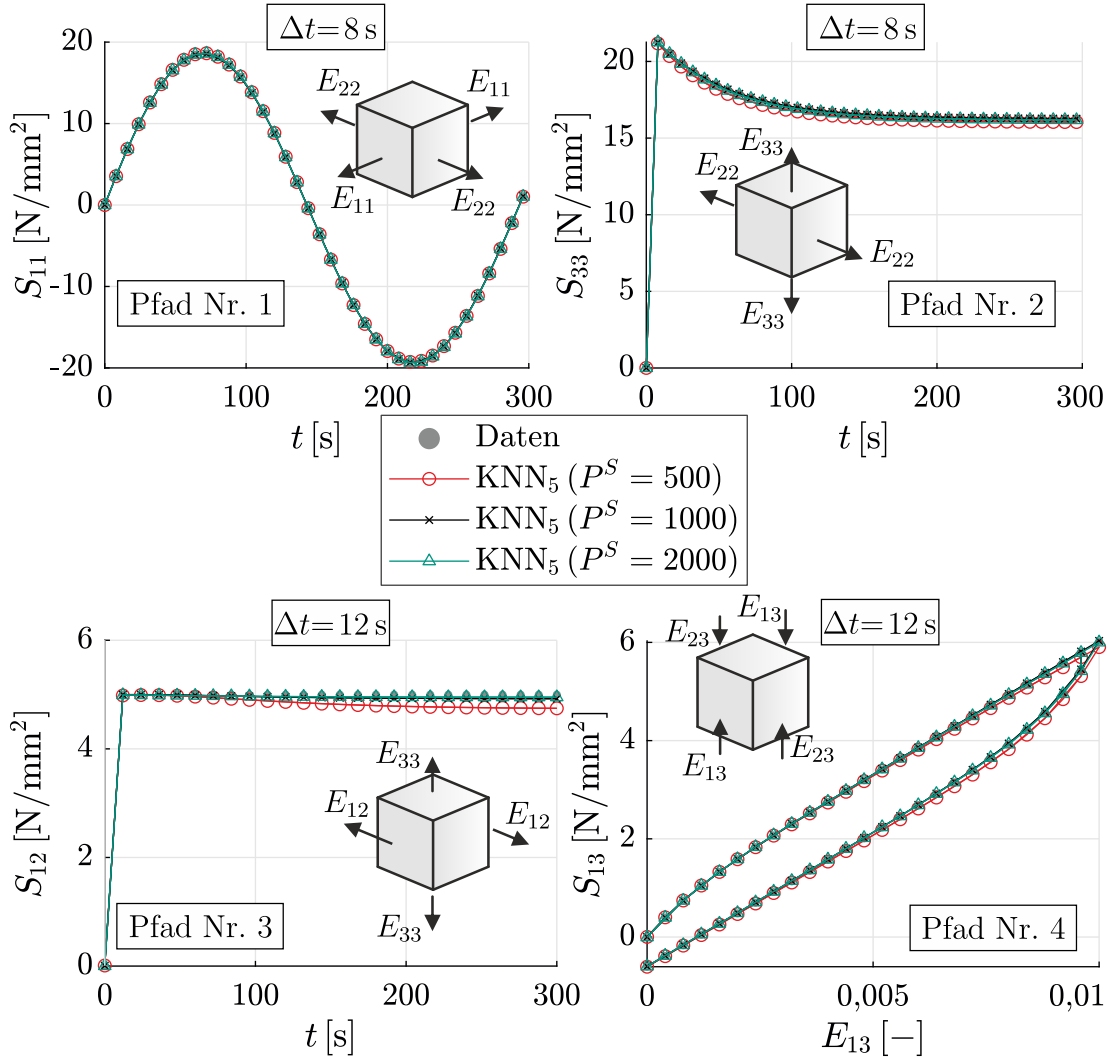


Abbildung 8.11: Darstellung der Approximation der in Tabelle 8.15 gegebenen Pfade mit drei KNN, trainiert auf verschiedenen Datenmengen

8.2.3 Hydrostatische und deviatorische Materialantwort

Zum Abschluss der mehrdimensionalen Untersuchungen am Materialpunkt wird geprüft, ob das KNN rein elastische hydrostatische von viskoelastischen deviatorischen Belastungszuständen unterscheiden kann. Dafür wird der in Tabelle 8.18 dargestellte Verzerrungspfad zur Validierung herangezogen. Es handelt sich um einen Zugversuch, bei dem eine maximale Volumendehnung von $\Theta = 0,03$ erreicht wird. Die Dehnungsraten sind ausreichend groß gewählt, dass bei nicht rein hydrostatischer Belastung viskoelastische Effekte auftreten. Der Versuch umfasst sowohl Be- als auch Entlastung, wobei im Gegensatz zu Pfad-Nr. 4 aus dem vorherigen Kapitel keine Spannungs-Dehnungs-Hysterese auftritt.

Für das Training wird das KNN_5 aus dem vorherigen Kapitel unter Berücksichtigung der Nebenbedingung mit $P^S = 1000$ Datenpunkten und einer Architektur $19 - 80 - 6$ verwendet. Die Ergebnisse für die Spannungskomponenten S_{11} über die Verzerrung E_{11} sowie S_{22} über die Zeit sind in Abbildung 8.12 dargestellt. Die Ergebnisse zeigen, dass das KNN in der Lage ist, die hydrostatische von der deviatorischen Materialantwort zu unterscheiden, obwohl entsprechende Belastungszustände nicht explizit in den Trainingsdaten enthalten waren. Erwähnt wird dabei, dass die hydrostatische Achse innerhalb des Trainingsraums liegt und somit zumindest implizit erlernt werden kann. Damit zeigt sich, dass das Modell auch für unbekannte Belastungsarten eine konsistente Materialantwort liefert. Angemerkt sei, dass kein KNN, welches ohne die Nebenbedingung trainiert wurde – unabhängig von der verwendeten Datenmenge – zu brauchbaren Approximationen gelangte. Der entscheidende Informationsgehalt darüber, wie sich die Spannungen bei volumetrischen und deviatorischen Verzerrungsänderungen entwickeln, liegt somit nicht allein in den uniaxialen Spannungs–Dehnungs-Daten, sondern im Wesentlichen in der Materialtangente, die dem KNN über die Nebenbedingung zugänglich gemacht wird.

Versuchsart	Komponenten	Werte	Δt [s]	T [s]
Hydrostatisch	E_{11}, E_{22}, E_{33}	0,01, 0,01, 0,01 [–]	8	600

Tabelle 8.18: Verwendete Validierungspfade unter Angabe der gewählten Parameter

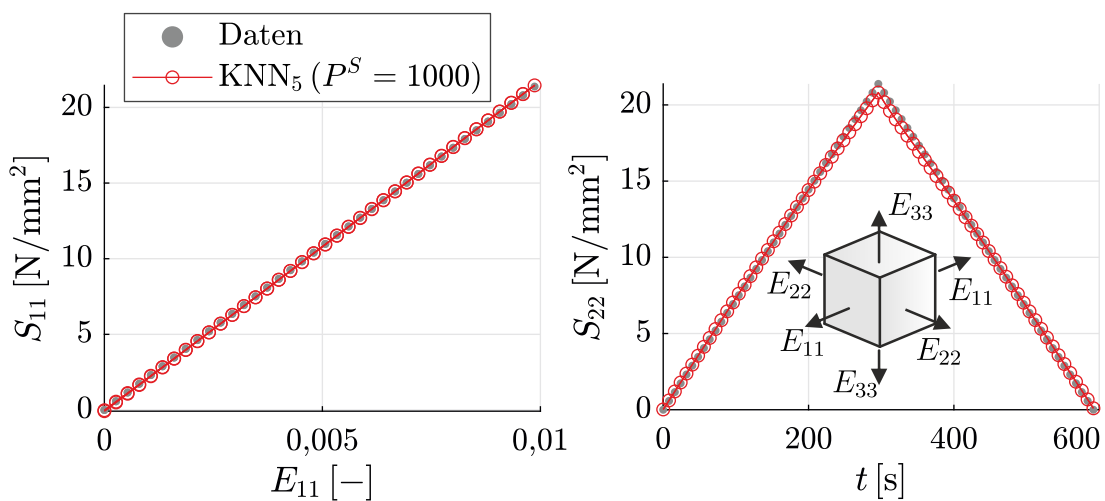


Abbildung 8.12: Darstellung der Approximation des in Tabelle 8.18 gegebenen Pfads mit dem KNN_5 und $P^S = 1000$ sowie die zugehörige Referenzlösung

8.2.4 Numerisches Beispiel: Ringförmige Platte mit Schlitz

Neben den Untersuchungen am Materialpunkt wird im Folgenden die Eignung des mehrdimensionalen KNN-Materialmodells im Rahmen der FEM analysiert. Als Referenzbeispiel dient eine dünne ringförmige Platte (vgl. Abbildung 8.13) mit Schlitz. Die Geometrie orientiert sich an den Arbeiten von BASAR ET AL. [11] sowie GRUTTMANN und WAGNER [44], in denen Schalenelemente für die Modellierung des Systems verwendet wurden. In dieser Arbeit wird das System dagegen mit Solid-Schalenelementen als dreidimensionales Modell abgebildet. Details zur verwendeten Elementformulierung sind in KLINKEL [69] und KLINKEL ET. AL. [70] zu finden.

Die Platte besitzt einen Innenradius von $R_i = 6$ mm, einen Außenradius von $R_a = 10$ mm sowie eine Dicke von $t = 0,04$ mm. Die Abmessungen sind als akademisch zu verstehen, eine Anpassung an reale Bauteilgrößen wäre durch geeignete Wahl der Materialparameter jederzeit möglich. Diskretisiert wird die Geometrie mit 30 Elementen in Umfangsrichtung, 6 Elementen in radialer Richtung sowie einem Element über die Dicke, womit eine auskonvergierte Lösung erreicht wird.

Die Belastung erfolgt durch eine Linienlast p am Schlitz der Platte, die mit einem Lastfaktor λ gesteigert und an den Knoten der Unterseite ($z = -t/2$) im Bereich $x \in [R_i, R_a]$ und $y = 0$ aufgebracht wird. Das Koordinatensystem ist im Mittelpunkt der Mittelfläche des Kreisrings definiert. Die gegenüberliegenden, unbelasteten Knoten am Schlitz werden vollständig unverschieblich gelagert ($u_x = u_y = u_z = 0$).

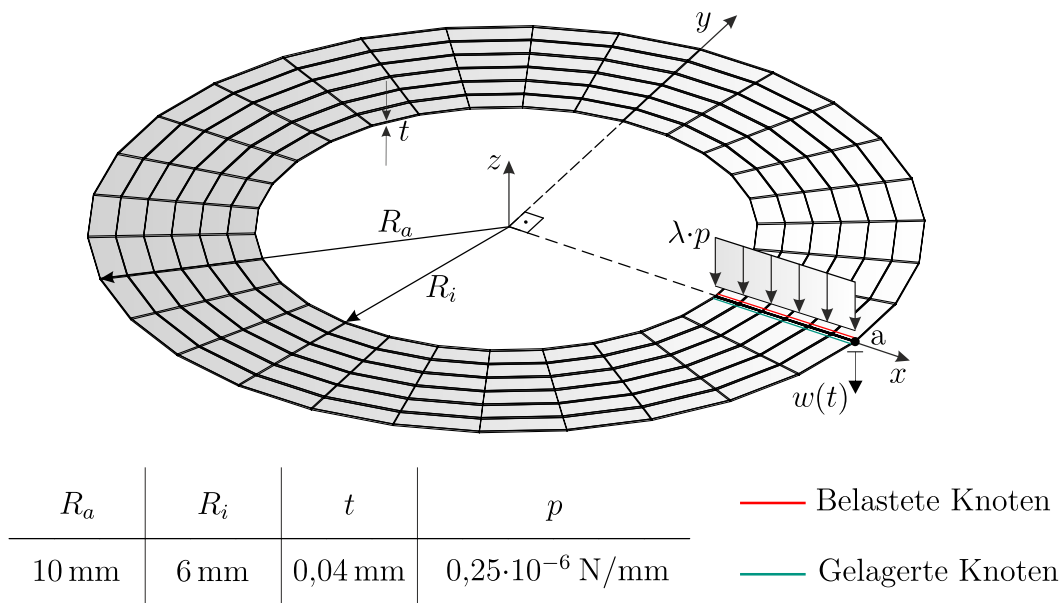


Abbildung 8.13: Geometrie und Belastung der ringförmigen Platte mit Schlitz

Die gegebene Belastung führt dazu, dass bei der Platte die Verzerrungen klein bleiben, jedoch große Verformungen auftreten, sodass eine geometrisch nichtlineare Theorie erforderlich ist. Das Materialverhalten wird als linear viskoelastisch auf Basis eines verallgemeinerten MAXWELL-Modells mit den in Tabelle 8.14 angegebenen Parametern beschrieben. Alternativ erfolgt die Modellierung mit dem KNN₅ aus dem vorherigen Kapitel, welches mit $P^S = 1000$ Nebenbedingungspunkten trainiert wurde. Für die Solid-Schale werden fünf EAS-Parameter (*engl.: Enhanced Assumed Strain*) berücksichtigt. Zudem wird der Ansatz nach BETSCH und STEIN [15] eingesetzt, um numerischen Versteifungseffekten in Dickenrichtung entgegenzuwirken, sowie der Ansatz nach BATHE und DVORKIN [12], um Schubversteifungen zu vermeiden.

Ein bislang nicht betrachteter Aspekt ist die zeitliche Invarianz des KNN-Materialmodells. Diese lässt sich als Forderung nach Energieerhaltung im Ruhezustand interpretieren, da in Abwesenheit äußerer Lasten keine zusätzliche Energie in das System eingetragen werden darf. Während dies beim analytischen Materialmodell per Konstruktion gewährleistet ist, gilt dies beim KNN-Materialmodell nicht notwendigerweise. Entsprechend wird überprüft, ob sich das Modell wie ein „Perpetuum Mobile“ verhält. Zu diesem Zweck wird das System zunächst im Zeitintervall $T \in [0, 80]$ s ohne Belastung betrachtet. Anschließend wird im Intervall $T \in [80, 160]$ s die Belastung linear auf $\lambda \cdot p = 2 \cdot 10^{-5}$ N/mm gesteigert und danach bis $T = 320$ s konstant gehalten. Alle Berechnungen erfolgen mit einer Zeitschrittweite von $\Delta t = 8$ s. Ausgewertet wird die vertikale Verschiebung $w(t)$

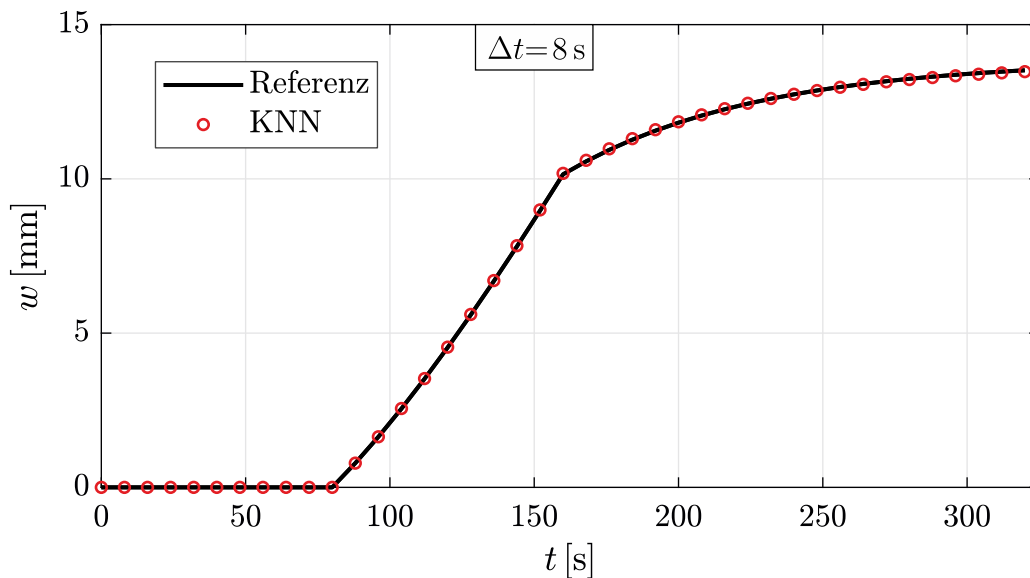


Abbildung 8.14: Ruhezustand sowie Belastungs- und Kriechverläufe der Durchbiegung $w(t)$ für das analytische und KNN-basierte Materialmodell

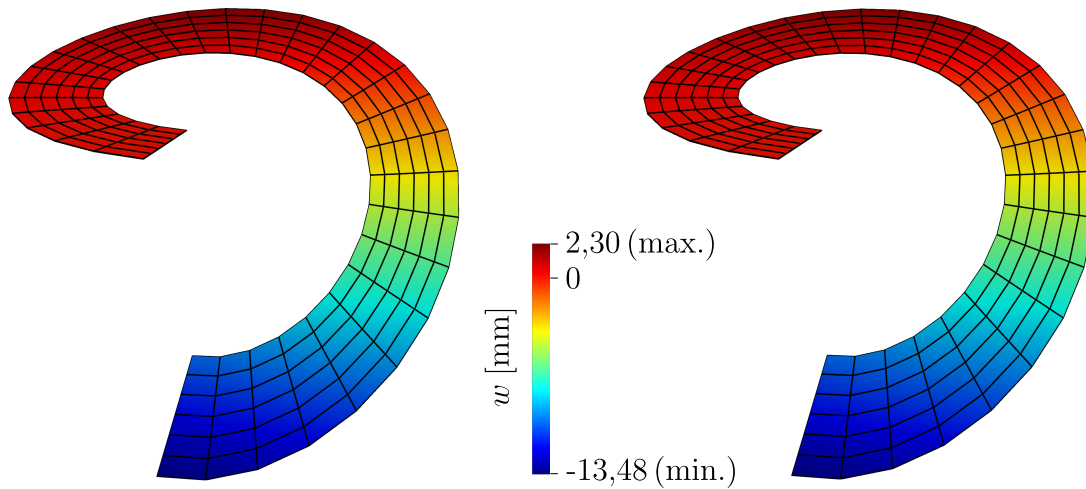


Abbildung 8.15: Verformungsfiguren zum Zeitpunkt $T = 320$ s für das analytische (links) und KNN-basierte Materialmodell (rechts)

an der Unterseite der Platte im Punkt a (vgl. Abbildung 8.13).

Die Ergebnisse beider Materialmodelle sind in Abbildung 8.14 dargestellt sowie zusätzlich deren Verformungsfiguren zum Zeitpunkt $T = 320$ s in Abbildung 8.15. Es zeigt sich eine sehr hohe Übereinstimmung für alle drei Belastungsphasen, so dass das KNN im Rahmen des Beispiels kein unphysikalisches Verhalten in Form von Energieeintrag bei Abwesenheit externer Lasten aufweist. Damit bestätigt sich, dass sich mit den in dieser Arbeit vorgeschlagenen Methoden auch für komplexe, mehrdimensionale Anwendungsbeispiele brauchbare KNN-Materialmodelle konstruieren lassen.

8.3 Modellierung von Schalenstrukturen

Im letzten Teil der Untersuchungen wird das Potenzial KNN-basierter viskoelastischer Materialmodelle für Schalenprobleme analysiert. Die Datenerzeugung erfolgt dabei ausschließlich auf Grundlage von Schalen-RVEs mittels numerischer Homogenisierung (vgl. Kapitel 5.3) mit dem Ausblick auf komplexe Mikrostrukturen. Geschichtete Strukturen, die beispielsweise im numerischen Beispiel des Unterkapitels 8.3.2 diskutiert werden, lassen sich grundsätzlich auch mit einer Schalenformulierung unter Berücksichtigung von 3D-Stoffgesetzen und Dickenintegration (vgl. KLINKEL ET AL. [71]) modellieren.

Die in den Kapiteln 8.1 und 8.2 vorgestellten Untersuchungen bilden die methodische Basis, sodass einige Ausführungen hier nur in verkürzter Form dargestellt werden. Das verwendete Schalenelement entspricht der Beschreibung in

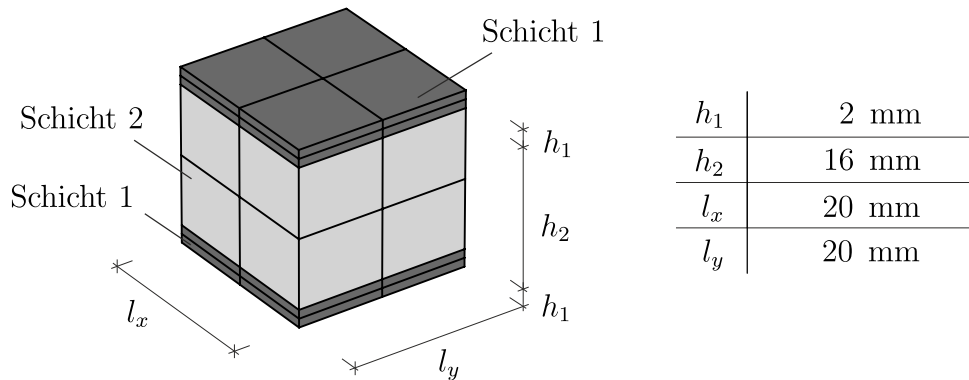


Abbildung 8.16: Geometrie und Abmessungen des geschichteten Schalen-RVEs

Kapitel 2.4.2 und liefert jeweils $n_{\text{str}} = 8$ arbeitskonforme Spannungs- und Verzerrungskomponenten. Ohne zusätzliche Parameter ergibt sich damit für das KNN eine Eingabedimension von $n_i = 25$ und eine Ausgabedimension von $n_o = 8$.

