
THERMOMECHANICAL MODELING OF SHAPE MEMORY ALLOY-BASED MICROACTUATORS

**in partial fulfillment of the requirements for the degree of
Doctor of Engineering
(Dr.-Ing.)**

Dissertation
by Marian Hörsting
from Greven

Submitted to the Faculty of Engineering of
Kiel University

Examiners: Prof. Dr.-Ing. Stephan Wulfinghoff
Prof. Dr.-Ing. Johanna Waimann

Date of submission: 29.10.2022
Date of oral examination: 17.03.2023

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Mir wurde in der Vergangenheit kein akademischer Grad entzogen. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen. Der Inhalt der Arbeit wurde in Teilen bereits in meinen wissenschaftlichen Publikationen veröffentlicht. Dies ist in der Arbeit entsprechend vermerkt. Die Arbeit ist nach bestem Wissen und Gewissen konform mit den Regeln guter wissenschaftlicher Praxis, welche durch die Deutsche Forschungsgemeinschaft festgelegt sind.

Kiel, der 19.10.2022

(Marian Hörsting)

To my father, in loving memory.

Acknowledgments

The present dissertation is the result of my time at the working group for Computational Materials Science at Kiel University. First of all, I want to thank my supervisor and mentor Prof. Dr.-Ing. Stephan Wulfinghoff for his support, everlasting patience and guidance. In particular, I enjoyed his trust in my work and freedom beyond example he provided at his new working group, where I was given every opportunity one needs to strive in mechanics. Second, I would like to thank my colleagues from our research project on *Cooperative Actuator Systems for Nanomechanics and Nanophotonics*, who broadened my field of view for problems in research outside of the world of mechanics. Thank you for the fruitful collaboration and various data you provided!

I also would like to thank Prof. Dr.-Ing. Johanna Waimann for being the second examiner to this thesis.

A big "Thank you" to my colleagues Christian Dorn, Jan Dittmann and Muhammad Babar Shamim. We all shared one modest office (though sometimes only virtual, just like the forces we dealt with), which, fueled by discussions on mechanics as well as everybody and his dog, never got boring. I hope from the bottom of my heart that there will be even more cake in the future!

My love and sincerest gratitude goes out to my friends and family. In particular, I owe a big thank you to my mother Dorothee, my brother Julian and my sister Anna. Your unconditional support since I can remember means a lot to me.

Finally, I want to thank my wife Lena. You support me in every imaginable way in life and never fail to cheer me up. I am looking forward to everything the future holds for us.

Zusammenfassung

Die Modellierung von Inelastizität unter finiten Deformationen beinhaltet häufig eine Inkompressibilitätsbeschränkung für die inelastischen Dehnungen, welche sich aus physikalischen Überlegungen ergibt. In der Regel wird diese Bedingung durch die Verwendung der Exponentialabbildung als geometrischer Integrator für die Evolutionsgleichungen exakt erfüllt. In dieser Dissertation wird jedoch ein neuer geometrischer Integrator für die Unimodularitätsbeschränkung entwickelt und analysiert. Er baut auf der Arbeit von Hurtado et al. aus dem Jahr 2014 auf, wo dieses Projektionsschema für die Kristallplastizität eingeführt wurde. Um dieses Projektionsschema in einem Finite-Elemente-Kontext effizient nutzen zu können, müssen jedoch zusätzliche numerische Probleme überwunden werden.

Die vorliegende Arbeit soll dazu beitragen und bestehende Arbeiten für Formgedächtnislegierungen erweitern. Sie ist ein Zusammenschluss von drei Publikationen des Autors und seiner Mitautoren, die sich auf die Modellierung von Materialien mit einer Inkompressibilitätsbeschränkung konzentrieren. Das übergeordnete Ziel ist es, ein effizientes Modell für Formgedächtnislegierungen zur Simulation von kooperativen bistabilen Formgedächtnis-Nanoaktuatoren zu implementieren.

Diese Arbeit beginnt mit einer kurzen Einführung, die einen Einblick in den Stand der Technik bei der Modellierung der Inkompressibilität sowie der Modellierung von Formgedächtnislegierungen im Allgemeinen, einschließlich der zugrunde liegenden Physik, gibt.