Der Fokus liegt im Folgenden auf der Eignung des KNN für ausgewählte numerische Beispiele, wobei relevante Materialpunktuntersuchungen dabei im Zusammenhang mit den Beispielen behandelt werden. Zunächst wird anhand einer geschichteten Schale die Approximationsfähigkeit des KNN für die hochdimensionale Problemstellung untersucht. Anschließend wird am Beispiel einer faserverstärkten Platte das Potenzial zur Reduktion der Rechenzeit im Vergleich zu einem Mehrskalenmodell demonstriert. Das Kapitel schließt mit der Erweiterung des KNN-Materialmodells um geometrische Parameter des Schalen-RVEs, womit die volle Leistungsfähigkeit des Ersatzmodells ausgeschöpft werden kann.

8.3.1 Schalen-RVE mit Schichtung: Voruntersuchungen am Materialpunkt

Geschichtete Strukturen mit viskoelastischen Eigenschaften finden sich beispielsweise bei Verbundglas, bei dem die Klebeschicht zwischen den beiden steifen Decklagen zeitabhängiges Materialverhalten zeigt (vgl. KUNTSCHKE [76]) oder beim gezielten Einsatz zur Verbesserung der Dämpfungseigenschaften von Bauteilen (vgl. KERWIN [65]). Im Folgenden wird ein Schalen-RVE eingeführt, das aus zwei dünnen rein elastischen Deckschichten und einer viskoelastischen Zwischenschicht besteht und in Abbildung 8.16 dargestellt ist. Die Höhe der Deckschichten (Schicht 1) beträgt $h_1 = 2$ mm, die der Zwischenschicht (Schicht 2) $h_2 = 16$ mm, sodass sich eine Gesamthöhe von $h = 20$ mm ergibt. Die Längenabmessungen sind hier durch die unendliche Anzahl an möglichen Einheitszellen beliebig und werden zu $l_x = l_y = 20$ mm gewählt. Diskretisiert wird mit $2 \times 2 \times 2$ Elementen pro Schicht unter Verwendung von 27-Knoten Volumenelementen. Für die syn-

Schicht	Initialer E-Modul E_0 [N/mm ²]	Poisson- Zahl ν_0 [-]	MAXWELL Element i	Modul- anteil $\mu_i = E_i/E_0$ [-]	Relaxations- zeit τ_i [s]
1	6850	0,2	-	-	-
2	2600	0,3	-	-	-
			1	0,90	33,33

Tabelle 8.19: Verwendete Materialparameter für das dreidimensionale verallgemeinerte MAXWELL-Modell

thetischen Materialversuche werden periodische Verschiebungsrandbedingungen gemäß Kapitel 5.2 zusammen mit fünf Lagrange-Parametern zur Einhaltung der kinematischen Zwangsbedingungen eingesetzt.

Die Materialparameter der beiden Schichten sind in Tabelle 8.19 gegeben. Der initiale E-Modul sowie die Relaxationszeit entsprechen dabei den Parametern aus dem in den dreidimensionalen Untersuchungen verwendeten Modell, mit dem Modulanteil $\mu_1 = 0,9$ weist das hier verwendete jedoch deutlich ausgeprägtere viskose Eigenschaften auf. Damit zeigen sich hier ähnliche Abklingeigenschaften wie im vorherigen Kapitel, womit sowohl bei Relaxation als auch bei Kriechen asymptotisches Verhalten in den Materialantworten nach $\bar{T} \approx 300$ s zu verzeichnen ist. Die Grenzen für die Zeitschrittweite werden damit ebenfalls zu $\Delta\bar{t} \in [5\text{ s}, 15\text{ s}]$ gewählt. Die Membran- und Schubverzerrungen werden auf das Intervall $\bar{\varepsilon}_{\alpha\beta}, \bar{\gamma}_{\alpha} \in [-0,03, 0,03]$ begrenzt. Für kleine Krümmungen gilt gemäß Gleichung (7.21) die Einschränkungen $|\bar{\kappa}_{\alpha\beta}| \leq 3 \cdot 10^{-3}/\text{mm}$, eine Begrenzung auf das Intervall $\bar{\kappa}_{\alpha\beta} \in [-3 \cdot 10^{-4}, 3 \cdot 10^{-4}] 1/\text{mm}$ erweist sich jedoch praktikabler, da dadurch unnötig große Biegemomente vermieden werden, was zu einem verbesserten KNN-Training führt. Um entsprechend bei spannungsgesteuerten Kriechversuchen die Grenzen der Schalenverzerrungen nicht zu überschreiten, werden die Streckenschnittgrößen wie folgt festgelegt:

$$\begin{aligned}
 |\bar{n}_{ii}| &\leq 100 \frac{\text{N}}{\text{mm}}, & |\bar{n}_{ij}| &\leq 50 \frac{\text{N}}{\text{mm}}, & |\bar{m}_{ii}| &\leq 2000 \frac{\text{Nmm}}{\text{mm}}, \\
 |\bar{m}_{ij}| &\leq 2000 \frac{\text{Nmm}}{\text{mm}}, & |\bar{q}_i| &\leq 50 \frac{\text{N}}{\text{mm}}, & i &\neq j.
 \end{aligned} \tag{8.5}$$

Mit dem Schalen-RVE werden die gleichen synthetischen Materialversuche durchgeführt wie im vorherigen Kapitel mit den analytischen Materialmodellen (vgl. Kapitel 7.2.6). Durch die Verwendung des numerischen Homogenisierungsschemas ist für jeden zu generierenden Datenpunkt ein FE-Problem zu lösen. Während die Rechenzeiten bei der Datengenerierung in den bisherigen Kapiteln keine

Name	Topologie	Anzahl uniaxialer Pfade pro Verzerrungskomponente			Datenpunkte
		Relaxation	Zyklisch	Kriechen	
KNN ₅	25-60-8	5	5	5	2601
KNN ₁₀	25-60-60-8	10	10	10	5415
KNN ₂₀	25-80-80-8	20	20	20	10682

Tabelle 8.20: Übersicht der verwendeten KNN und der zugrunde liegenden Trainingsdaten

wesentliche Rolle spielten, macht sich dies je nach Komplexität des RVEs deutlich bemerkbar. Umso wichtiger ist es, eine effiziente Methode einzusetzen, mit der sich aus möglichst wenigen Daten ein KNN-Materialmodell konstruieren lässt.

Zunächst wird untersucht, wie sich das Training eines KNN rein datenbasiert – also ohne Verwendung von Nebenbedingungen – verhält. Dazu werden drei KNN betrachtet, die sich in der Topologie sowie in der Anzahl der verwendeten Pfade (5,10,20) unterscheiden. Die Netztopologien sowie die Anzahl der uniaxialen Pfade pro Verzerrungskomponente ($n_{\text{str}} = 8$) sind in Tabelle 8.20 dargestellt. Es wird lediglich ein Datensatz erzeugt, auf dem die KNN trainiert werden. Pro Netz erfolgen fünf Trainingsdurchläufe, wobei jeweils das KNN mit dem kleinsten Testfehler zur Auswertung herangezogen wird. Die Generalisierungsfähigkeit wird mit den in Tabelle 8.21 gegebenen sechs Versuchen geprüft, wobei Pfad-Nr. 1-3 uniaxial und Pfad-Nr. 4-6 biaxial gefahren werden.

Die Ergebnisse der drei Netze sind in Abbildung 8.17 dargestellt. Für die uniaxialen Validierungspfade liefert bereits das mit wenigen Daten trainierte KNN₅ brauchbare Ergebnisse. Im Gegensatz zu den dreidimensionalen Untersuchungen aus Kapitel 8.2.1 zeigt sich hier, dass ein KNN rein datenbasiert – also ohne Ne-

Pfad Nr.	Versuchsart	Komponenten	Werte	Δt [s]	T [s]
1	Zyklisch	ε_{11}	0,01 [–]	8	300
2	Kriechen	m_{11}	100 [Nmm/mm]	8	300
3	Relaxation	γ_1	0,01 [–]	8	300
1	Zyklisch	$\varepsilon_{11}, \varepsilon_{22}$	0,01, 0,01 [–]	8	300
1	Kriechen	m_{11}, m_{22}	100, 100 [Nmm/mm]	8	300
4	Relaxation	γ_1, γ_2	0,01, 0,01 [–]	8	300

Tabelle 8.21: Verwendete Validierungspfade unter Angabe der gewählten Parameter, mit Pfaden Nr. 1-3 uniaxial und Pfad-Nr. 4-6 biaxial

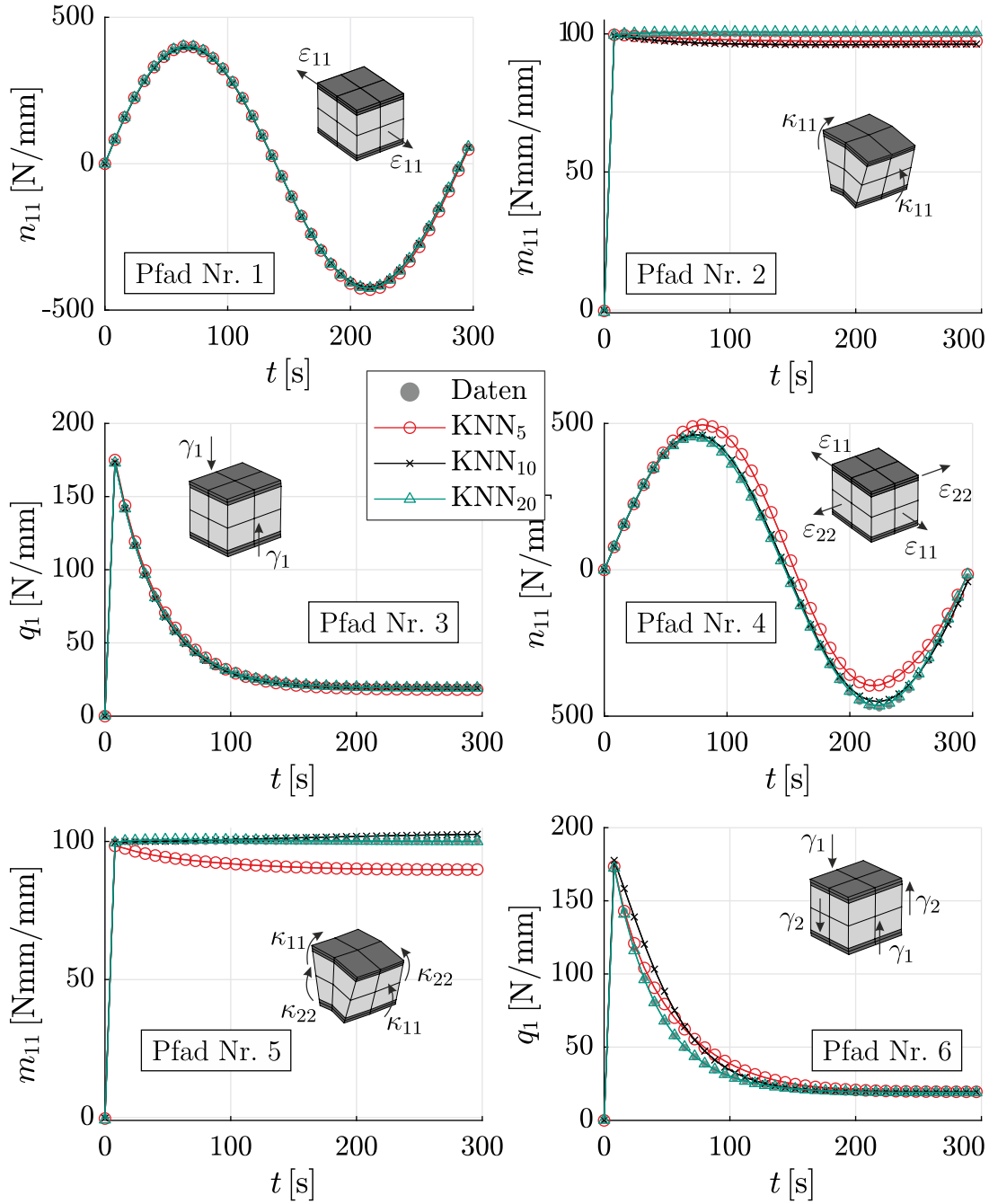


Abbildung 8.17: Approximation der in Tabelle 8.21 gegebenen Pfade mit drei rein auf Datenbasis trainierten KNN

benbedingung – auch für biaxiale Pfade gute Approximationen liefert. Mit dem KNN₂₀ sind bereits kaum noch Abweichungen von der Referenzlösung erkennbar. Dies lässt die Vermutung zu, dass die Approximation der viskoelastischen Konstitutivbeziehung der Schale zwar eine höherdimensionale Problemstellung darstellt, der Trainingsraum jedoch durch die Entkopplung der verschiedenen Ver-

zerrungsanteile (vgl. Gleichungen (7.31) und (7.30)) und die gewählten uniaxialen Trainingspfade gut abgedeckt wird. Es kann daher festgehalten werden, dass sich bei ausreichender Datenmenge – im vorliegenden Fall etwa 11,000 Datenpunkte – rein datenbasiert ein brauchbares KNN-Materialmodell am Materialpunkt konstruieren lässt.

Im Rahmen dieser Untersuchung zeigt sich, dass die Tangenten-Nebenbedingung theoretisch nicht zwingend erforderlich ist, um ein funktionsfähiges Materialmodell zu konstruieren. Für die Generierung von 11,000 Datenpunkten sind jedoch ebenso viele FE-Rechnungen notwendig, um die homogenisierten Größen auf dem RVE zu berechnen. Daher ist es von besonderem Interesse zu prüfen, ob sich durch die Einbindung zusätzlicher Tangenten-Informationen – analog zu Kapitel 8.2.2 – die Anzahl der erforderlichen Primärdaten reduzieren lässt.

Zu diesem Zweck wird das KNN_5 zusätzlich mit $P^S = 1000$ Nebenbedingungspunkten trainiert, die mit einem Penalty-Parameter $\rho = 1$ berücksichtigt werden. Zur Abgrenzung wird dieses Netz als $\text{KNN}_{5_{\text{sob}}}$ bezeichnet, wobei die zugrunde liegende Topologie mit 25–80–80–8 derjenigen des KNN_{20} entspricht. Im Gegensatz zu den bisherigen Untersuchungen werden hier nicht die jeweils besten Approximationen betrachtet, sondern zur Beurteilung der Robustheit des Trainings die gemittelten Ergebnisse aus fünf Durchläufen sowie die zugehörigen Standardabweichungen s_{20} und $s_{5_{\text{sob}}}$ ausgewertet.

Die Resultate sind in Abbildung 8.18 dargestellt. Es wird deutlich, dass das SOBOLEV-Training bei geringerer Datenbasis ein robusteres Verhalten ermöglicht. Das $\text{KNN}_{5_{\text{sob}}}$ liefert unabhängig vom Trainingslauf stabile Approximationen mit deutlich reduzierter Streuung gegenüber der Referenzlösung. Zudem genügt hierfür etwa ein Viertel der für das KNN_{20} erforderlichen FE-Rechnungen zur Datengenerierung. Gerade bei komplexeren RVEs kann dies zu erheblichen Einsparungen in der absoluten Rechenzeit führen. Abschließend sei angemerkt, dass die hier angegebenen Mittelwerte und Standardabweichungen aufgrund der beschränkten Anzahl an Trainingsdurchläufen nicht als vollständig konvergent anzusehen sind. Da im Rahmen dieser Arbeit jedoch die Zielsetzung besteht, unabhängig von zufälligen Einflüssen wie Initialisierungen der Gewichte oder Pfadvariationen verlässliche KNN-Approximationen zu erzielen, wird auf eine sehr große Zahl von Trainingsläufen verzichtet.

8.3.2 Numerisches Beispiel: Geschichtete Zylinderschale

Für die Prüfung der Brauchbarkeit der zuvor trainierten KNN-Modelle auf Strukturebene wird das in Abbildung 8.23 dargestellte Beispiel einer geschichteten Zylinderschale untersucht. Der Radius vom Ursprung zur Mittelfläche der Schale beträgt $R = 200$ mm, die Länge $L = 600$ mm und die Gesamthöhe $h = 20$ mm.

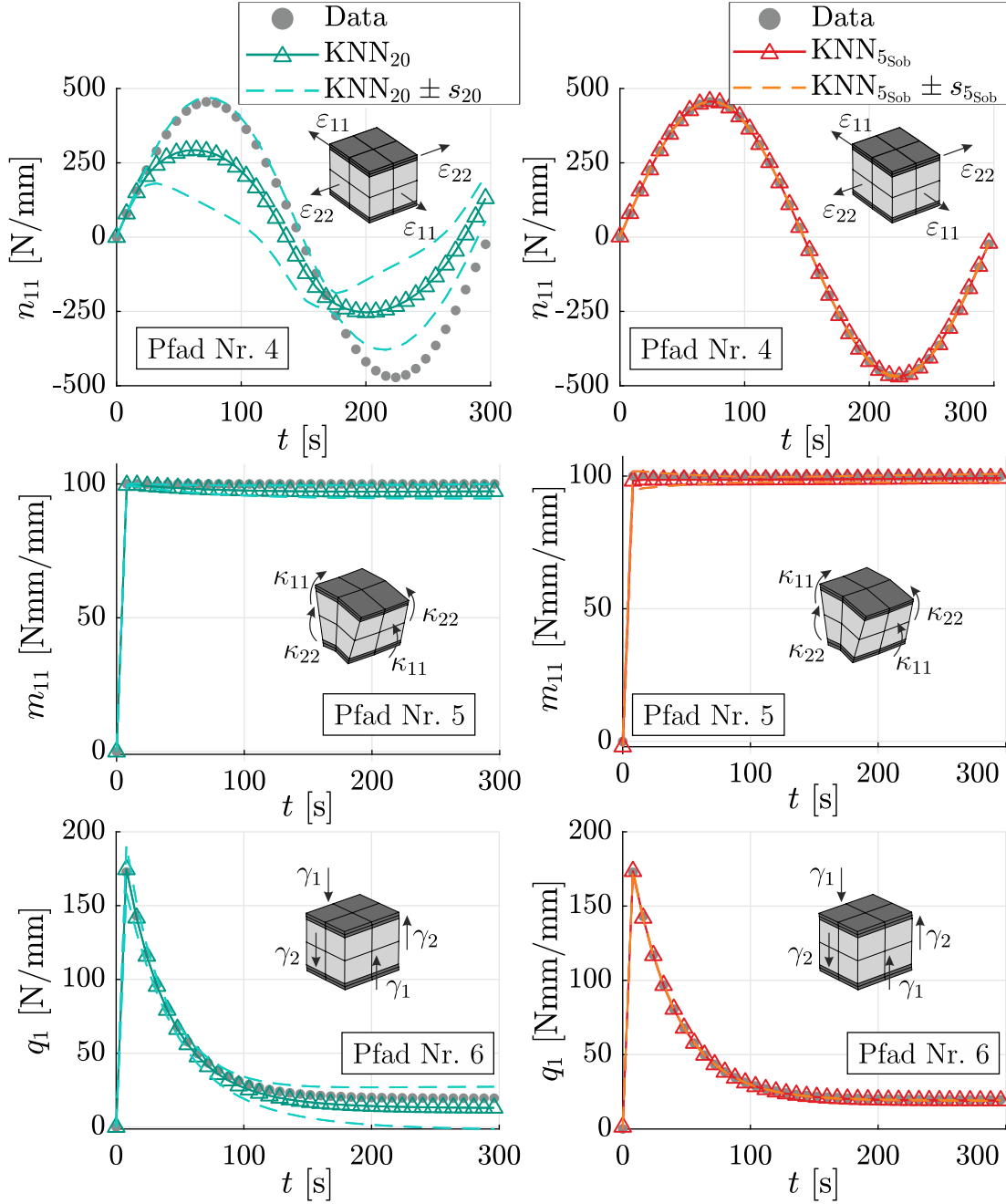


Abbildung 8.18: Approximation der in Tabelle 8.21 gegebenen Pfade mit dem „rein auf Datenbasis“ (links) und mit Nebenbedingungen trainierten KNN (rechts)

Wegen der Doppelsymmetrie des Systems wird lediglich ein Viertel der Struktur modelliert, wozu geeignete Randbedingungen auf den Symmetrieachsen gesetzt werden. Der Rand bei $y = 600$ mm ist vollständig eingespannt. Auf der Oberseite der Schale wird eine zeitabhängige Linienlast $q(t)$ aufgebracht, deren zeitlicher Verlauf weiter unten präzisiert wird. Die Schichtung entspricht dem in

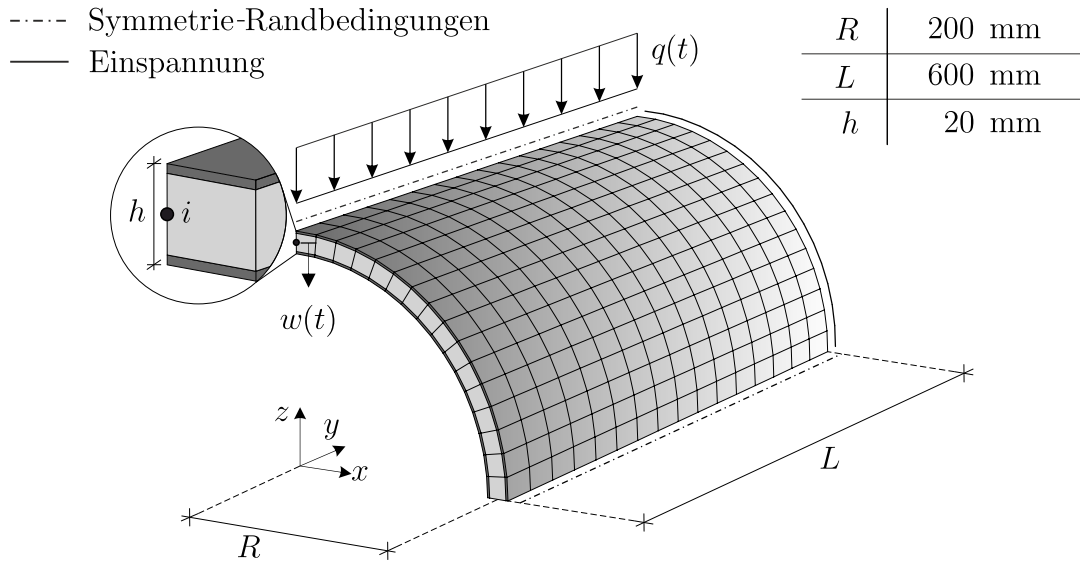


Abbildung 8.19: Geometrie, Belastung und Randbedingungen der doppel-symmetrischen geschichteten Zylinderschale; hier dargestellt Diskretisierung mit 3D-Volumenelementen

Abbildung 8.16 verwendeten RVE, die dazugehörigen Materialparameter sind in Tabelle 8.19 angegeben.

Im Folgenden werden verschiedene Modellierungsvarianten untersucht: zum einen eine vollständige 3D-Lösung, bei der das System mit 16×16 27-Knoten-Volumenelementen in Längen- und Umfangsrichtung sowie je einem Element pro Schicht über die Höhe diskretisiert wird. Zum anderen wird eine Schalen-Mehrskalenlösung (FE^2) verwendet, bei der die Schale mit 16×16 Elementen in der Schalenebene diskretisiert und an jedem der 2×2 GAUSS-Punkte das oben genannte RVE definiert wird (Schalen-Formulierung vgl. Kapitel 2.4). Zuletzt wird die entsprechende Schalenlösung, bei der das Materialverhalten durch die zuvor trainierten KNN-Materialmodelle (KNN_{20} und $KNN_{5_{\text{Sob}}}$) approximiert wird, berücksichtigt. Bei den Schalenmodellen wird die Belastung in der Mittelfläche aufgebracht. Im 3D-Modell erfolgt eine Aufteilung mit jeweils $q(t)/2$ auf die Knoten der Ober- und Unterseite (oberer Knoten: $z = R + h/2$, unterer Knoten: $z = R - h/2$). Alle Berechnungen werden unter der Annahme kleiner Verformungen mit einer geometrisch linearen Theorie durchgeführt.

Von Interesse sind zwei Lastfälle. Im ersten Fall wird ein Kriechversuch mit konstanter Last und anschließender vollständiger Entlastung über

$$q(t) = \begin{cases} 10 \text{ N/mm}, & t \in [10, 350] \text{ s}, \\ 0, & t \in (350, 700] \text{ s}, \end{cases} \quad (8.6)$$

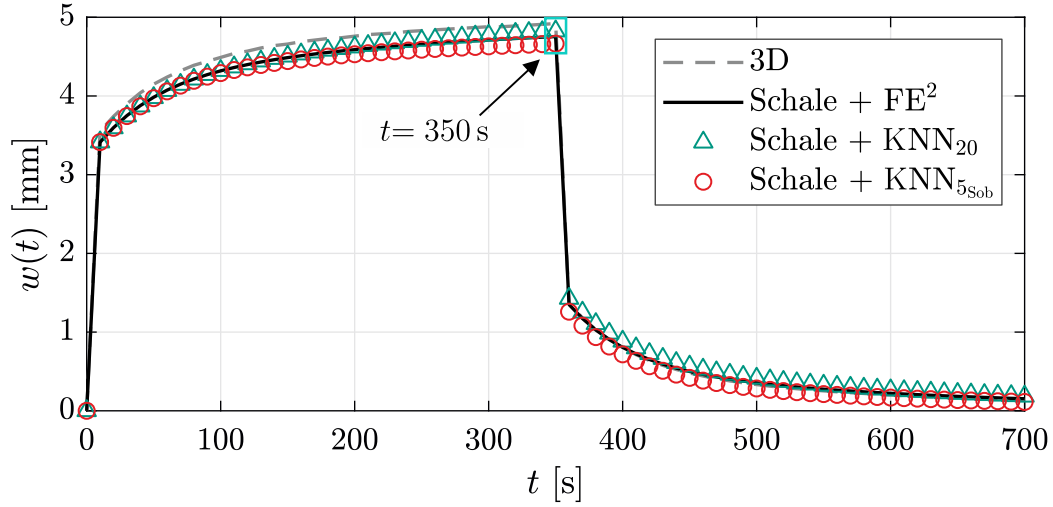


Abbildung 8.20: Ergebnisse Kriechverläufe der Durchbiegung während der Be- und Entlastungsphase mit den vier Modellierungsvarianten

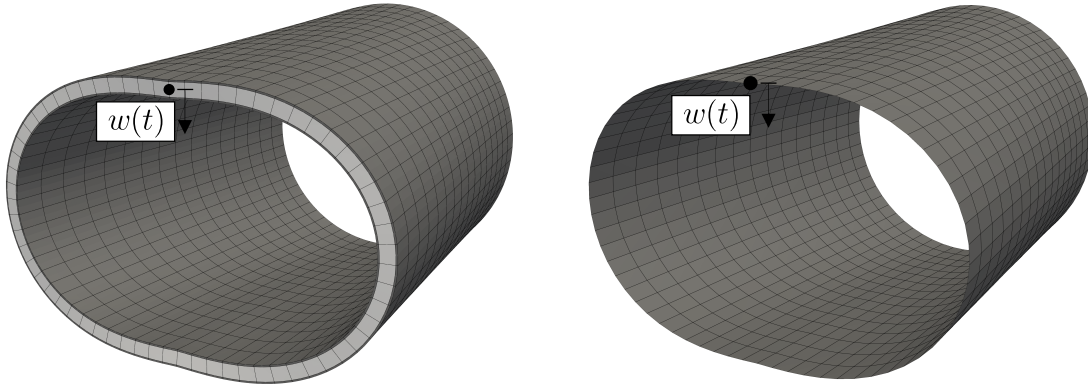


Abbildung 8.21: Zehnfach überhöhte Verformungsfigur des vollständigen Systems für $t = 350$ s beim Kriechversuch; Ergebnisse des 3D-Modells (links) und Schale mit dem $\text{KNN}_{5\text{Sob}}$ -Materialmodell (rechts)

vorgegeben. Im zweiten Fall folgt die Last einem Sägezahnverlauf

$$q(t) = \begin{cases} \pm 10 \text{ N/mm}, & t \in [0, 200] \text{ s}, \\ \pm 5 \text{ N/mm}, & t \in (200, 400] \text{ s}, \\ \pm 2,5 \text{ N/mm}, & t \in (400, 600] \text{ s}, \\ 0, & t \in (600, 700] \text{ s}. \end{cases} \quad (8.7)$$

In beiden Fällen wird die vertikale Verschiebung $w(t)$ am Knoten der Mittelfläche (vgl. Abbildung 8.23) ausgewertet. Alle Berechnungen werden mit einer Zeitschrittweite von $\Delta t = 10$ s durchgeführt.

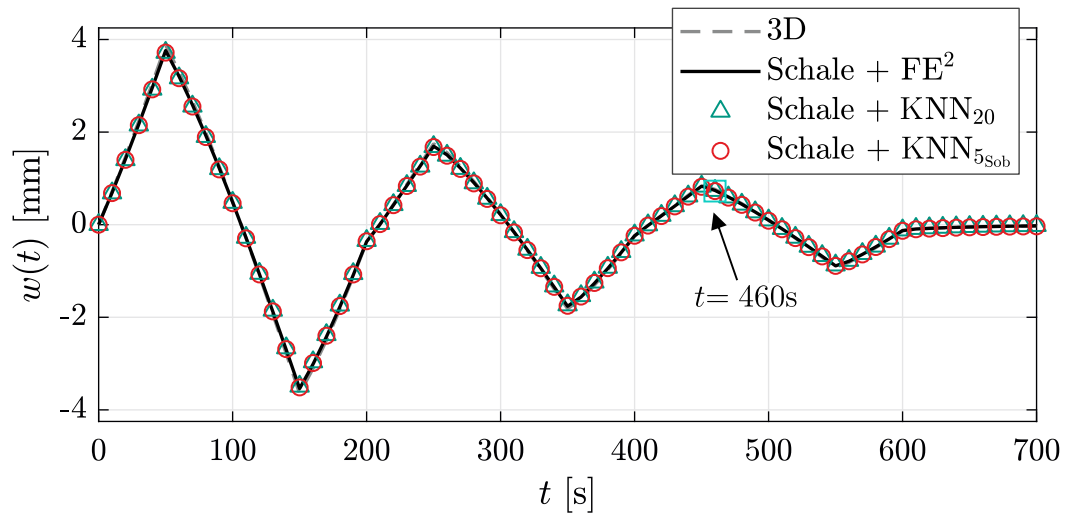


Abbildung 8.22: Ergebnisse der Durchbiegung über die Zeit für den Sägezahn-Belastungsverlauf mit den vier Modellierungsvarianten

In Abbildung 8.20 sind die Verläufe der Durchbiegung über die Zeit für die Be- und Entlastungsphasen dargestellt. Es zeigt sich eine grundsätzlich gute Übereinstimmung zwischen allen vier Modellen. Die Unterschiede zwischen der 3D- und der FE^2 -Lösung sind auf lokale Effekte, insbesondere im Bereich der Lasteinleitung, zurückzuführen, die bei der Schalenmodellierung nur in der Mittelfläche berücksichtigt werden können. Beide KNN-Materialmodelle zeigen in diesem Beispiel eine hohe numerische Stabilität und liefern sehr gute Approximationen. Damit lässt sich schließen, dass die Einbindung der Tangenteninformationen für ein robustes Verhalten innerhalb der FEM nicht zwingend erforderlich ist und ein rein datenbasiertes KNN bei hinreichender Datenmenge im Rahmen dieses Beispiels ebenfalls ein brauchbares Materialmodell darstellt. In Abbildung 8.21 sind die Verformungsfiguren des Gesamtsystem zum Zeitpunkt $t = 350$ s für das 3D- und Schalen-KNN-Modell überhöht dargestellt.

Ein bislang nicht betrachteter Belastungspfad ist der in Gleichung (8.7) definierte Sägezahnverlauf, der über die Zeit zu unterschiedlichen Dehnraten führt. Auch hier zeigt Abbildung 8.22 eine sehr gute Übereinstimmung der Ergebnisse aller Modellvarianten. Da die Belastung hier nicht sprunghaft, sondern linear aufgebracht wird, ist der Einfluss lokaler Effekte weniger ausgeprägt, sodass das 3D-Modell und die FE^2 -Lösung nahezu identische Ergebnisse liefern.

Ein weiterer zu untersuchender Aspekt betrifft das numerische Lösungsverhalten: Das KNN stellt per Konstruktion ein nichtlineares Materialmodell dar, weshalb im Rahmen der FEM – im Gegensatz zu den beiden linearen Referenzmodellen – stets eine (Makro) NEWTON-Iteration erforderlich ist. Die hier verwendete Schalenformulierung basiert auf einer FE-Formulierung, die symmetrische

Iterationsschritt	Norm des Residualvektors $\ \mathbf{g}\ $	
	KNN ₂₀	KNN _{5_{Sob}}
1	2.67E+03	1.25E+05
2	1.36E+00	1.78E+00
3	5.00E-04	8.48E-04
4	2.27E-07	5.36E-07
5	3.80E-10	6.93E-10
6	2.99E-10	4.36E-10

Tabelle 8.22: Konvergenz des globalen Residualvektors für die beiden KNN-Materialmodelle zum Zeitpunkt $t = 450$ s auf $t = 460$ s, vgl. Abbildung 8.22

Systemmatrizen liefert. Daher wird ausschließlich der symmetrische Anteil der KNN-Materialmatrix berücksichtigt

$$\mathbf{D}_{\text{sym}}^{\text{KNN}} = \frac{1}{2} \left[(\mathbf{D}^{\text{KNN}})^T + \mathbf{D}^{\text{KNN}} \right], \quad (8.8)$$

wobei eine exakt symmetrische Matrix das Konvergenzverhalten nicht beeinflusst und das NEWTON-Verfahren quadratisch konvergiert. Ist die Materialmatrix dagegen unsymmetrisch, passt die modifizierte Tangente nicht mehr exakt zu dem berechneten Residuum, was zu subquadratischer Konvergenz führt. Das Konvergenzverhalten des globalen Residualvektors unter Verwendung der beiden KNN-Modelle ist in Tabelle 8.22 für den Zeitschritt $t = 450$ s auf $t = 460$ s dargestellt. Beide Modelle zeigen ein quadratisches Konvergenzverhalten, wobei das Gleichgewicht nach fünf Iterationen erreicht wird. Für das KNN_{5_{Sob}} ist dies erwartbar, da die Tangenten-Nebenbedingung nicht nur die Einträge, sondern implizit auch die Symmetrie der KNN-Materialmatrix schwach einfordert. Für das rein datenbasierte KNN₂₀ ist dies hingegen bemerkenswert, da diese Information nicht in den Trainingsdaten enthalten ist – ein Hinweis darauf, dass bei ausreichender Datenmenge auch die Materialtangente selbst zuverlässig approximiert wird. Hierbei sei jedoch nochmals erwähnt, dass für das KNN_{5_{Sob}} lediglich ein Viertel der für das KNN₂₀ verwendeten Daten benötigt wurden.

Abschließende Bemerkungen zum Beispiel

Das Ziel der Arbeit ist die Entwicklung einer Methodik, die in Bezug auf Rechenzeiten gegenüber einem vollständigen 3D-Modell oder einer FE²-Lösung eine effiziente Alternative darstellt. Offensichtlich besteht keine Notwendigkeit, das hier

vorgestellte Beispiel mit der besagten Methodik zu modellieren. Durch die einfache Mikrostruktur, die mit wenigen Elementen diskretisiert werden kann, ist ein vollständige aufgelöstes Modell hier völlig ausreichend. Es sei jedoch angemerkt, dass trotz der benötigten NEWTON-Iteration durch die geringere Anzahl an Freiheitsgrade ähnliche Zeiten von unter einer Minute für die vollständige Berechnung der Verschiebungs-Zeitkurven für das 3D- und die KNN-Modelle benötigt werden. Die Berechnungen wurden in der verwendeten numerischen Umgebung auf einem PC mit einem 12th Gen Intel(R) Core(TM) i9-12900 (2.40 GHz) und 32 GB Arbeitsspeicher durchgeführt. Als weitere Alternative kann eine geschichtete Schalenformulierung nach KLINKEL ET AL. [71] verwendet werden, die hier als Vergleichslösung nicht berücksichtigt wurde, jedoch vergleichbare, wenn nicht sogar geringere Rechenzeiten für das entsprechende Beispiel aufweisen würde.

8.3.3 Numerisches Beispiel: Glasfaserverstärkte Platte

Das folgende Beispiel demonstriert das Potential der vorgeschlagenen Methodik an einer realitätsnahen technischen Anwendung, bei der das Werkstoffverhalten auf Basis realer Materialparameter beschrieben wird und die Mikrostruktur eine deutlich höhere Komplexität aufweist. In diesem Fall führen vollaufgelöste 3D-Simulationen bereits zu deutlich erhöhten Rechenzeiten. Betrachtet wird die in Abbildung 8.23 dargestellte faserverstärkte Platte unter einer konstanten Flächenlast. Faserverstärkte Verbundkonstruktionen treten häufig im Leichtbau auf,

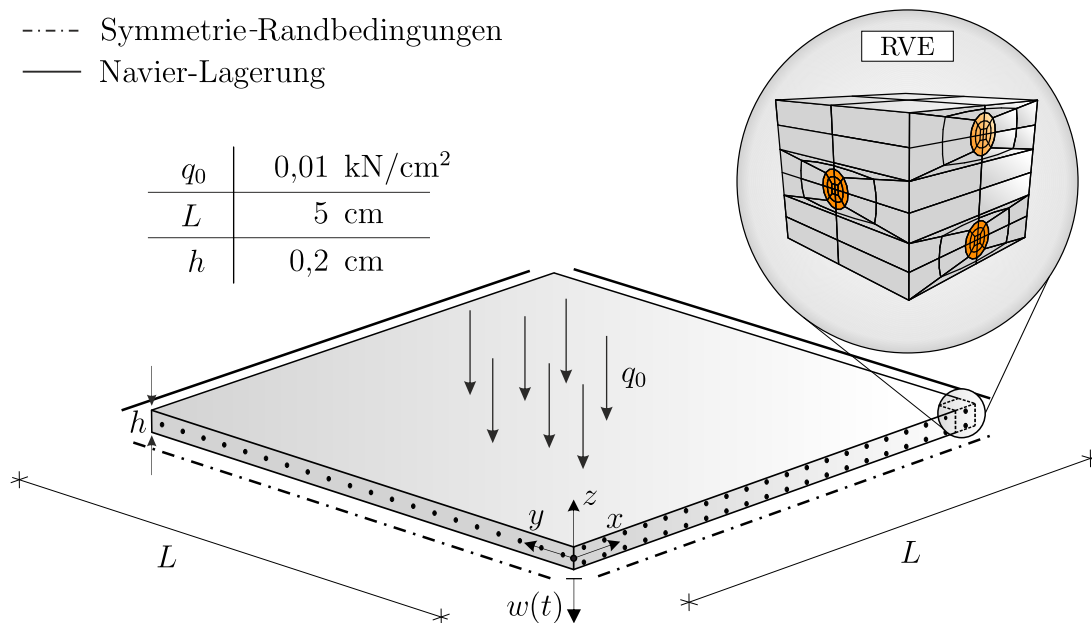


Abbildung 8.23: Geometrie, Belastung und Randbedingungen der doppel-symmetrischen faserverstärkten Platte

	Initialer E-Modul E_0 [kN/cm ²]	Poisson- Zahl ν_0 [-]	MAXWELL Element i	Modul- anteil $\mu_i = E_i/E_0$ [-]	Relaxations- zeit τ_i [s]
Faser	7300	0,18	-	-	-
Matrix	125	0,30	-	-	-
			1	0,40	565
			2	0,17	83
			3	0,10	21
			4	0,03	4

Tabelle 8.23: Verwendete Materialparameter der elastischen Fasern und viskoelastischen Matrix

eine ausführliche Übersicht hierzu findet sich beispielsweise in SCHÜRMANN [120]. Typischerweise werden elastische Fasern in eine Polymer- oder Kunststoffmatrix eingebettet, die viskoelastisches und häufig auch temperaturabhängiges Verhalten zeigt.

Die Platte besitzt makroskopisch die Abmessungen $2L = 10$ cm in x - und y -Richtung sowie eine Gesamthöhe von $h = 0,2$ cm. Aufgrund der Symmetrie wird nur ein Viertel des Gesamtsystems modelliert, wobei auf den Knoten der xz - und yz -Ebene die entsprechenden Randbedingungen gesetzt werden. Auf den verbleibenden Rändern wird eine Navier-Lagerung mit $w(z) = 0$ vorgegeben. Das Grundmaterial der Platte besteht aus einer weichen Kunststoffmatrix, die mit Glasfasern verstärkt wird.

Die Materialparameter der beiden Werkstoffe sind in Tabelle 8.23 zusammengefasst. Die Fasern werden mit typischen Werten für Glasfasern modelliert, während die Parameter der Matrix aus experimentellen Relaxationsversuchen an 3D-gedruckten Proben des Kunststoffes *VeroBlackPlus*TM stammen. Sie wurden am Institut für Statik und Dynamik der Universität der Bundeswehr München bestimmt und mittels linearer Regression für ein verallgemeinertes MAXWELL-Modell angepasst.⁵⁷

Abbildung 8.24 (links) zeigt den Aufbau der Mikrostruktur und zugleich das RVE, das zur Generierung der Trainingsdaten dient. Die Abmessungen betragen $l_x = l_y = 0,25$ cm bei einer Gesamthöhe von $h = 0,2$ cm. Das RVE ist durch den Schichtaufbau $90^\circ/0^\circ/90^\circ$ charakterisiert, wobei 90° der x -Richtung zugeordnet wird und alle Schichten gleiche Dicken von $h/3$ besitzen. Jede Schicht enthält eine

⁵⁷Da nur fünf Relaxationsversuche zur Verfügung standen, war aufgrund der begrenzten Datenmenge keine Konstruktion eines KNN-basierten Materialmodells möglich.

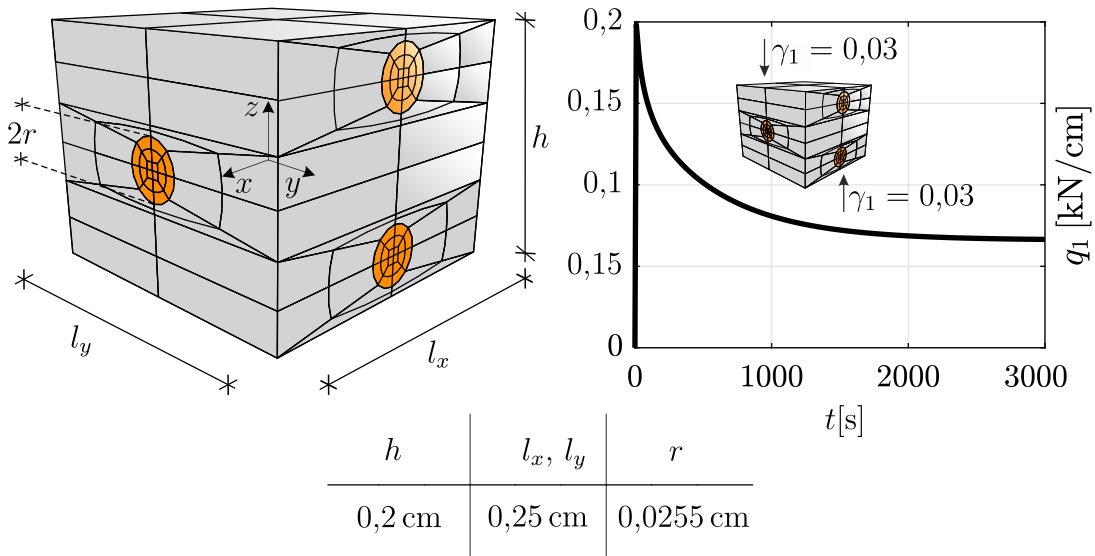


Abbildung 8.24: Geometrie und Abmessungen des verwendeten Schalen-RVEs (links) sowie Darstellung eines exemplarischen Relaxationsversuchs unter Vorgabe der Schubverzerrung γ_1 (rechts)

Faser mit Radius $r = 0,0255 \text{ cm}$, was einem Faservolumenanteil von $v_F = 13 \%$ entspricht.

Das RVE wird mit 216 27-Knoten Volumenelementen diskretisiert. Zusätzlich werden fünf Lagrange-Parameter zur Einforderung der kinematischen Zwangsbedingungen sowie periodische Verschiebungsrandbedingungen verwendet. Das vollständige 3D-Modell setzt sich aus jeweils zwanzig RVEs in x - und y -Richtung zusammen und umfasst somit etwa 86,000 Elemente und rund 2,1 Millionen Freiheitsgrade. Für die Schalenmodelle auf der Makroskala – sowohl für die FE²- als auch für die KNN-Lösung – haben sich im Rahmen von Konvergenzuntersuchungen 12×12 Elemente als ausreichend erwiesen.

Für die Konstruktion des KNN-Materialmodells werden die vorgeschriebenen Membran- und Schubverzerrungen auf $[-0,03, 0,03]$ sowie die Krümmung auf $[-0,3, 0,3] \text{ 1/cm}$ beschränkt. Der exemplarisch dargestellte Relaxationsversuch in Abbildung 8.24 (rechts) weist kaum wahrnehmbare Spannungsänderungen nach $\bar{T} \approx 3000 \text{ s}$ auf, was als Dauer aller Versuche festgelegt wird. Angemerkt sei dabei, dass hier aufgrund mehrerer MAXWELL-Terme mit unterschiedlichen Relaxationszeiten grundsätzlich ein unterschiedliches Abklingverhalten für Kriechen und Relaxation zu verzeichnen ist, hier jedoch keine individuelle Anpassung der Versuchsdauern vorgenommen wird, da ohnehin nur eine geringe Datenmenge erforderlich ist, um in Kombination mit der Tangenten-Nebenbedingung hinreichende Genauigkeit zu erzielen. Die Versuchsdauern werden mit $\Delta \bar{t} \in [15, 50] \text{ s}$ diskretisiert. Durch die Orthotropie des RVEs muss bei spannungsgesteuerten

Schalenschnittgröße		Grenzen des Eingangsraums
$\bar{n}_{11}$	$\in$	$[-4, 4]$ kN/cm
$\bar{n}_{22}$	$\in$	$[-2, 2]$ kN/cm
$\bar{n}_{12}$	$\in$	$[-0,3, 0,3]$ kN/cm
$\bar{m}_{11}$	$\in$	$[-0,17, 0,17]$ kNcm/cm
$\bar{m}_{22}$	$\in$	$[-0,02, 0,02]$ kNcm/cm
$\bar{m}_{12}$	$\in$	$[-0,01, 0,01]$ kNcm/cm
$\bar{q}_1$	$\in$	$[-0,20, 0,20]$ kN/cm
$\bar{q}_2$	$\in$	$[-0,20, 0,20]$ kN/cm

Tabelle 8.24: Grenzen des Eingangsraums für die acht Schalenschnittgrößen, ermittelt aus Kriechversuchen am RVE

Kriechversuchen berücksichtigt werden, dass unterschiedliche Steifigkeiten vorliegen. Entsprechend muss für jede Richtung im Rahmen einer Voruntersuchung die maximal zulässige Größe der Streckenschnittgrößen geprüft werden. Die acht gewählten Wertebereiche sind in Tabelle 8.24 angegeben.

Für dieses Beispiel wird lediglich ein KNN mit zwei verdeckten Schichten und 80 Neuronen pro Schicht verwendet, was sich auch hier als geeignete Topologie gezeigt hat. Es werden wieder lediglich fünf uniaxiale Relaxations-, Kriech-, und zyklische Versuche pro Verzerrungskomponente gefahren, was hier 4173 Datenpunkte ergab. Zusätzlich wird die Tangenten-Nebenbedingung an 1000 zufälligen Datenpunkten mit einem Penalty-Parameter $\rho = 1$ eingefordert. Sonstige, rein auf Daten trainierte KNN werden im Weiteren nicht betrachtet.

Im Rahmen eines Kriechversuchs soll zunächst die Lösung der drei Modelle (3D, Schale mit FE², Schale mit KNN) verglichen werden. Dafür wird die Last $q_0 = 0,01$ kN/cm² im ersten Zeitschritt $t_1 = 75$ s aufgebracht und bis $t = 3000$ s konstant gehalten, mit einer Zeitdiskretisierung von $\Delta t = 75$ s. Auch hier zeigt sich in den in Abbildung 8.25 dargestellten Ergebnissen eine gute Übereinstimmung der drei Modelle für die in Abbildung 8.23 dargestellte Durchbiegung in der Mitte der Platte $w(t)$. Für die initiale sowie finale Verschiebung im Rahmen der gewählten Dauer ergeben sich die in Tabelle 8.25 zusammengefassten Durchbiegungen. Dabei überschätzt die KNN-Lösung die initiale (+4,4 %) und unterschätzt die finale Verformung (−2,0 %), verglichen mit der FE²-Lösung, die hier als Referenzlösung dient. Beide Ergebnisse sind jedoch als zufriedenstellende Approximation einzuordnen.

Im Rahmen dieses Beispiels zeigt sich zudem eine deutliche Reduktion der Re-

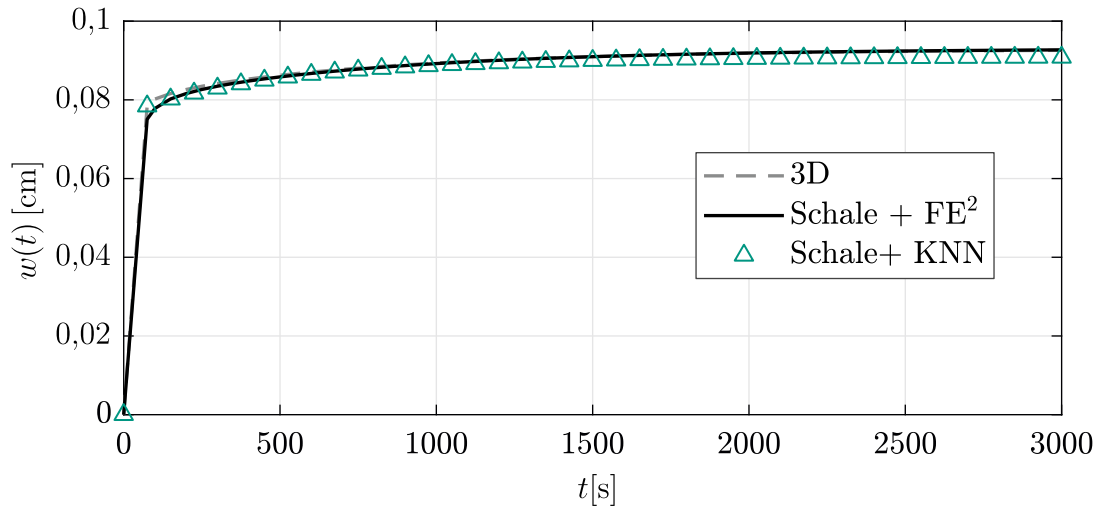


Abbildung 8.25: Ergebnisse für die Kriechverformung $w(t)$ für die drei verwendeten Modelle

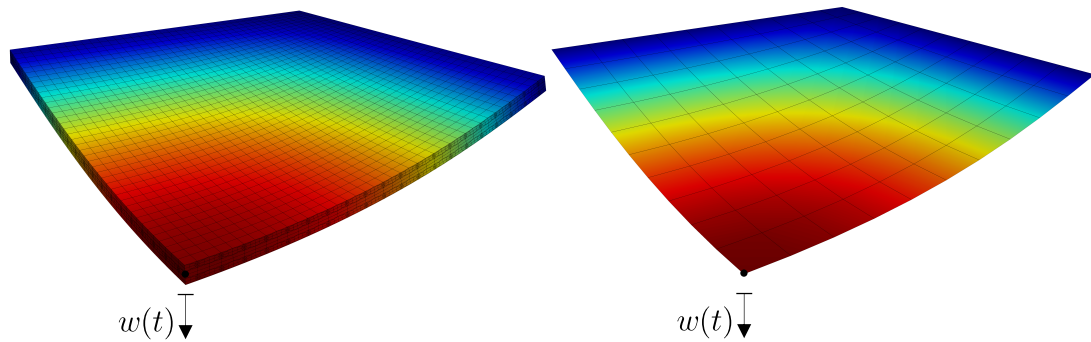


Abbildung 8.26: Verformungsfiguren des Viertelsystems der 3D- und Schalen-Lösung unter Verwendung des KNN für den Kriechversuch zum Zeitpunkt $t = 3000$ s

chenzeit unter Verwendung des KNN-Ersatzmodells. Dafür sind in Tabelle 8.26 die relativen Zeiten für die Berechnung eines Lastschrittes bis zum Erreichen der Konvergenz angegeben. Erneut wird darauf hingewiesen, dass für die 3D- und FE^2 -Lösung das Gleichungssystem im Rahmen der FEM nur einmal, beim KNN aufgrund der Nichtlinearität des Modells mehrfach gelöst werden muss. Dennoch ergeben sich um Faktor 100 verbesserte Rechenzeiten im Vergleich zu FE^2 und um Faktor 450 im Vergleich zum 3D-Modell. In absoluten Zahlen gesprochen, bewegt sich im Rahmen der verwendeten Computerumgebung die Berechnung eines Zeitschritts mit dem KNN im Bereich von etwa einer Sekunde, während die FE^2 -Lösung einige Minuten und die 3D-Lösung mehrere Minuten pro Zeitschritt benötigt. Dabei ist selbstverständlich zu berücksichtigen, dass auch das Training des KNN Rechenzeit in Anspruch nimmt. Diese liegt ungefähr in der Größenordnung der vollständigen Berechnung der Zeit-Verformungskurve der FE^2 -Lösung.

	$w(t=75 \text{ s}) [\text{cm}]$	$w(t=3000 \text{ s}) [\text{cm}]$
3D-Modell	7.84E-02	9.26E-02
Schale + FE ²	7.51E-02	9.27E-02
Schale + KNN	7.84E-02	9.08E-02

Tabelle 8.25: Anfangs- und Endverformung der Durchbiegung im Kriechversuch

	KNN	FE ²	3D
$t_{\text{rel}} = t/t_{\text{KNN}}$	1	≈ 100	≈ 450

Tabelle 8.26: Relative Rechenzeiten pro Zeitschritt bis zur Konvergenz, normiert auf das KNN-Modell

Ein entscheidender Vorteil besteht jedoch darin, dass das KNN nur einmal trainiert werden muss und anschließend beliebig oft als effizientes Ersatzmodell in der FEM eingesetzt werden kann.

Ein weiterer Aspekt betrifft die Verwendung unterschiedlicher Zeitschrittweiten in der FEM. Dafür wird als zweiter Lastfall eine zyklische Belastung

$$q(t) = q_0 \sin\left(\frac{4\pi t}{3000}\right), \quad t \in [0, 3000] \text{ s} \quad (8.9)$$

mit der Amplitude $q_0 = 0,01 \text{ kN/cm}^2$ aufgebracht, was gerade zwei vollen Belastungsperioden entspricht. Die Berechnungen mit dem KNN werden mit Zeitschrittweiten $\Delta t = \{50; 100; 125; 150\} \text{ s}$ durchgeführt. Die Ergebnisse sind in Abbildung 8.27 dargestellt, wobei als Referenz die FE²-Lösung verwendet wird.

Erwartungsgemäß liefert das KNN am Rande des Trainingsraums der Zeitschrittweite ($\Delta \bar{t} \in [15, 50] \text{ s}$) schlechtere Approximationen, was sich an der leichten Abweichung der Lösung für $\Delta t = 50 \text{ s}$ erkennen lässt. Für alle weiteren Zeitschrittenweite stimmen die KNN und die Referenzlösung zufriedenstellend überein, auch mit der Zeitschrittweite $\Delta t = 150 \text{ s}$, die den oberen Rand des Trainingsraums darstellt. Entsprechend kann festgehalten werden, dass mit dem KNN auch unabhängig von der Zeitschrittweite innerhalb der FE-Simulation zuverlässige Berechnungen durchgeführt werden können.

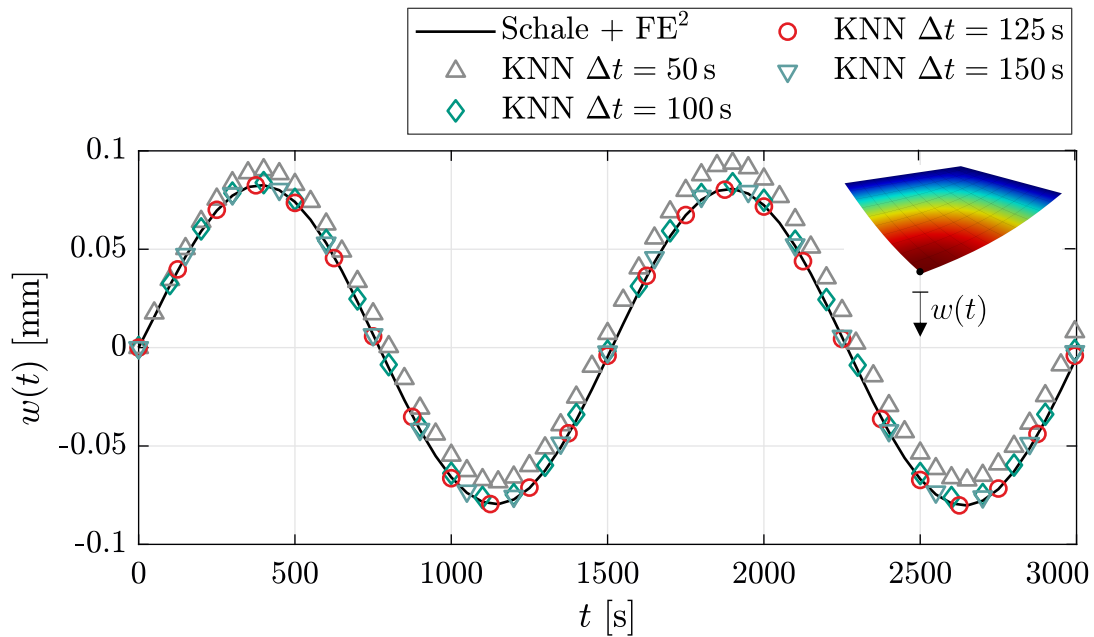


Abbildung 8.27: Verlauf der Durchbiegung $w(t)$ infolge einer zyklischen Belastung; Vergleich der Ergebnisse des FE^2 - und KNN-Schalenmodells bei verschiedenen Zeitschrittweiten

8.3.4 Schalen-RVE mit Einschluss: Voruntersuchungen am Materialpunkt

Im Rahmen des letzten Beispiels soll geprüft werden, ob der bereits hochdimensionale Eingangsraum des KNN – bei der Verwendung als Schalen-Materialmodell – um zusätzliche Parameter des RVEs erweitert werden kann. Dass dies grundsätzlich möglich ist, wurde bereits erfolgreich im eindimensionalen Fall am Beispiel des weichen Aktuators in Kapitel 8.1.5 gezeigt. Dies ermöglicht die Konstruktion parametrisierbarer effektiver Stoffgesetze, was eine Motivation der vorliegenden Arbeit darstellt. Im Rahmen des Designprozesses von Bauteilen können damit effizient die Einflüsse verschiedener mikrostruktureller geometrischer oder materieller Eigenschaften auf das makroskopische Verhalten der Struktur untersucht werden.

Im Folgenden wird ein Beispiel in Anlehnung an eine baupraktische Problemstellung aus dem Stahlbetonbau gewählt. Eine aktuell verfolgte Innovation betrifft die Verwendung von Hohlkörpern als verlorene Schalung innerhalb von Deckenkonstruktionen, um eine Material- und Gewichtsersparnis zu erzielen. Aus statischer Sicht kann sich zwischen den Hohlkörpern ein Raumfachwerk in der Stahlbetonmatrix ausbilden, in das die erforderliche Bewehrung eingebaut wird. Als Hersteller sei beispielsweise die Firma COBIAX [23] genannt, auf die für weitere Details zu derartigen Deckensystemen verwiesen wird.

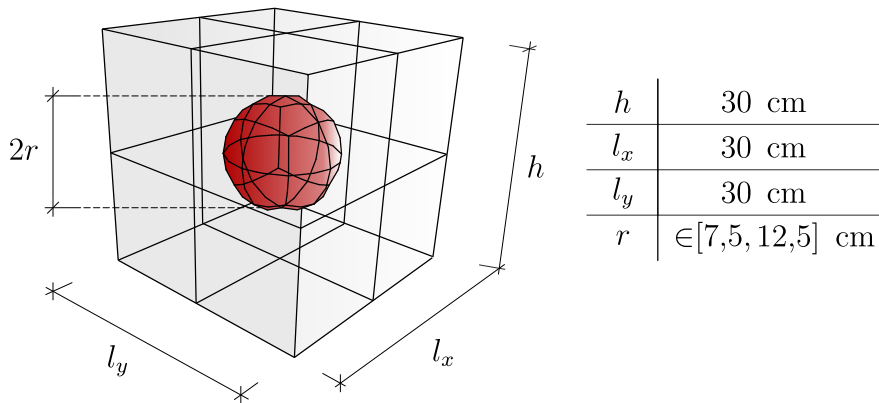


Abbildung 8.28: Geometrie des Schalen-RVEs; der Radius des Einschlusses wird als Intervallgröße gewählt

	Initialer E-Modul E_0 [kN/cm ²]	Poisson- Zahl ν_0 [-]	MAXWELL Element i	Modul- anteil $\mu_i = E_i/E_0$ [-]	Relaxations- zeit τ_i [s]
Einschluss	500	0,20	-	-	-
Matrix	3000	0,20	-	-	-
			1	0,25	10 000
			2	0,25	100 000

Tabelle 8.27: Verwendete Materialparameter des elastischen Einschluss und viskoelastischen Stahlbeton-Matrix

Zunächst wird ein geeignetes Schalen-RVE definiert, das in Abbildung 8.28 dargestellt ist. Es besitzt eine Länge l_x , Breite l_y und Höhe h von jeweils 30 cm. Die Höhe entspricht gängigen Abmessungen für Deckenplatten mit Hohlkörpern, während Länge und Breite auf der Annahme der Periodizitätseigenschaften der Einheitszelle gewählt wird. Im Zentrum des RVEs befindet sich ein weicher, elastischer Einschluss, dessen Radius auf das Intervall $r \in [7,5, 12,5]$ cm festgelegt wird. Dies entspricht typischen Abmessungen kugelförmiger Einschlüsse der genannten Deckensysteme. Der Radius r wird somit als geometrischer Parameter des RVEs behandelt und als zusätzliche Variable in den Eingangsvektor des KNN aufgenommen. Der Einschluss wird nicht als Hohlraum, sondern als im Verhältnis zur Matrix deutlich weiches, voll ausgefülltes Material modelliert, um die Eigenschaften der verlorenen Schalung zu berücksichtigen. Das RVE wird mit 104 27-Knoten Volumenelementen diskretisiert, mit periodischen Verschiebungsrandbedingungen beaufschlagt und zusätzlich durch fünf Lagrange-Parameter für die Einforderung der kinematischen Zwangsbedingungen ergänzt.

Für die Wahl der Materialparameter wird angenommen, dass die Stahlbetonmatrix viskoelastische und der weiche Einschluss rein elastische Eigenschaften aufweist. Es wird dabei ein vereinfachtes Modell zur Beschreibung des rein lastabhängigen und reversiblen Kriechverhaltens des Betons verwendet. Weitere Ausführungen zum zeitabhängigen Verformungsverhalten von Beton sind beispielsweise in KVISSEL [77] zu finden. Die gewählten Materialparameter sind in Tabelle 8.27 zusammengefasst. Durch die Wahl der Modulanteile μ_i ergeben sich für den Beton Steifigkeitsverluste von bis zu 50 % bezogen auf den initialen E-Modul $E_0 = 3000 \text{ kN/cm}^2$.

Der Radius der Kugel beeinflusst zwar die initiale Steifigkeit, jedoch nicht das zeitabhängige Verhalten der effektiven viskoelastischen Eigenschaften. Entsprechend wird die Versuchsdauer für alle Pfade zu $\bar{T} = 1 \cdot 10^6 \text{ s}$ gewählt, was etwa 11,5 Tagen entspricht. Die Zeitachse wird mit minimal 20 und maximal 80 Zeitschritten diskretisiert, woraus $\Delta \bar{t} \in [12\,500, 50\,000] \text{ s}$ resultiert. Die Membran- und Schubverzerrungen werden auf $[-0,03, 0,03]$ begrenzt. Mit $h/2 = 15 \text{ cm}$ ergibt sich für die Krümmungen ein Bereich von $[-2 \cdot 10^{-3}, 2 \cdot 10^{-3}] \text{ 1/cm}$. Die aus Gleichung (7.21) resultierenden Werte wurden um den Faktor 10 reduziert, um übermäßig große Biegemomente zu vermeiden, was zu einem stabileren KNN-Training führt. Aufgrund der richtungsunabhängigen Eigenschaften des RVEs muss keine Differenzierung der zulässigen Streckenschnittgrößen für die spannungsgesteuerten Versuche vorgenommen werden. Diese werden zu

$$\begin{aligned} |\bar{n}_{ii}| &\leq 1000 \frac{\text{kN}}{\text{cm}}, & |\bar{n}_{ij}| &\leq 400 \frac{\text{kN}}{\text{cm}}, & |\bar{m}_{ii}| &\leq 600 \frac{\text{kNcm}}{\text{cm}}, \\ |\bar{m}_{ij}| &\leq 250 \frac{\text{kNcm}}{\text{cm}}, & |\bar{q}_i| &\leq 300 \frac{\text{kN}}{\text{cm}}, & i &\neq j, \end{aligned} \quad (8.10)$$

gewählt. Dabei wurde berücksichtigt, dass sich die resultierenden Kriechdehnungen unabhängig von der Wahl des Radius im Rahmen der genannten Verzerrungsgrenzen bewegen.

Für das KNN wurde in Voruntersuchungen eine Topologie mit 80 Neuronen in zwei verdeckten Schichten bestimmt, womit die Architektur unter Einbeziehung des Radius des Einschlusses 26–80–80–8 lautet. Durch die zusätzliche Eingangsvariable sind mehr Trainingspfade erforderlich, wobei sich zehn zufällige Pfade pro Verzerrungskomponente und Versuchstyp (Kriechen, Relaxation, zyklischer Versuch) als geeignet erwiesen haben. Insgesamt wurden damit 240 Pfade erzeugt und 9199 Trainingsdaten generiert. Zusätzlich werden 1000 zufällige Datenpunkte genutzt, um die Tangenten-Nebenbedingung mit einem Penalty-Faktor von $\rho = 1$ einzufordern.

Bevor die verwendeten Validierungspfade erläutert werden, sei zunächst angemerkt, dass sich das Stoffgesetz zur Beschreibung der Matrix des RVEs zwar rein

Pfad Nr.	Versuch	Kompo- nenten	Werte	Δt [s]	T [s]	r [cm]
1	Zug	$\varepsilon_{11}, \varepsilon_{22}$	0,02, 0,02 [–]	17500	250000	7,5
2	Zug	$\varepsilon_{11}, \varepsilon_{22}$	0,02, 0,02 [–]	47500	2000000	7,5
3	Zug	κ_{11}, κ_{22}	$1 \cdot 10^{-4}, 1 \cdot 10^{-4}$ [cm ⁻¹]	17500	250000	10
4	Zug	κ_{11}, κ_{22}	$1 \cdot 10^{-4}, 1 \cdot 10^{-4}$ [cm ⁻¹]	47500	2000000	10
5	Zug	γ_1, γ_2	0,01, 0,01 [–]	17500	250000	12,5
6	Zug	γ_1, γ_2	0,01, 0,01 [–]	47500	2000000	12,5

Tabelle 8.28: Übersicht der auf dem RVE durchgeführten Versuche mit den jeweiligen maximalen Verzerrungen, Zeitparametern und Radien

deviatorisch viskoelastisch verhält, dieser Deformationszustand im Rahmen der Schalenkinematik jedoch nicht berücksichtigt werden kann. Dies steht im Widerspruch zur Annahme der inextensiblen Dickenrichtung in der makroskopischen Schalenformulierung. Entsprechend existiert aufgrund der kinematischen Annahmen kein hydrostatischer Spannungszustand, über den – wie im Beispiel 8.2.3 – geprüft werden könnte, ob das KNN die rein elastische Materialantwort abbilden kann. Dies kann jedoch auf anderem Wege erfolgen: Durch die Vorgabe kleiner Dehnraten werden die viskosen Effekte gering, wodurch sich der dissipative Anteil in den Hysteresen reduziert.

Auf Basis dieser Diskussion erfolgen die Untersuchungen am Materialpunkt mit den in Tabelle 8.28 zusammengefassten Validierungspfaden. Dabei werden ausschließlich biaxiale Zugversuche (dem Training unbekannt) unter Berücksichtigung von Be- und Entlastung bei unterschiedlichen Dehnraten und variierenden Radien durchgeführt, um die Generalisierungsfähigkeit des Modells zu überprüfen.

In Abbildung 8.29 sind die Ergebnisse des KNN im Vergleich zur Referenzlösung dargestellt. Für alle Pfade zeigt sich eine sehr gute Übereinstimmung der Modellantworten. Deutlich zu erkennen ist die verkleinerte Hysterese bei Versuchen mit geringeren Dehnraten (rechte Spalte). Damit kann bereits am Materialpunkt festgehalten werden, dass die Berücksichtigung des hier gewählten zusätzlichen Parameters mit der verwendeten Methodik grundsätzlich möglich ist.

8.3.5 Numerisches Beispiel: Hohlkörperdecke

Als letztes numerisches Beispiel wird die Approximationsfähigkeit des KNN-Materialmodells am System einer Deckenplatte untersucht. Das in Abbildung 8.30

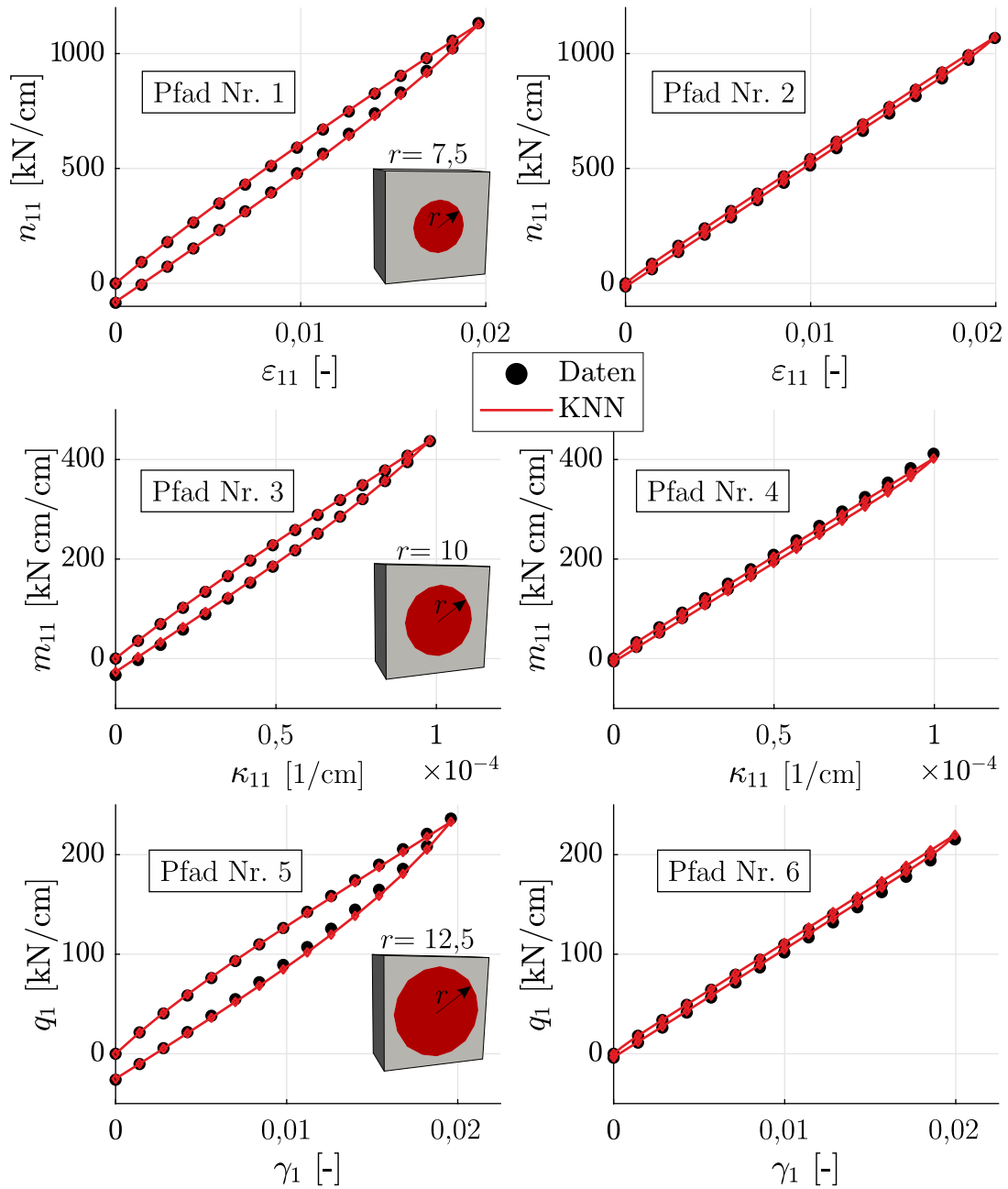


Abbildung 8.29: Materialpunktversuche mit dem KNN für verzerrungsgesteuerten Zugversuchen unter Vorgabe verschiedener Dehnraten sowie Radien des zugrundeliegenden effektiven viskoelastischen Stoffgesetzes und zugehörige Referenzlösung

dargestellte makroskopische System wird mit 10×10 (hier nicht dargestellten) Schalenelementen diskretisiert. Es wird ein Viertel des Gesamtsystems mit jeweiliger Länge und Breite $2L = 900$ cm sowie einer Höhe $h = 30$ cm unter Vorgabe von Symmetrie-Randbedingungen modelliert. Auf den verbleibenden Rändern wird eine Navier-Lagerung gewählt. An jedem makroskopischen GAUSS-Punkt ist

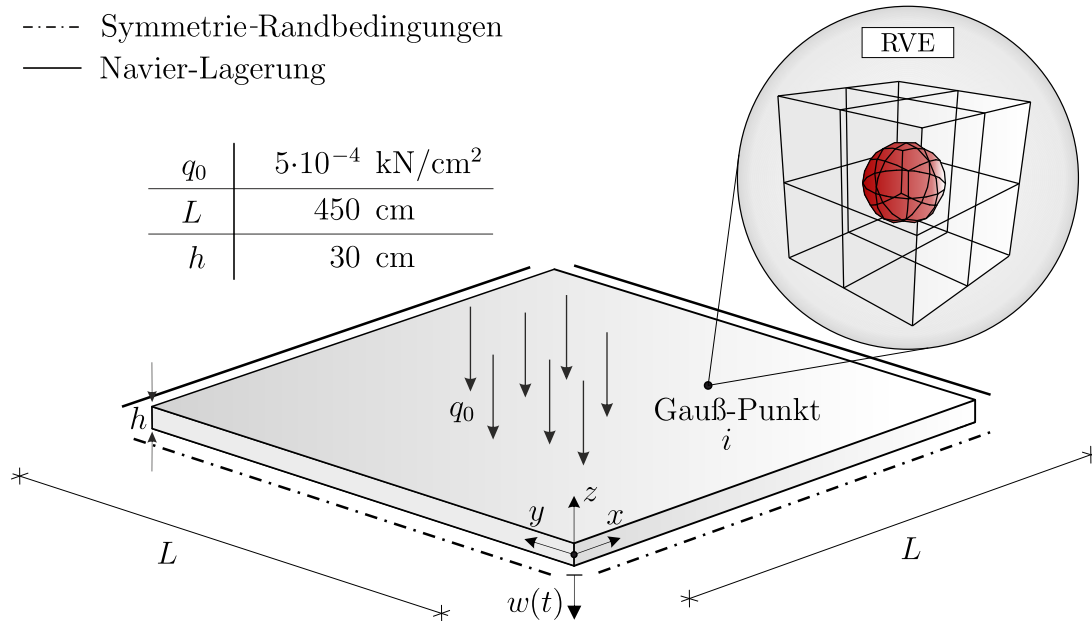


Abbildung 8.30: Geometrie, Randbedingungen sowie Belastung des Viertel-Systems des FE²-Modells der Schale mit zugehörigem RVE; RVE-Details vgl. Abbildung 8.28

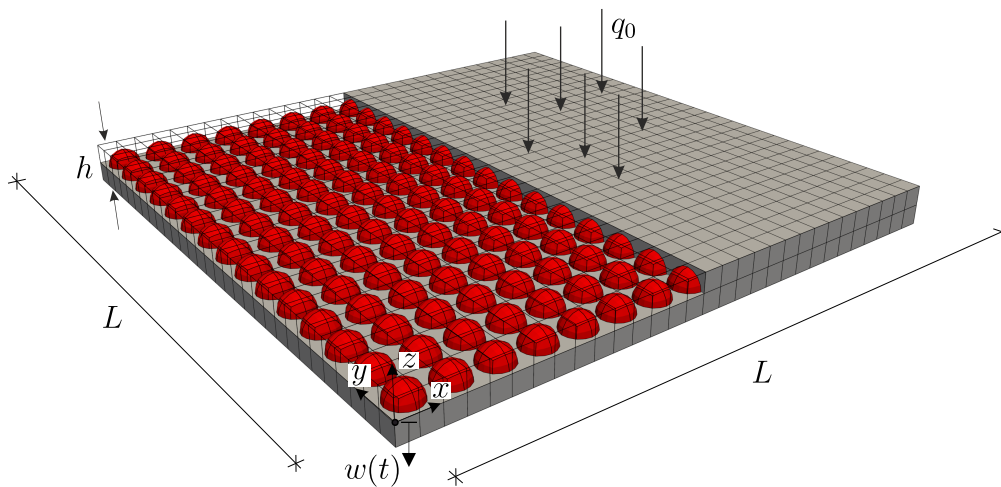


Abbildung 8.31: Geometrie sowie Belastung des Viertel-Systems des 3D-Modells; Randbedingungen und Abmessungen analog zu Abbildung 8.30, Mikrostruktur-Details vgl. Abbildung 8.28. Der Radius der Kugeln beträgt hier $r = 12,5 \text{ cm}$

das zuvor eingeführte RVE definiert.

Der Lastfall besteht aus einer zeitlich und räumlich konstanten Streckenlast $q_0 = 5 \cdot 10^{-4} \text{ kN/cm}^2 (= 5 \text{ kN/m}^2)$. Zusätzlich wird eine Vergleichslösung auf Basis eines vollaufgelösten 3D-Modells berücksichtigt, dargestellt in Abbildung 8.31. Auch hier wird das Viertel-System modelliert, das aus 15×15 Schalen-RVEs zusam-

Modell	Durchbiegung $w(t=1 \cdot 10^5 \text{ s})$ [cm]		
	$r = 7,5$	$r = 10$	$r = 12,5$
KNN	$3,86 \cdot 10^{-1}$	$3,99 \cdot 10^{-1}$	$4,37 \cdot 10^{-1}$
FE ²	$3,94 \cdot 10^{-1}$	$4,05 \cdot 10^{-1}$	$4,31 \cdot 10^{-1}$
3D	$3,89 \cdot 10^{-1}$	$4,00 \cdot 10^{-1}$	$4,33 \cdot 10^{-1}$

Tabelle 8.29: Vergleich der Durchbiegung $w(t=1 \cdot 10^5 \text{ s})$ am Ende des Kriechversuchs für verschiedene Modelle und Radien r

mengesetzt ist. Für die 3D-Lösung wird die bereits beschriebene Diskretisierung des RVEs mit 104 27-Knoten Volumenelementen (ohne Nebenbedingungen) verwendet. Die Navier-Lagerung der Platte erfolgt auf den Mittenknoten bei $z = 0$, während die Belastung auf die Knoten der Oberfläche $z = h/2$ aufgebracht wird. Dargestellt ist exemplarisch das System für einen Radius der Kugeln von $r = 12,5 \text{ cm}$.

Im Folgenden wird das zeitabhängige Verhalten der Platte unter der Flächenlast, die beispielsweise das Eigengewicht der Konstruktion repräsentieren kann, untersucht. Von Interesse ist dabei die Durchbiegung in der Mitte $w(t)$, wobei bei der 3D-Lösung der Mittenknoten bei $x = y = z = 0$ betrachtet wird. Die Belastung wird innerhalb eines Zeitschrittes $\Delta t = 20\,000 \text{ s} (\approx 5,5 \text{ h})$ aufgebracht und anschließend innerhalb von 25 Zeitschritten mit gleicher Zeitschrittweite bis $T = 5 \cdot 10^5 \text{ s} (\approx 11,6 \text{ Tage})$ konstant gehalten. Es werden drei Kriechkurven für Kugelradien $r \in \{7,5; 10; 12,5\} \text{ cm}$ mit der Schalen-FE²-, der 3D- und der Schalen-KNN-Lösung berechnet. Dabei sei angemerkt, dass für das FE²-Modell jeweils ein neues RVE definiert und für das 3D-Modell das gesamte System neu vernetzt werden muss, während beim Schalen-KNN-Modell lediglich der Radius im Materialmodell angepasst wird. Zwar ist auch bei den beiden erstgenannten Ansätzen der Radius lediglich ein variabler Eingabeparameter, bei komplexen 3D-Geometrien kann jedoch bereits die Netzerzeugung einen nicht zu vernachlässigenden Rechenaufwand darstellen.

Die Kriechkurven der drei Modelle für die unterschiedlichen Radien sind in Abbildung 8.32 dargestellt. Zusätzlich ist eine zwanzigfach überhöhte Verformungsfigur des 3D-Modells gezeigt. Qualitativ ergibt sich ein erwartungsgemäßes Verhalten der Struktur, unabhängig vom Modell: Mit zunehmendem Radius und damit geringerer Steifigkeit nehmen die Verformungen zu. Modellabhängig ergeben sich Unterschiede in den Durchbiegungen von etwa 10 % bis 15 %. Insgesamt zeigt sich eine sehr gute Übereinstimmung der Ergebnisse der drei Modelle. Für $r = 7,5 \text{ cm}$ überschätzt das KNN die Verformungen zwischenzeitlich leicht, nähert sich jedoch

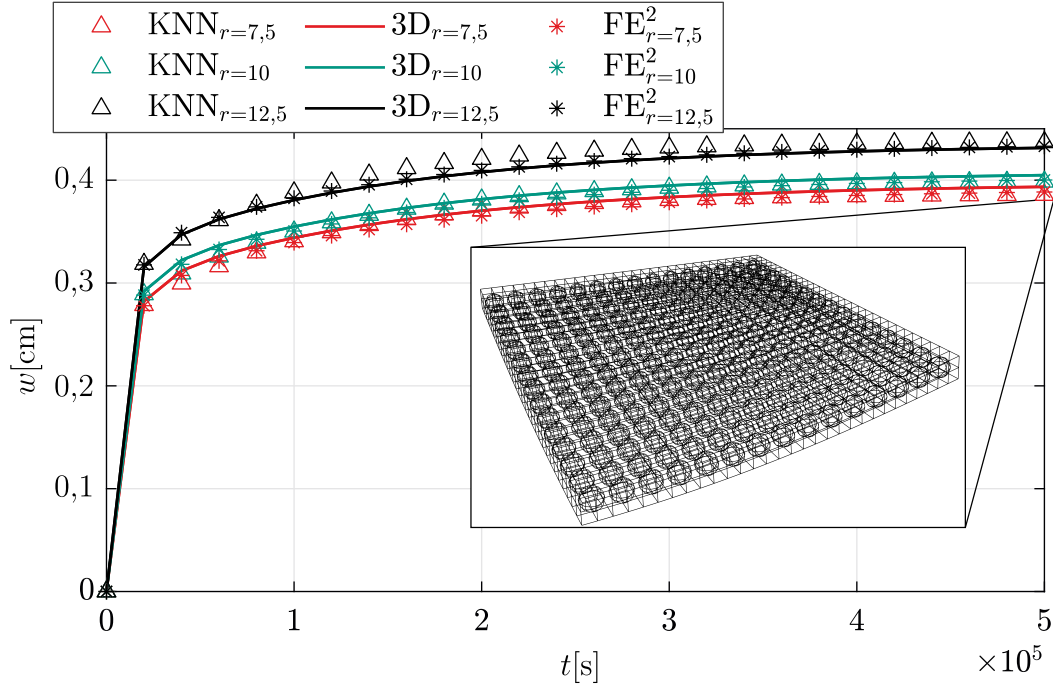


Abbildung 8.32: Ergebnisse der drei verwendeten Modelle für die Durchbiegung $w(t)$ bei konstanter Belastung und drei Radien der Kugeln sowie zwanzigfach überhöhte Verformungsfigur des 3D-Modells für $r = 7,5$ cm und $t = 5 \times 10^5$ s

im zeitlichen Verlauf wieder den anderen Lösungen an. Es ist hervorzuheben, dass die verwendeten Radien jeweils die obere und untere Grenze des Trainingsraums des KNN darstellen und dennoch gute Approximationen erzielt werden. Somit kann festgehalten werden, dass das KNN auch unter Berücksichtigung des zusätzlichen Parameters zuverlässige Ergebnisse liefert. Die absoluten Werte der berechneten Verformungen zum Zeitpunkt $t = 5 \cdot 10^5$ s sind in Tabelle 8.29 zusammengefasst.

Zwei abschließende Aspekte sind bei diesem Beispiel hervorzuheben: Obwohl das System vergleichsweise komplex erscheint, bleiben die Rechenzeiten für alle Modelle in einem akzeptablen Rahmen. Normiert auf das KNN benötigt das 3D-Modell die etwa 10-fache und die FE^2 -Lösung die etwa 35-fache Zeit pro Zeitschritt. Durch die erhöhte Anzahl an benötigten Trainingsdaten steigt entsprechend auch die Zeit für das KNN-Training. Das volle Potenzial der vorgeschlagenen Methode entfaltet sich vor allem bei komplexeren Mikrostrukturen mit mehr Freiheitsgraden oder bei großen Verformungen, bei denen im Rahmen einer geometrisch nichtlinearen Theorie auch für die 3D- und FE^2 -Modelle eine NEWTON-Iteration erforderlich wird.

Weiterhin geht der hier verwendete zusätzliche Parameter vermutlich nur mit mäßiger Nichtlinearität in das Problem ein. Die Grenzen des KNN dürften mit den

vorgestellten Methoden erreicht werden, wenn beispielsweise unterschiedliche Relaxationszeiten auftreten, wodurch tiefere Überlegungen zur Wahl der Zeitschrittweiten und Versuchsdauern notwendig würden. Für rein geometrische Parameter – etwa Volumengehalte einzelner Phasen – ist das KNN mit der vorgestellten Methodik hingegen zweifelsfrei einsetzbar, was die sehr zufriedenstellenden Ergebnisse sowohl auf Materialpunkt- als auch auf Strukturebene eindrucksvoll zeigen.

9 Zusammenfassung und Ausblick

In der vorliegenden Arbeit wurden KNN-basierte Materialmodelle für Schalen mit viskoelastischer Mikrostruktur auf Basis einer effizienten Strategie zur Datengenerierung entwickelt. Mit Rückblick auf die eingangs formulierten Aspekte werden im Folgenden die wesentlichen Erkenntnisse der Arbeit zusammenfassend dargelegt:

- i) Es zeigt sich, dass auf Grundlage gängiger eindimensionaler viskoelastischer experimenteller Versuche eine Datenbasis erzeugt werden kann, mit der zuverlässige Trainingserfolge der verwendeten KNN erzielt werden können. Dies hat den Vorteil, dass grundsätzlich auch auf Basis vorhandener experimenteller Daten, die zur Kalibrierung analytischer Materialmodelle dienen, KNN-basierte Materialmodelle aufgebaut werden können. Insbesondere für eindimensionale Materialmodelle konnte gezeigt werden, dass Relaxations-, Kriech- und zyklische Versuche den Trainingsraum deutlich besser abdecken als zufällige Verzerrungspfade. Durch eine adaptive Diskretisierung der Zeitachse können vor allem bei geringer Datenbasis bessere Approximationen im Vergleich zu einer konstanten zeitlichen Diskretisierung erzielt werden.
- ii) Auch für den mehrdimensionalen Fall können die zuvor genannten Versuchsarten genutzt werden, um das KNN-Materialmodell zu konstruieren. Dabei zeigt sich, dass uniaxiale Versuche unter Kenntnis der Querdehnung in bestimmten Fällen bereits zur Approximation beliebiger mehrachsialer Spannungs-Verzerrungs-Zustände ausreichend sein können. Somit könnte auch im mehrdimensionalen Fall die Konstruktion des KNN-Modells auf Basis uniaxialer, experimentell üblicher Versuche erfolgen. Der Trainingserfolg ist jedoch nicht in jedem Fall garantiert.
- iii) Bei der Verwendung synthetischer Daten, z. B. auf Basis von RVEs, stellt die Materialtangente einen entscheidenden Informationszuwachs dar, durch den zuverlässig gute Generalisierungsfähigkeiten des KNN erzielt werden. Darüber hinaus kann die benötigte Menge an Primärdaten (reinen Spannungs-Verzerrungs-Daten) deutlich reduziert werden. Durch die Möglichkeit, die Nebenbedingungen im normalisierten Trainingsraum einzufordern, wird a priori eine gleiche Gewichtung von Trainings- und Nebenbedingungsfehler gewährleistet, wodurch keine explizite Wahl des Penalty-Parameters erforderlich ist. Es zeigt sich ferner, dass sich mit zunehmender Anzahl an Nebenbedingungsdaten die Approximationsgüte des KNN verbessert, während eine übermäßige Erhöhung dieser Datenmenge keinen negativen Einfluss auf

das Training hat. Mit der Tangenteninformation wird auch die rein elastische, hydrostatische Achse erlernt, die nicht explizit in den verwendeten Trainingspfaden enthalten ist.

- iv) Neben den Untersuchungen auf Materialpunktebene wird gezeigt, dass auch beim Einsatz in anspruchsvollen FE-Berechnungen gute Approximationen und ein stabiles numerisches Verhalten erzielt werden. Für den eindimensionalen Fall wird dies am Beispiel eines parametrisierbaren Materialmodells für viskoelastische Fasern weicher Aktuatoren demonstriert, im dreidimensionalen Fall an einer geschlitzten Platte. Dabei wird außerdem gezeigt, dass in Abwesenheit äußerer Belastung kein unphysikalisches Verhalten des KNN auftritt.
- v) Für die konstitutive Beschreibung von Schalen ergibt sich aufgrund der acht arbeitskonformen Spannungs- und Verzerrungskomponenten – unter Berücksichtigung der Raten- und Geschichtsabhängigkeit durch weitere Variablen im Eingangsvektor des KNN – ein hochdimensionaler Funktionsraum, der durch das KNN approximiert werden muss. Der sogenannte „Fluch der Dimension“ zeigt sich hierbei trotz 25 Eingangs- und 8 Ausgangsvariablen nicht. Die im ein- und dreidimensionalen Fall angewandten Methoden zur Konstruktion des KNN-Materialmodells lassen sich ohne Einschränkungen übertragen. An einfachen numerischen Beispielen wurde die Brauchbarkeit erfolgreich demonstriert. Bei komplexen Mikrostrukturen, deren vollständig aufgelöste makroskopische Systeme mehrere Millionen Freiheitsgrade umfassen können, entfaltet das KNN als Ersatzmodell für FE^2 -Lösungen und als Alternative zu 3D-Modellen sein volles Potenzial. Dabei werden unter Verwendung des KNN signifikante Reduktionen in der Berechnungszeit erreicht.
- vi) Der zuvor genannte – bereits hochdimensionale – zu approximierende Funktionsraum kann durch geometrische Merkmale, wie beispielsweise den Radius von Einschlüssen, erweitert werden. Damit ist es möglich, mit einem KNN die Eigenschaften von RVEs mit unterschiedlichen Geometrien abzubilden. Zukünftig können so im Designprozess von Bauteilen die Einflüsse zusätzlicher Parameter innerhalb eines Modells erfasst und effizient analysiert werden.

Die erzielten Ergebnisse stellen einen vielversprechenden Ansatz für den Einsatz von KNN zur Approximation des Mikro-Makro-Übergangs – nicht nur viskoelastischen, sondern auch beliebigen Materialverhaltens – dar. Im Rahmen der Untersuchungen zeigen sich jedoch auch Einschränkungen der hier verwendeten Methodik, die kritisch eingeordnet werden müssen und zugleich einen Ausblick

auf weiterführende Arbeiten liefern. Im Wesentlichen sind dabei die folgenden Punkte zu nennen:

- i) Die verwendete KNN-Architektur eines Feedforward-Netzes garantiert per Konstruktion keine thermodynamische Konsistenz. Physikalische Grundprinzipien wie die Normalisierung, die Unabhängigkeit gegenüber der Zeitschrittweite im Ruhezustand sowie die Einhaltung der Dissipationsungleichung können strukturell nicht gewährleistet werden. Es zeigt sich lediglich, dass bei gut trainierten KNN im Rahmen üblicher Anwendungen die Fehler toleranzen klein sind. In dieser Arbeit stand nicht das Ziel im Vordergrund, die FE^2 -Methode zu ersetzen, sondern diese insbesondere in frühen Phasen des Designprozesses durch eine effiziente Approximation des effektiven Mikroverhaltens zu ergänzen. Die Verwendung strukturell thermodynamisch konsistenter Architekturen, wie sie beispielsweise jüngst von ROSENKRANZ ET AL. [114] vorgeschlagen wurden, stellt einen naheliegenden nächsten Schritt dar, um zukünftig möglicherweise eine Alternative zu klassischen Mehrskalenansätzen zu bieten.
- ii) Für die Wahl der Grenzen der Zeitschrittweiten wird in dieser Arbeit ein ingenieurmäßiger Ansatz verfolgt. Die Zeitachse wird dabei mit einer selbst gewählten minimalen und maximalen Anzahl von Zeitpunkten diskretisiert. Aufgrund der fehlenden Extrapolationsfähigkeit des KNN ist die Anwendung des Modells auf diesen Bereich beschränkt, was einen signifikanten Nachteil gegenüber klassischen analytischen Modellen darstellt. Eine Eliminierung der Zeitschrittweite aus dem Eingangsvektor – beispielsweise über implizite Verfahren – kann hierbei Abhilfe schaffen, erfordert jedoch die Einführung eines NEWTON-Verfahrens am Materialpunkt, was mit erhöhtem Rechenaufwand verbunden ist.
- iii) Bei der Verwendung der FE^2 -Methode liegen die Informationen der Mikroebene an jedem makroskopischen GAUSS-Punkt vor. Dadurch können beispielsweise auch lokale Spannungen ausgewertet werden, die im Designprozess möglicher Mikrostrukturen von Bedeutung sind. Bei der Verwendung des KNN gehen diese lokalen Informationen verloren. Ein Ansatz bestünde darin, bereits während der Datengenerierung für die verwendeten Spannungs-Verzerrungspfade die Materialverträglichkeit zu prüfen. Außerhalb der generierten Pfade steht diese Information jedoch nicht zur Verfügung.

In Hinblick auf zukünftige Anwendungen, unter Berücksichtigung der zuvor genannten Verbesserungspotenziale, eröffnet sich ein weites Feld weiterer Fragestellungen. Exemplarisch seien hier die Erweiterung auf temperaturabhängiges und

nichtlineares Materialverhalten – wie es häufig bei Polymeren auftritt – oder die Anwendung auf Strukturen genannt, bei denen auf der Mikroebene geometrische Instabilitäten entstehen, die auf der Makroebene materielle Instabilitäten am Materialpunkt induzieren. Auch multiphysikalische Phänomene stellen eine vielversprechende Perspektive dar. Neben Schalen können dabei künftig auch Balkenelemente (vgl. z. B. KLARMANN ET AL. [67]) auf der Makroebene eingesetzt werden, für die erste KNN-Ansätze derzeit von SCHOMMARTZ [119] entwickelt werden. Darüber hinaus rücken mit den von LIU ET AL. [85] eingeführten *Kolmogorov–Arnold*-Netzen neuartige KI-Architekturen in den Fokus, die eine stärkere physikalische Interpretierbarkeit ermöglichen.

Aus Sicht des Autors bestätigen sich zwei Punkte der eingangs geführten Diskussion: Auch in der Baustatik kann der Einsatz von KI dem Ingenieur als zusätzliches Werkzeug in der Tragwerksplanung dienen, um die Grenzen des Machbaren künftig weiter auszureizen. Der Weg bis zur praktischen Anwendung wird jedoch – wie so oft – noch einige Zeit in Anspruch nehmen. Weiterhin sind die im Rahmen der Arbeit erzielten zufriedenstellenden Ergebnisse nicht zuletzt auf die ausführlich beschriebenen theoretischen Grundlagen zurückzuführen, die für die Entwicklung der Methodik erforderlich waren. Auch hier bestätigt sich, dass das volle Potenzial der KI nur ausgeschöpft werden kann, wenn das notwendige Hintergrundwissen vorhanden ist. Der Prompt „Konstruiere mir ein Materialmodell für viskoelastische Schalenstrukturen“ wird daher auch künftig für vergleichbare Problemstellungen wohl nur begrenzten Erfolg haben.

Literatur

- [1] ABDOLAZIZI, K.P. / AYDIN, R.C. / CYRON, C.J. / LINKA, K.: *Constitutive Kolmogorov–Arnold Networks (CKANs): Combining accuracy and interpretability in data-driven material modeling*. Journal of the Mechanics and Physics of Solids 203 (2025), S. 106212.
- [2] ABDOLAZIZI, K.P. / LINKA, K. / CYRON, C.J.: *Viscoelastic constitutive artificial neural networks (vCANNs) – A framework for data-driven anisotropic nonlinear finite viscoelasticity*. Journal of Computational Physics 499 (2024), S. 112704.
- [3] ABOUDI, J.: *Mechanics of composite materials*. Nummer 29 in der Reihe: Studies in applied mechanics Elsevier, Amsterdam, 1991.
- [4] AL-HAIK, M.S. / HUSSAINI, M.Y. / GARMESTANI, H.: *Prediction of nonlinear viscoelastic behavior of polymeric composites using an artificial neural network*. International Journal of Plasticity 22(7) (2006), S. 1367–1392.
- [5] ALDAKHEEL, F. / ELSAYED, E.S. / ZOHDİ, T.I. / WRIGGERS, P.: *Efficient multiscale modeling of heterogeneous materials using deep neural networks*. Computational Mechanics 72(1) (2023), S. 155–171.
- [6] AMOS, B. / XU, L. / KOLTER, J. Z.: *Input Convex Neural Networks*. In: PRECUP, DOINA / TEH, YEE WHY, (HRSG.), *Proceedings of the 34th International Conference on Machine Learning*, Band 70, in der Reihe: Proceedings of Machine Learning Research, S. 146–155 PMLR, 06–11 Aug 2017.
- [7] ARGYRIS, J.H.: *The computer shapes the theory*. Journal of the Royal Aeronautical Society (1965), S. XXXII.
- [8] ARGYRIS, J.H. / PISTER, K.S. / SZIMMAT, J. / WILLAM, K.J.: *Unified concepts of constitutive modelling and numerical solution methods for concrete creep problems*. Computer Methods in Applied Mechanics and Engineering 10(2) (1977), S. 199–246.
- [9] AS’AD, F. / FARHAT, C.: *A mechanics-informed deep learning framework for data-driven nonlinear viscoelasticity*. Computer Methods in Applied Mechanics and Engineering 417 (2023), S. 116463.
- [10] BABUŠKA, I. / SURI, M.: *The p- and h-p versions of the finite element method, an overview*. Computer Methods in Applied Mechanics and Engineering 80(1–3) (1990), S. 5–26.
- [11] BASAR, Y. / MONTAG, U. / DING, Y.: *On an isoparametric finite-element for composite laminates with finite rotations*. Computational Mechanics 12(6) (1993), S. 329–348.

- [12] BATHE, K.-J. / DVORKIN, E.N.: *A four-node plate bending element based on Mindlin/Reissner plate theory and a mixed interpolation*. International Journal for Numerical Methods in Engineering 21(2) (1985), S. 367–383.
- [13] BATTITI, R.: *First- and Second-Order Methods for Learning: Between Steepest Descent and Newton's Method*. Neural Computation 4(2) (1992), S. 141–166.
- [14] BENADY, A. / BARANGER, E. / CHAMOIN, L.: *Unsupervised learning of history-dependent constitutive material laws with thermodynamically-consistent neural networks in the modified Constitutive Relation Error framework*. Computer Methods in Applied Mechanics and Engineering 425 (2024), S. 116967.
- [15] BETSCH, P. / STEIN, E.: *An assumed strain approach avoiding artificial thickness straining for a non-linear 4-node shell element*. Communications in Numerical Methods in Engineering 11(11) (1995), S. 899–909.
- [16] BISHARA, D. / XIE, Y. / LIU, W.K. / LI, S.: *A State-of-the-Art Review on Machine Learning-Based Multiscale Modeling, Simulation, Homogenization and Design of Materials*. Archives of Computational Methods in Engineering 30(1) (2022), S. 191–222.
- [17] BISHOP, C.M.: *Curvature-driven smoothing: a learning algorithm for feedforward networks*. IEEE Transactions on Neural Networks 4(5) (1993), S. 882–884.
- [18] BOZKURT, M.O. / TAGARIELLI, V.L.: *Data-driven homogenisation of viscoelastic porous elastomers: Feedforward versus knowledge-based neural networks*. International Journal of Mechanical Sciences 286 (2025), S. 109824.
- [19] CAUCHY, A.: *Allgemeine Methode zum Lösen von Gleichungssystemen*. Comp. Returns. Sci. Paris 25 (1847), S. 536–538.
- [20] CHEN, G.: *Recurrent neural networks (RNNs) learn the constitutive law of viscoelasticity*. Computational Mechanics 67(3) (2021), S. 1009–1019.
- [21] CHRISTENSEN, R.M.: *Theory of viscoelasticity*. Academic Press, New York, 2. Auflage, 1982.
- [22] CLOUGH, R.W.: *Early history of the finite element method from the view point of a pioneer*. International Journal for Numerical Methods in Engineering 60(1) (2004), S. 283–287.
- [23] COBIAX DEUTSCHLAND GMBH: *How to Cobiax – Der Quick Guide zu Cobiax SL*. <https://www.cobiax.com/intl/de/downloads/> (2023) (Zugriff am 10. Oktober 2025).
- [24] COENEN, E.W.C. / KOUZNETSOVA, V.G. / GEERS, M.G.D.: *Computational homogenization for heterogeneous thin sheets*. International Journal for Numerical Methods in Engineering 83(8–9) (2010), S. 1180–1205.

-
- [25] CONG, Y. / NEZAMABADI, S. / ZAHROUNI, H. / YVONNET, J.: *Multiscale computational homogenization of heterogeneous shells at small strains with extensions to finite displacements and buckling*. International Journal for Numerical Methods in Engineering 104(4) (2015), S. 235–259.
- [26] CZARNECKI, W.M. / OSINDERO, S. / JADERBERG, M. / SWIRSZCZ, G. / PASCANU, R.: *Sobolev Training for Neural Networks*. In: *Advances in Neural Information Processing Systems*, 2017.
- [27] DA SILVA, I.N.: *Artificial Neural Networks: A practical course*. Springer International Publishing AG Switzerland, 2017.
- [28] DORNHEIM, J. / MORAND, L. / NALLANI, H.J. / HELM, D.: *Neural Networks for Constitutive Modeling: From Universal Function Approximators to Advanced Models and the Integration of Physics*. Archives of Computational Methods in Engineering 31(2) (2023), S. 1097–1127.
- [29] DUCHI, J. / HAZAN, E. / SINGER, Y.: *Adaptive Subgradient Methods for Online Learning and Stochastic Optimization*. Journal of Machine Learning Research 12 (2011), S. 2121–2159.
- [30] FERRY, J.D.: *Viscoelastic properties of polymers*. Wiley, New York, 2. Auflage, 1970.
- [31] FEYEL, F. / CHABOCHE, J.-L.: *FE2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre SiC/Ti composite materials*. Computer Methods in Applied Mechanics and Engineering 183(3–4) (2000), S. 309–330.
- [32] GEERS, M.G.D. / COENEN, E.W.C. / KOUZNETSOVA, V.G.: *Multi-scale computational homogenization of structured thin sheets*. Modelling and Simulation in Materials Science and Engineering 15(4) (2007), S. S393–S404.
- [33] GEIGER, C.: *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer, Berlin, Heidelberg, 2002.
- [34] GEIGER, J. / WAGNER, W. / FREITAG, S.: *Effective Material Modeling of Parameterizable Viscoelastic Shell Structures with Artificial Neural Networks*. PAMM 25(3) (2025), S. e70015.
- [35] GEIGER, J. / WAGNER, W. / FREITAG, S.: *Multiscale modeling of viscoelastic shell structures with artificial neural networks*. Computational Mechanics 76(3) (2025), S. 541–565.
- [36] GHABOUSSI, J. / GARRETT, J.H. / WU, X.: *Knowledge-Based Modeling of Material Behavior with Neural Networks*. Journal of Engineering Mechanics 117(1) (1991), S. 132–153.
- [37] GHABOUSSI, J. / JOGHATAIE, A.: *Active Control of Structures Using Neural Networks*. Journal of Engineering Mechanics 121(4) (1995), S. 555–567.

-
- [38] GHABOUSSI, J. / PECKNOLD, D.A. / ZHANG, M. / HAJ-ALI, R.M.: *Autoprogessive training of neural network constitutive models*. International Journal for Numerical Methods in Engineering 42(1) (1998), S. 105–126.
- [39] GHABOUSSI, J. / SIDARTA, D.E.: *New nested adaptive neural networks (NANN) for constitutive modeling*. Computers and Geotechnics 22(1) (1998), S. 29–52.
- [40] GHAVAMIAN, F. / SIMONE, A.: *Accelerating multiscale finite element simulations of history-dependent materials using a recurrent neural network*. Computer Methods in Applied Mechanics and Engineering 357 (2019), S. 112594.
- [41] GOETTLER, L.A. / SHEN, K.S.: *Short Fiber Reinforced Elastomers*. Rubber Chemistry and Technology 56(3) (1983), S. 619–638.
- [42] GOODFELLOW, I.: *Deep learning*. Adaptive computation and machine learning The MIT Press, Cambridge, Massachusetts, 2016.
- [43] GROSS, D.: *Bruchmechanik*. Springer, Berlin, Heidelberg, 6. Auflage, 2016.
- [44] GRUTTMANN, F. / WAGNER, W.: *Structural analysis of composite laminates using a mixed hybrid shell element*. Computational Mechanics 37(6) (2005), S. 479–497.
- [45] GRUTTMANN, F. / WAGNER, W.: *A coupled two-scale shell model with applications to layered structures*. International Journal for Numerical Methods in Engineering 94(13) (2013), S. 1233–1254.
- [46] GRUTTMANN, F. / WAGNER, W.: *A FE2 shell model with periodic boundary conditions for thin and thick shells*. International Journal for Numerical Methods in Engineering 125(11) (2024).
- [47] GRUTTMANN, F. / WAGNER, W.: *A nonlinear 4-node shell element with one point quadrature and stabilization based on a Hu–Washizu variational formulation*. Computational Mechanics 76(3) (2025), S. 613–633.
- [48] HAGAN, M.T. / MENHAJ, M.B.: *Training feedforward networks with the Marquardt algorithm*. IEEE Transactions on Neural Networks 5(6) (1994), S. 989–993.
- [49] HAMBLI, R.: *Application of Neural Networks and Finite Element Computation for Multiscale Simulation of Bone Remodeling*. Journal of Biomechanical Engineering 132(11) (2010).
- [50] HASHASH, Y.M.A. / JUNG, S. / GHABOUSSI, J.: *Numerical implementation of a neural network based material model in finite element analysis*. International Journal for Numerical Methods in Engineering 59(7) (2004), S. 989–1005.
- [51] HASHIN, Z.: *Analysis of Composite Materials—A Survey*. Journal of Applied Mechanics 50(3) (1983), S. 481–505.

-
- [52] HEBB, D.O.: *The Organization of Behavior: A Neuropsychological Theory*. Wiley, New York, 1949.
- [53] HERRMANN, L. / KOLLMANNSSBERGER, S.: *Deep learning in computational mechanics: a review*. Computational Mechanics 74(2) (2024), S. 281–331.
- [54] HILL, R.: *Elastic properties of reinforced solids: Some theoretical principles*. Journal of the Mechanics and Physics of Solids 11(5) (1963), S. 357–372.
- [55] HOLZAPFEL, G.A.: *Nonlinear solid mechanics: a continuum approach for engineering*. Wiley, Chichester, 2010.
- [56] HOLZAPFEL, G.A. / GASSER, T.C. / STADLER, M.: *A structural model for the viscoelastic behavior of arterial walls: Continuum formulation and finite element analysis*. European Journal of Mechanics - A/Solids 21(3) (2002), S. 441–463.
- [57] HOPFIELD, J.J.: *Neural networks and physical systems with emergent collective computational abilities*. Proceedings of the National Academy of Sciences 79(8) (1982), S. 2554–2558.
- [58] HSU, F.-H.: *Behind Deep Blue*. Princeton University Press, Princeton, 2022.
- [59] HUANG, S. / HE, Z. / CHEM, B. / REINA, C.: *Variational Onsager Neural Networks (VONNs): A thermodynamics-based variational learning strategy for non-equilibrium PDEs*. Journal of the Mechanics and Physics of Solids 163 (2022), S. 104856.
- [60] HUGHES, T.J.R.: *The finite element method*. Prentice-Hall Inc, Englewood Cliffs, New Jersey, 1987.
- [61] HUGHES, T.J.R. / COTTRELL, J.A. / BAZILEVS, Y.: *Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement*. Computer Methods in Applied Mechanics and Engineering 194(39–41) (2005), S. 4135–4195.
- [62] JUNG, S. / GHABOUSSI, J.: *Neural network constitutive model for rate-dependent materials*. Computers & Structures 84(15–16) (2006), S. 955–963.
- [63] KALINA, K.A. / BRUMMUND, J. / SUN, W. / KÄSTNER, M.: *Neural networks meet anisotropic hyperelasticity: A framework based on generalized structure tensors and isotropic tensor functions*. Computer Methods in Applied Mechanics and Engineering 437 (2025), S. 117725.
- [64] KALINA, K.A. / LINDEN, L. / BRUMMUND, J. / KÄSTNER, M.: *FE^{ANN}: an efficient data-driven multiscale approach based on physics-constrained neural networks and automated data mining*. Computational Mechanics 71(5) (2023), S. 827–851.
- [65] KERWIN, E.M.: *Damping of Flexural Waves by a Constrained Viscoelastic Layer*. The Journal of the Acoustical Society of America 31(7) (1959), S. 952–962.

-
- [66] KINGMA, D.P. / BA, J.: *Adam: A Method for Stochastic Optimization*. 2014.
- [67] KLARMANN, S. / GRUTTMANN, F. / KLINKEL, S.: *Homogenization assumptions for coupled multiscale analysis of structural elements: beam kinematics*. Computational Mechanics 65(3) (2019), S. 635–661.
- [68] KLEIN, D.K. / ORTIGOSA, R. / MARTÍNEZ-FRUTOS, J. / WEEGER, O.: *Finite electro-elasticity with physics-augmented neural networks*. Computer Methods in Applied Mechanics and Engineering 400 (2022), S. 115501.
- [69] KLINKEL, S.: *Theorie und Numerik eines Volumen-Schalen-Elementes bei finiten elastischen und plastischen Verzerrungen*. Dissertation, Institut für Baustatik Universität Karlsruhe (TH), 2000.
- [70] KLINKEL, S. / GRUTTMANN, F. / WAGNER, W.: *A robust non-linear solid shell element based on a mixed variational formulation*. Computer Methods in Applied Mechanics and Engineering 195(1–3) (2006), S. 179–201.
- [71] KLINKEL, S. / GRUTTMANN, F. / WAGNER, W.: *A mixed shell formulation accounting for thickness strains and finite strain 3d material models*. International Journal for Numerical Methods in Engineering 74(6) (2007), S. 945–970.
- [72] KRAWIETZ, A.: *Materialtheorie*. Springer, Berlin, 1986.
- [73] KRIESEL, D.: *Ein kleiner Überblick über Neuronale Netze*. 2007.
- [74] KUHN, H.W. / TUCKER, A.W.: *Nonlinear Programming*. S. 247–258, Springer Basel, 2013.
- [75] KUMAR, S. / TAN, S. / ZHENG, L. / KOCHMANN, D.M.: *Inverse-designed spinodoid metamaterials*. Computational Materials 6(73) (2020).
- [76] KUNTSCHKE, JOHANNES K.: *Mechanisches Verhalten von Verbundglas unter zeitabhängiger Belastung und Explosionsbeanspruchung*. Nummer 46 in der Reihe: Mechanik, Werkstoffe und Konstruktion im Bauwesen, Springer Berlin Heidelberg, 2015.
- [77] KVITSEL, V.: *Zur Vorhersage des Schwindens und Kriechens von normal- und hochfestem Konstruktionsleichtbeton mit Blähtongesteinskörnung*. Dissertation, Institut für Massivbau und Baustofftechnologie, Karlsruher Institut für Technologie, 2011.
- [78] LE, B.A. / YVONNET, J. / HE, Q.-C.: *Computational homogenization of nonlinear elastic materials using neural networks*. International Journal for Numerical Methods in Engineering 104(12) (2015), S. 1061–1084.
- [79] LECUN, Y.A. / BOTTOU, L. AND ORR, G.B. / MÜLLER, K.-R.: *Efficient BackProp*. S. 9–48, Springer, Berlin, Heidelberg, 2012.

-
- [80] LEFIK, M. / BOSO, D.P. / SCHREFLER, B.A.: *Artificial Neural Networks in numerical modelling of composites*. Computer Methods in Applied Mechanics and Engineering 198(21–26) (2009), S. 1785–1804.
- [81] LEFIK, M. / SCHREFLER, B.A.: *Artificial neural network as an incremental non-linear constitutive model for a finite element code*. Computer Methods in Applied Mechanics and Engineering 192(28–30) (2003), S. 3265–3283.
- [82] LEMAITRE, J. / CHABOCHE, J.-L.: *Mechanics of solid materials*. Cambridge Univ. Press, Cambridge, 2002.
- [83] LEVENBERG, K.: *A Method for the solution of certain non-linear problems in least squares*. Quarterly of Applied Mathematics 2 (1944), S. 164–168.
- [84] LINKA, K. / HILLGÄRTNER, M. / ABDOLAZIZI, K.P. / AYDIN, R.C. / ITSKOV, M. / CYRON, C.J.: *Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning*. Journal of Computational Physics 429 (2021), S. 110010.
- [85] LIU, Z. / WANG, Y. / VAIDYA, S. / RUEHLE, F. / HALVERSON, J. / SOLJAČIĆ, M. / HOU, T. Y. / TEGMARK, M.: *KAN: Kolmogorov-Arnold Networks*. arXiv preprint, 2024.
- [86] LUBLINER, J.: *On the thermodynamic foundations of non-linear solid mechanics*. International Journal of Non-Linear Mechanics 7(3) (1972), S. 237–254.
- [87] LUBLINER, J.: *Plasticity Theory*. Dover Books on Engineering Series Dover Publications, Incorporated, Newburyport, 2008.
- [88] MALVERN, L.E.: *Introduction to the mechanics of a continuous medium*. Prentice-Hall Series in Engineering of the Physical Sciences, 2007.
- [89] MARQUARDT, D.W.: *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*. Journal of the Society for Industrial and Applied Mathematics 11(2) (1963), S. 431–441.
- [90] MARSDEN, J.E.: *Mathematical Foundations of Elasticity*. Dover Civil and Mechanical Engineering, Dover Publications, Newburyport, 2012.
- [91] MASI, F. / STEFANOU, I. / VANNUCCI, P. / MAFFI-BERTHIER, V.: *Thermodynamics-based Artificial Neural Networks for constitutive modeling*. Journal of the Mechanics and Physics of Solids 147 (2021), S. 104277.
- [92] MCCULLOCH, W.S. / PITTS, W.: *A logical calculus of the ideas immanent in nervous activity*. The Bulletin of Mathematical Biophysics 5(4) (1943), S. 115–133.
- [93] MICHEL, J.C. / MOULINEC, H. / SUQUET, P.: *Effective properties of composite materials with periodic microstructure: a computational approach*. Computer Methods in Applied Mechanics and Engineering 172(1–4) (1999), S. 109–143.

-
- [94] MIEHE, C.: *Numerical computation of algorithmic (consistent) tangent moduli in large-strain computational inelasticity*. Computer Methods in Applied Mechanics and Engineering 134(3–4) (1996), S. 223–240.
- [95] MIEHE, C. / KOCH, A.: *Computational micro-to-macro transitions of discretized microstructures undergoing small strains*. Archive of Applied Mechanics 72(4–5) (2002), S. 300–317.
- [96] MIEHE, C. / SCHRÖDER, J. / BECKER, M.: *Computational homogenization analysis in finite elasticity: material and structural instabilities on the micro- and macro-scales of periodic composites and their interaction*. Computer Methods in Applied Mechanics and Engineering 191(44) (2002), S. 4971–5005.
- [97] MINSKY, M. / PAPERT, S.: *Perceptrons*. MIT Press, Cambridge, Mass., 5. Auflage, 1990.
- [98] MØLLER, M.F.: *A scaled conjugate gradient algorithm for fast supervised learning*. Neural Networks 6(4) (1993), S. 525–533.
- [99] NEMAT-NASSER, S.: *Averaging theorems in finite deformation plasticity*. Mechanics of Materials 31(8) (1999), S. 493–523.
- [100] NOCEDAL, J.: *Numerical optimization*. Springer series in operation research and financial engineering Springer, New York, NY, 2. Auflage, 2006.
- [101] NOLL, W.: *A new mathematical theory of simple materials*. Archive for Rational Mechanics and Analysis 48(1) (1972), S. 1–50.
- [102] OESER, M. / FREITAG, S.: *Modeling of materials with fading memory using neural networks*. International Journal for Numerical Methods in Engineering 78(7) (2008), S. 843–862.
- [103] OGDEN, R.W.: *Non-linear elastic deformations*. Dover books on physics, Dover Publ., Mineola, NY, 1. Auflage, 1997.
- [104] PAGOLI, A. / CHAPPELLE, F. / CORRALES-RAMON, J.-A. / MEZOUAR, Y. / LAPUSTA, Y.: *Review of soft fluidic actuators: classification and materials modeling analysis*. Smart Materials and Structures 31(1) (2021), S. 013001.
- [105] PEEL, L.D.: *Fabrication and mechanics of fiber-reinforced elastomers*. Dissertation, Brigham Young University, 1998.
- [106] PENG, G.C.Y. / ALBER, M. / BUGANZA TEPOLE, A. / CANNON, W.R. / DE, S. / DURA-BERNAL, S. / GARIKIPATI, K. / KARNIADAKIS, G. / LYTTON, W.W. / PERDIKARIS, P. / PETZOLD, L. / KUHL, E.: *Multiscale Modeling Meets Machine Learning: What Can We Learn?* Archives of Computational Methods in Engineering 28(3) (2020), S. 1017–1037.

-
- [107] PRECHELT, L.: *Early Stopping — But When?* S. 53–67, Springer, Berlin, Heidelberg, 2012.
- [108] RAISSI, M. / PERDIKARIS, P. / KARNIADAKIS, G.E.: *Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*. Journal of Computational Physics 378 (2019), S. 686–707.
- [109] REUSS, A.: *Berechnung der Fließgrenze von Mischkristallen auf Grund der Plastizitätsbedingung für Einkristalle*. ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik 9(1) (1929), S. 49–58.
- [110] RIEDMILLER, M. / BRAUN, H.: *A direct adaptive method for faster backpropagation learning: the RPROP algorithm*. In: *IEEE International Conference on Neural Networks* IEEE, 1993.
- [111] ROCHA, I.B.C.M. / KERFRIDEN, P. / VAN DER MEER, F.P.: *On-the-fly construction of surrogate constitutive models for concurrent multiscale mechanical analysis through probabilistic machine learning*. Journal of Computational Physics: X 9 (2021), S. 100083.
- [112] ROSENBLATT, F.: *The perceptron: A probabilistic model for information storage and organization in the brain*. Psychological Review 65(6) (1958), S. 386–408.
- [113] ROSENKRANZ, M. / KALINA, K.A. / BRUMMUND, J. / KÄSTNER, M.: *A comparative study on different neural network architectures to model inelasticity*. International Journal for Numerical Methods in Engineering 124(21) (2023), S. 4802–4840.
- [114] ROSENKRANZ, M. / KALINA, K.A. / BRUMMUND, J. / SUN, W. / KÄSTNER, M.: *Viscoelasticity with physics-augmented neural networks: model formulation and training methods without prescribed internal variables*. Computational Mechanics 74(6) (2024), S. 1279–1301.
- [115] RUMELHART, D.E. / HINTON, G.E. / WILLIAMS, R.J.: *Learning representations by back-propagating errors*. Nature 323(6088) (1986), S. 533–536.
- [116] RUST, W.: *Nichtlineare Finite-Elemente-Berechnungen*. Springer Vieweg, Wiesbaden, 3. Auflage, 2016.
- [117] SAEB, S. / STEINMANN, P. / JAVILI, A.: *Aspects of Computational Homogenization at Finite Deformations: A Unifying Review From Reuss' to Voigt's Bound*. Applied Mechanics Reviews 68(5) (2016).
- [118] SALETAN, E.J.: *Theoretische Mechanik*. Oldenbourg, München, 1974.

-
- [119] SCHOMMARTZ, J.O. / KLEIN, D.K. / ALZATE COBO, J.C. / WEEGER, O.: *Physics-augmented neural networks for constitutive modeling of hyperelastic geometrically exact beams*. Computer Methods in Applied Mechanics and Engineering 435 (2025), S. 117592.
- [120] SCHÜRMANN, H.: *Konstruieren mit Faser-Kunststoff-Verbunden*. VDI-Buch Springer, Berlin, Heidelberg, 2. Auflage, 2007.
- [121] SIMO, J.C.: *On a fully three-dimensional finite-strain viscoelastic damage model: Formulation and computational aspects*. Computer Methods in Applied Mechanics and Engineering 60(2) (1987), S. 153–173.
- [122] SIMO, J.C.: *Computational Inelasticity*. Springer New York, 2000.
- [123] SIMO, J.C. / RIFAI, M.S.: *A class of mixed assumed strain methods and the method of incompatible modes*. International Journal for Numerical Methods in Engineering 29(8) (1990), S. 1595–1638.
- [124] SIMO, J.C. / TAYLOR, R.L. / PISTER, K.S.: *Variational and projection methods for the volume constraint in finite deformation elasto-plasticity*. Computer Methods in Applied Mechanics and Engineering 51(1–3) (1985), S. 177–208.
- [125] STRANG, G.: *An analysis of the finite element method*. Wellesley-Cambridge Press, Wellesley, Mass., 2. Auflage, 2008.
- [126] STROEVEN, M. / ASKES, H. / SLUYS, L.J.: *Numerical determination of representative volumes for granular materials*. Computer Methods in Applied Mechanics and Engineering 193(30–32) (2004), S. 3221–3238.
- [127] TANG, B.: *Orthogonal Array-Based Latin Hypercubes*. Journal of the American Statistical Association 88(424) (1993), S. 1392–1397.
- [128] TAYLOR, R.L.: *FEAP - Finite Element Analysis Program*. Version IBS-KIT, University of California, Berkeley (2025).
- [129] TAYLOR, R.L. / PISTER, K.S. / GOUDREAU, G.L.: *Thermomechanical analysis of viscoelastic solids*. International Journal for Numerical Methods in Engineering 2(1) (1970), S. 45–59.
- [130] TEMIZER, İ. / WRIGGERS, P.: *On the computation of the macroscopic tangent for multiscale volumetric homogenization problems*. Computer Methods in Applied Mechanics and Engineering 198(3–4) (2008), S. 495–510.
- [131] THE MATHWORKS INC., NATICK, MASSACHUSETTS UNITED STATES: *MATLAB version: 9.14.0 (R2023a)*. 2023.
- [132] TORQUATO, S.: *Random heterogeneous materials*. Nummer 16 in der Reihe: Interdisciplinary Applied Mathematics, Springer, New York, 2002.

-
- [133] TRELOAR, L.R.G.: *Stress-strain data for vulcanised rubber under various types of deformation*. Transactions of the Faraday Society 40 (1944), S. 59.
- [134] TRUESDELL, C.: *The Non-Linear Field Theories of Mechanics*. Springer, Berlin, Heidelberg, 3. Auflage, 2004.
- [135] TURNER, M.J. / CLOUGH, R.W. / MARTIN, H.C. / TOPP, L.J.: *Stiffness and Deflection Analysis of Complex Structures*. Journal of the Aeronautical Sciences 23 (1956), S. 805–823.
- [136] UNGER, J.F. / KÖNKE, C.: *Coupling of scales in a multiscale simulation using neural networks*. Computers & Structures 86(21–22) (2008), S. 1994–2003.
- [137] UNGER, J.F. / KÖNKE, C.: *Neural networks as material models within a multiscale approach*. Computers & Structures 87(19–20) (2009), S. 1177–1186.
- [138] VLASSIS, N.N. / SUN, W.: *Sobolev training of thermodynamic-informed neural networks for interpretable elasto-plasticity models with level set hardening*. Computer Methods in Applied Mechanics and Engineering 377 (2021), S. 113695.
- [139] VOIGT, W.: *Theoretische Studien über die Elasticitätsverhältnisse der Krystalle. I.* Abhandlungen der Königlich Gesellschaft der Wissenschaften in Göttingen 34 (1887), S. 3–52.
- [140] WAGNER, W. / GRUTTMANN, F.: *A robust non-linear mixed hybrid quadrilateral shell element*. International Journal for Numerical Methods in Engineering 64(5) (2005), S. 635–666.
- [141] WAGNER, W. / GRUTTMANN, F.: *An improved quadrilateral shell element based on the Hu–Washizu functional*. Advanced Modeling and Simulation in Engineering Sciences 7(28) (2020).
- [142] WAGNER, W. / WRIGGERS, P.: *A simple method for the calculation of postcritical branches*. Engineering Computations 5(2) (1988), S. 103–109.
- [143] WASHIZU, K.: *Variational methods in elasticity and plasticity*. Pergamon Press, Oxford, 3. Auflage, 1982.
- [144] WASZCZYSZYN, Z.: *Neural Networks in the Analysis and Design of Structures*. CISM International Centre for Mechanical Sciences, Springer, Vienna, 1999.
- [145] WEBER, P.: *Physikalisch konsistente Modellierung von elastischem und plastischem Materialverhalten mit künstlichen neuronalen Netzen*. Dissertation, Institut für Baustatik, Karlsruher Institut für Technologie, 2023.
- [146] WEBER, P. / GEIGER, J. / WAGNER, W.: *Constrained neural network training and its application to hyperelastic material modeling*. Computational Mechanics 68(5) (2021), S. 1179–1204.

-
- [147] WEBER, P. / WAGNER, W. / FREITAG, S.: *Physically enhanced training for modeling rate-independent plasticity with feedforward neural networks*. Computational Mechanics 72(4) (2023), S. 827–857.
- [148] WEBER, P. / WAGNER, W. / FREITAG, S.: *ANN material modeling for SMA fibers enhanced with a physical constraint and its application to FE computations*. Computational Mechanics 76(1) (2025), S. 93–116.
- [149] WHITE, A. / ARRIGHI, W.J. / KUDO, J. / WATTS, S.E.: *Multiscale topology optimization using neural network surrogate models*. Computer Methods in Applied Mechanics and Engineering 346 (2019), S. 1118–1135.
- [150] WIDROW, B. / HOFF, M.E.: *Adaptive switching circuits*. In: *In Proceedings WESCON S. 96–104*, 1960.
- [151] WILAMOWSKI, B.M. / IRWIN, J.D.: *Intelligent Systems: The Industrial Electronics Handbook*. CRC Press, 2018.
- [152] WRIGGERS, P.: *Nichtlineare Finite-Element-Methoden*. Springer, Berlin, Heidelberg, 2001.
- [153] XAVIER, M.S. / TAWK, C.D. / ZOLFAGHARIAN, A. / PINSKIER, J. / HOWARD, D. / YOUNG, T. / LAI, J. / HARRISON, S. M. / YONG, Y.K. / BODAGHI, M. / FLEMING, A.J.: *Soft Pneumatic Actuators: A Review of Design, Fabrication, Modeling, Sensing, Control and Applications*. IEEE Access 10 (2022), S. 59442–59485.
- [154] ZHANG, Q. B. / ZHAO, J.: *A Review of Dynamic Experimental Techniques and Mechanical Behaviour of Rock Materials*. Rock Mechanics and Rock Engineering 47(4) (2013), S. 1411–1478.
- [155] ZIENKIEWICZ, O.C. / TAYLOR, R.L. / S., GOVINDJEE: *The finite element method*. Lecture Notes in Applied and Computational Mechanics Butterworth-Heinemann, London, 8. Auflage, 2025.
- [156] ZOHDI, T.I. / WRIGGERS, P.: *An introduction to computational micromechanics*. Lecture Notes in Applied and Computational Mechanics, Springer, Berlin, 2. Auflage, 2008.

Berichte des Instituts für Baustatik, Karlsruher Institut für Technologie

Verzeichnis der bisher erschienenen Berichte:

- 51/1: Beitrag zur Ermittlung der zeitabhängigen Eigenspannungen in Verbundträgern, Dissertation, Hansjürgen Sonntag, 1951.
- 54/2: Beiträge zur Ermittlung der zeitabhängigen Spannungen in Verbundträgern, Dissertation, Eugen Müller, 1954.
- 54/3: Zur Statik räumlich gekrümmter Träger, Dissertation, Heinrich Bechert, 1954.
- 55/4: Bemessungsverfahren für Verbundträger, Dissertation, Günter Utescher, 1955.
- 57/5: Einflußfelder für die Hauptträgerschnittkräfte zweistegiger Plattenbalkensysteme, Dissertation, Otto Köller, 1957.
- 60/6: Beitrag zur numerischen Berechnung der Schnittkräfte von rechteckigen und schiefen randversteiften Platten, Dissertation, Klaus Stiglat, 1960.
- 61/7: Allgemeines und einheitliches Berechnungsverfahren für Verbundkonstruktionen aus Stahl und Beton, Dissertation, Herbert Wippel, 1961.
- 62/8: Beitrag zur Bemessung vorgespannter stählerner Fachwerkträger mit besonderer Berücksichtigung der Tragsicherheit, Dissertation, Wolfgang Heidecker, 1962.
- 63/9: Beitrag zur Berechnung der Schnittkräfte und Durchbiegungen von umfanggelagerten Rechteckplatten mit veränderlichem Querschnitt in einer Spannrichtung, Dissertation, Ernst Buchholz, 1963.
- 63/10: Der winkelrecht zu seiner Ebene belastete, kreisförmig gekrümmte Träger mit starrer Torsionseinspannung an den Auflagerpunkten, Dissertation, Hans Wittfoht, 1963.
- 64/11: Ein Verfahren zur Berechnung biegsamer rechteckiger Gründungsplatten mit elektronischen Rechengegeräten, Dissertation, Artur Deninger, 1964.
- 64/12: Anwendung des Momentenausgleichsverfahrens für durchlaufende schiefwinklige Platten, Dissertation, Jahidul Alum, 1964.

- 67/13: Beitrag zur Berechnung des Tragvermögens von Rahmentragwerken aus Stahl nach der Plastizitätstheorie II. Ordnung, Dissertation, Mieng-Siong Wang, 1967.
- 67/14: Zusammenwirken von biegesteifem Skelett und aussteifenden Wandscheiben in Hochhauskonstruktionen, Dissertation, Rolf Zickendraht, 1967.
- 68/15: Experimentelle Untersuchung von beliebig gelagerten Plattentragwerken mit Hilfe von elektrischen Analogieschaltungen, Dissertation, Werner Sitzler, 1968.
- 69/16: Operatorenkalkül zur Berechnung von Stab- und Flächentragwerken, Dissertation, Günter Eisenbiegler, 1969.
- 72/17: Umlagerungen der Schnittkräfte und Verformungen von biegebeanspruchten Stahlbetonkonstruktionen im Zustand II infolge Kriechens und Schwindens, Dissertation, Walter H. Schneider, 1972.
- 73/18: Berechnungssystematik für Rechteckscheiben, Habilitation, Günter Eisenbiegler, 1973.
- 1/75 : Berechnung einfeldriger und durchlaufender Kastenträger unter Berücksichtigung der Querschnittsdeformation, Dissertation, Jürgen Hanisch, 1975.
- 2/75 : Der Plattenbalken mit Lastangriff innerhalb des Steges, Dissertation, Hans-Erich Heining, 1975.
- 3/76 : Das Tragverhalten längsversteifter, vorverformter Rechteckplatten unter Axialbelastung nach der nichtlinearen Beultheorie (unter Berücksichtigung der orthotropen Struktur sowie einer elastischen Lagerung am Längsrand), Habilitation, Helmut Rubin, 1976.
- 4/78 : Traglastuntersuchungen ebener Rahmen aus Stahl unter Berücksichtigung teilplastischer Zonen, Dissertation, Ülgen Korkut, 1978.
- 5/79 : Traglastermittlung von räumlich belasteten Durchlaufträgern mit offenem, dünnwandigem Querschnitt bei beliebigem Werkstoffgesetz, Dissertation, Wolfgang Heil, 1979.
- 6/79 : Schnittgrößen im Verschneidungsbereich eines kreiszylindrischen Druckbehälters mit einem Stutzen beliebiger Länge, Dissertation, Kurt Schumer, 1979.
- 7/81 : Traglastberechnung räumlicher Rahmen aus Stahl- oder Leichtmetallprofilen mit dünnwandigen offenen Querschnitten, Dissertation, Theodor Ackermann, 1981.
- 8/83 : Berechnung orthotroper Rechteckplatten nach der nichtlinearen Elastizitätstheorie für beliebige Randbedingungen, Dissertation, Karl Morgen, 1983.

- 9/86 : Traglastermittlung räumlicher Stabtragwerke aus Stahl und Leichtmetall unter Berücksichtigung der Schubweichheit, Dissertation, Dietmar H. Maier, 1986.
- 10/89 : Bemessung von Raumfachwerken bezüglich ihrer Grenztragfähigkeit mit Hilfe der mathematischen Optimierung, Dissertation, Richard Harich, 1989.
- 11/89 : Quasi-Newton Verfahren und Kurvenverfolgungsalgorithmen für die Lösung nichtlinearer Gleichungssysteme in der Strukturmechanik, Forschungsbericht, Karl Schweizerhof, 1989.
- 12/90 : Finite Element Konzepte für schubelastische Schalen mit endlichen Drehungen, Dissertation, Heinrich Gebhardt, 1990.
- 13/91 : Theorie II. Ordnung des eben gekrümmten Stabes bei nicht richtungstreuen Kräften, Dissertation, Peter Brković, 1991.
- 14/91 : Nichtlineare Finite Element Berechnungen von Schalentragwerken mit geschichtetem anisotropen Querschnitt, Dissertation, Rudolf Klarmann, 1991.
- 15/94 : Adaptive Finite Element Konzepte zur Analyse von Schalentragwerken, Dissertation, Markus Baumann, 1994.
- 1 (1995) : Optimierung von Stockwerkrahmen aus Stahl mit biegesteifen und biegeweichen Rahmenknoten, Dissertation, Petra Höß, 1994, ISBN: 978-3-935322-00-3.
- 2 (1996) : Theorie und Numerik dünnwandiger Faserverbundstrukturen, Forschungsbericht, Friedrich Gruttmann, 1996, ISBN: 978-3-935322-01-1.
- 3 (1997) : Stabilität von Stahlzylindern mit unilateralen Randbedingungen bei statischen und dynamischen Beanspruchungen, Forschungsbericht, Klaus Knebel, 1997, ISBN: 978-3-935322-02-7.
- 4 (1998) : Eine einheitliche Finite-Element-Formulierung für Stab- und Schalentragwerke mit endlichen Rotationen, Dissertation, Roland Sauer, 1998, ISBN: 978-3-935322-03-4.
- 5 (1999) : Finite-Element-Modellierung von Starrkörper- und Übergangsbedingungen in der Statik, Dissertation, Joachim Kugler, 1999, ISBN: 978-3-935322-04-1.
- 6 (2000) : Delaminationsanalysen von Faserverbundstrukturen mit der Methode der finiten Elemente, Dissertation, Wolfgang Sprenger, 2000, ISBN: 978-3-935322-05-8.
- 7 (2000) : Theorie und Numerik eines Volumen-Schalen-Elementes bei finiten elastischen und plastischen Verzerrungen, Dissertation, Sven Klinkel, 2000, ISBN: 978-3-935322-06-5.

- 8 (2001) : Theorie und Numerik für den Entwurf von vorgespannten Membrantragwerken, Dissertation, René Ziegler, 2001, ISBN: 978-3-935322-07-2.
- 9 (2005) : Ein inelastisches 3D-Versagensmodell für Beton und seine Finite-Element-Implementierung, Dissertation, Jan Schütt, 2005, ISBN: 978-3-935322-08-9.
- 10 (2006) : Formfindung der Fließliniengeometrie für polygonale Platten im Traglastzustand, Dissertation, Jochen Wüst, 2006, ISBN: 978-3-935322-09-6.
- 11 (2006) : Nichtlineare Formulierung piezoelektrischer 3D-Stabstrukturen - Theorie und Finite-Element-Modellierung, Dissertation, Alexander Butz, 2006, ISBN: 978-3-935322-10-2.
- 12 (2007) : Nichtlineare Modellierung ferroelektrischer Keramiken und piezoelektrischer Strukturen – Analyse und Finite-Element-Formulierung, Habilitationsschrift, Sven Klinkel, 2007, ISBN: 978-3-935322-11-9.
- 13 (2007) : Ein geometrisch und materiell nichtlineares Cosserat-Modell – Theorie, Numerik und Anwendungsmöglichkeiten, Dissertation, Ingo Münch, 2007, ISBN: 978-3-935322-12-6.
- 14 (2008) : Magnetostriktive und piezoelektrische Materialien – Konstitutive Modellierung und Finite-Element-Formulierung, Dissertation, Konrad Linnemann, 2008, ISBN: 978-3-935322-13-3.
- 15 (2008) : Theorie und Numerik unidirektional verstärkter Faserverbundwerkstoffe: 3D Finite-Elemente-Untersuchungen der Faser-Matrix Mikroinstabilitäten, Dissertation, Jens Harich, 2008, ISBN: 978-3-935322-15-7.
- 16 (2009) : Finite Element Modeling of Intra- and Interlaminar Damage Growth in Composite Laminates, Dissertation, Claudio Balzani, 2009, ISBN: 978-3-935322-16-4.
- 17 (2010) : Theorie und Finite-Elemente-Modellierung geometrisch und materiell nichtlinearer piezoelektrischer Schalenstrukturen, Dissertation, Katrin Schulz, 2010, ISBN: 978-3-935322-17-1.
- 18 (2011) : Finite-Elemente-Formulierung mit abgestimmten Approximationsräumen für die Modellierung piezoelektrischer Stab- und Schalenstrukturen, Dissertation, Dieter Legner, 2011, ISBN: 978-3-935322-18-8.
- 19 (2014) : Nichtlineare elektromechanisch gekoppelte Viskoelastizität bei großen Verformungen am Beispiel dielektrischer Elastomere – Konstitutive Modellierung und Finite-Elemente-Formulierung, Dissertation, Alexander Büschel, 2014, ISBN: 978-3-935322-19-5.

- 20 (2014) : Duroplastische Kunststoffe mit Elastomerzusätzen - Konstitutive Modellierung und Finite-Elemente-Formulierung, Dissertation, Stefan Lacher, 2014, ISBN: 978-3-935322-20-1.
- 21 (2014) : Numerische Stabilitätsuntersuchung dünnwandiger Tragwerke unter dem Einfluss stochastischer Imperfektionen, Dissertation, Stefan Lauterbach, 2014, ISBN: 978-3-935322-21-8.
- 22 (2017) : Entwicklung robuster Ein- und Mehrskalen-Modelle zur Simulation von Delaminationsprozessen faserverstärkter Lamine, Dissertation, Tillmann Herwig, 2017, ISBN: 978-3-935322-22-5.
- 23 (2017) : Entwurf und Finite-Elemente-Modellierung ferroelektrischer Nanostrukturen zur Transformation mechanischer Deformationen in elektrische Energie, Dissertation, Matthias Krauß, 2017, ISBN: 978-3-935322-23-2.
- 24 (2018) : Entwicklung eines gradientenbasierten Modells zur netzunabhängigen Schädigungssimulation geschichteter Faserverbundstrukturen, Dissertation, Jonas Läufer, 2018, ISBN: 978-3-935322-24-9.
- 25 (2018) : Ein gemischtes FE-Modell zur Untersuchung des nichtlinearen Deformations- und Spannungsverhaltens schubweicher Lamine, Dissertation, Max Köppler, 2018, ISBN: 978-3-935322-25-6.
- 26 (2020) : Metaheuristische Form- und Systemoptimierung von Stabtragwerken mit Evolutionsstrategien, Dissertation, Alexander Keller, 2020, ISBN: 978-3-935322-26-3.
- 27 (2020) : Polymorphe Unschärfemodellierung in der nichtlinearen Strukturmechanik – Stabilität von Schalenträgwerken, räumliche Variabilität und Metamodellierung, Dissertation, Marc Fina, 2020, ISBN: 978-3-935322-27-0.
- 28 (2023) : Physikalisch konsistente Modellierung von elastischem und plastischem Materialverhalten mit künstlichen neuronalen Netzen, Dissertation, Patrick Weber, 2024, ISBN: 978-3-935322-28-7.
- 29 (2024) : Ein Phasenfeldmodell zur Simulation des Einflusses von Leitungsmechanismen in einem ferroelektrischen Nanogenerator, Dissertation, Franziska Wöhler, 2024, ISBN: 978-3-935322-29-4.
- 30 (2026) : Theorie und Numerik spektraler stochastischer Finite-Elemente-Formulierungen bei geometrisch nichtlinearem Deformationsverhalten, Dissertation, Lukas Panther, 2026, ISBN: 978-3-935322-30-0

Wissenschaftlicher Lebenslauf

Persönliche Daten

Name Jeremy Geiger
Geburtsdatum 17. Januar 1994
Geburtsort Stuttgart

Ausbildung

2000–2013 Schulbildung
 Ulrichsgymnasium, Norden
 Abitur

2013–2017 Bachelorstudium Bauingenieurwesen
 Karlsruher Institut für Technologie (KIT)
 Bachelorarbeit: Institut für Baustatik

2017–2019 Masterstudium Funktionaler und Konstruktiver Ingenieurbau
 - Engineering Structures
 Karlsruher Institut für Technologie (KIT)
 Profil Modellierung und Simulation im Ingenieurbau
 Masterarbeit: Institut für Baustatik

2017–2019 Werkstudent
 IBT Ingenieurbau GbR

2020–2026 Wissenschaftlicher Mitarbeiter
 Karlsruher Institut für Technologie (KIT)
 Institut für Baustatik

Preise und Auszeichnungen

2020 DYNAmore Preis (DYNAmore GmbH Stuttgart)
 Auszeichnung für die Masterarbeit

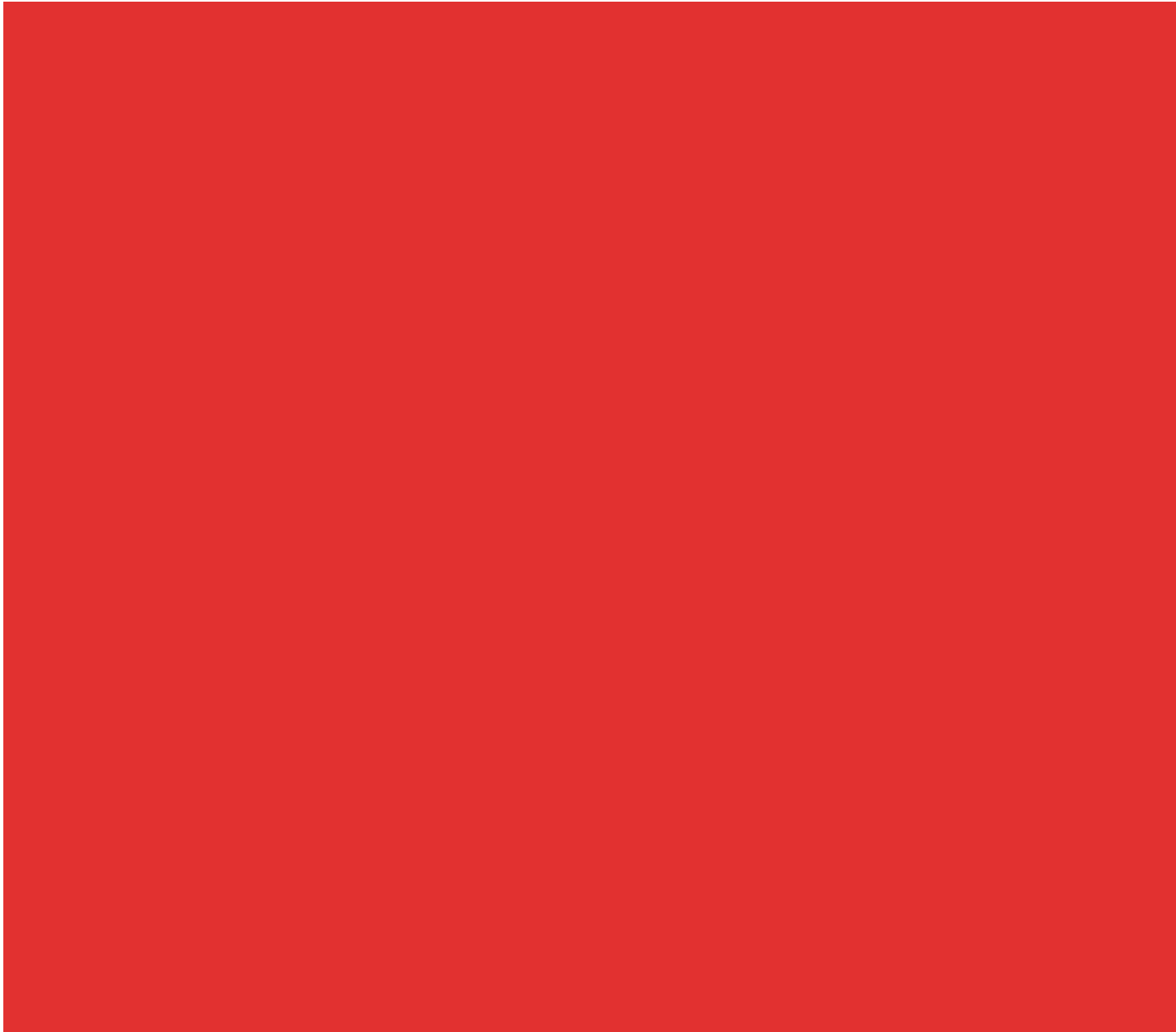
Ehrenamtliches Engagement

2024–2026 Vorstand BGU Doktorandenkonvent
 Mitglied des Vorstands (2024–2026)
 1. Vorsitzender (2025–2026)

Veröffentlichungen

- Zeitschriftenartikel (Peer Review)
 - **Geiger, J.**, Wagner, W., Freitag, S.: Effective Material Modeling of Parameterizable Viscoelastic Shell Structures with Artificial Neural Networks, Proceedings in Applied Mathematics and Mechanics, 25, Nr. 3, 2025, <https://doi.org/10.1002/pamm.70015>
 - **Geiger, J.**, Wagner, W., Freitag, S.: Multiscale modeling of viscoelastic shell structures with artificial neural networks, Computational Mechanics, 76, S. 541–565, 2025, <https://doi.org/10.1007/s00466-025-02613-5>
 - Weber, P., **Geiger, J.**, Wagner, W.: Constrained neural network training and its application to hyperelastic material modeling, Computational Mechanics, 68, S. 1179–1204, 2021, <https://doi.org/10.1007/s00466-021-02064-8>
- Konferenzbeiträge (Artikel)
 - **Geiger, J.**, Wagner, W., Freitag, S.: Effektive Modellierung von Flächentragwerken mit zeitabhängigem Materialverhalten auf Basis künstlicher neuronaler Netze, 16. Fachtagung Baustatik-Baupraxis (im Druck), 28.-29. Mai 2026, Dresden
- Weitere Konferenzbeiträge (Vortrag)
 - **Geiger, J.**; Wagner, W.; Freitag, S.: Effective material modeling of viscoelastic structures with artificial neural networks: from material point to FEM, Forschungskolloquium FE ohne Schnee, 20.-23. Juli, 2025, Hirschegg, Österreich
 - **Geiger, J.**; Wagner, W.; Freitag, S.: Effective material modeling of complex viscoelastic shell structures with artificial neural networks, 95th annual meeting of the International Association of Applied Mathematics and Mechanics (GAMM), 7-11 April, 2025, Posen, Polen
 - **Geiger, J.**; Wagner, W.; Freitag, S.: Multiscale modeling of viscoelastic shell structures with artificial neural networks, 9th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS), 03.-07. Juni, 2024, Lissabon, Portugal
 - **Geiger, J.**; Wagner, W.; Freitag, S.: Constrained Neural Network Training with Application to Multiscale Modeling of Inelastic and Time-Dependent Materials 2nd IACM Mechanistic Machine Learning and Digital Engineering for Computational Science Engineering and Technology, 24.-27. September, El Paso, USA

- **Geiger, J.**, Wagner, W., Freitag, S.: Sobolev Training for Artificial Neural Networks with Application to Multiscale Modeling of Shell Structures 9th GACM Colloquium on Computational Mechanics, 21.-23. September, 2022, Essen, Deutschland
- **Geiger, J.**, Wagner, W., Freitag, S.: Multiscale modeling of shell structures with artificial neural networks 8th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS), 05.-09. Juni, 2022, Oslo, Norwegen
- Weber, P., **Geiger, J.**, Wagner, W., Freitag, S.: Incorporating physical restrictions by constraint optimization techniques for ANN-based hyper-elastic material modeling, 7th GAMM AG Data Workshop, 15. Dezember, 2021, online



ISBN 978-3-935322-31-7