Anschließend werden im ersten Artikel mehrere neue numerische Erweiterungen des Modells von Hurtado et al. aus dem Jahr 2014 vorgestellt. Im Einzelnen werden numerische Probleme mit Singularitäten aufgrund eines unterbestimmten Gleichungssystems und nicht verfügbarer Ableitungen gelöst. Außerdem wird eine neue Methode zur Approximation der Logarithmusfunktion für Tensoren verwendet, um die Genauigkeit der Zeitintegration zu verbessern. Eine weitere Besonderheit dieses Modells ist die Symmetrie der algorithmischen Tangente, die sich aus der Implementierung des Modells in das Rahmenkonzept der generalisierten Standardmaterialien ergibt.

Der zweite Artikel verwendet den in der ersten Veröffentlichung entwickelten geometrischen Integrator, um das Modell für Formgedächtnislegierungen mit kleinen Deformationen von Sedláček et al. aus dem Jahr 2012 auf den Fall großer Dehnungen zu erweitern. Das Modell wird wieder im generalisierten Standardmaterialien Rahmenkonzept formuliert, welches erweitert wurde, um eine konsistente Berücksichtigung einer thermomechanischen Kopplung zu ermöglichen. Dadurch ist das Modell in der Lage, den Formgedächtniseffekt sowie die Superelastizität in Abhängigkeit von der Temperatur über die Optimierung des globalen Potenzials zu beschreiben, das zusätzlich zur Ableitung von Fließkriterien, Randbedingungstermen sowie Evolutionsgleichungen der internen Variablen verwendet wird. Ein weiterer wichtiger Beitrag ist die Verwendung der logarithmischen Dehnung im Dissipationspotential, die sicherstellt, dass inelastische Dehnungen bei der Rückwärtstransformation verschwinden.

Die letzte Veröffentlichung befasst sich mit den Herausforderungen bei der Herstellung von

bistabilen Formgedächtnis-Nanoaktuatoren auf der Grundlage von Dünnschichttechnologie. In mehreren Experimenten werden verschiedene Wärmebehandlungen auf funktionale Ermüdungseigenschaften, Umwandlungstemperaturen, elektrischen Widerstand und Spannungs-Dehnungs-Verhalten bei unterschiedlichen Temperaturen untersucht. Darüber hinaus werden die im dritten Artikel ermittelten Materialparameter in Kombination mit dem im zweiten Artikel entwickelten Modell verwendet, um die angestrebten bistabilen Aktoren aus Formgedächtnislegierungen zu simulieren.

Abstract

Modeling finite deformation inelasticity often involves an incompressibility constraint on the inelastic stretches, which arises from physical considerations. Regularly, this constraint is fulfilled by use of the exponential map as a geometric integrator for the evolution equations. However, in this dissertation, a new geometric integrator for the unimodularity constraint is developed and analyzed. It builds on the work of Hurtado et al. from 2014, where this projection scheme was introduced for crystal plasticity. However, to make use of this projection scheme efficiently in a finite element context, additional numerical problems have to be overcome.

The work at hand aims to contribute to this aim and extend existing works for shape memory alloys. It comprises of three publications of the author and his co-authors concentrating on the modeling of materials with an incompressibility constraint. The overall goal is to implement an efficient shape memory alloy model for the simulation of cooperative bistable shape memory nanoactuators.

This thesis begins with a short introduction, which yields insight into the state of the art in modeling incompressibility as well as shape memory alloy modeling in general, including the underlying physics.

Subsequently, the first article introduces several new numerical enhancements to the model of Hurtado et al. from 2014. In detail, numerical problems involving singularities due to an underdetermined equation system and ill-defined derivatives are resolved. Furthermore, a new way to approximate the logarithm function for tensors is used to improve the time integration accuracy. A further characteristic of this model is the symmetry of the algorithmic tangent, which arises from the implementation of the model in the generalized standard materials framework.

The second article adopts the geometric integrator developed in the first publication to extend the small strain shape memory alloy model by Sedláč et al. from 2012 to the finite strain case. The model is again formulated in the generalized standard materials framework, which was extended to include thermomechanical coupling consistently. It is capable to describe the shape memory effect as well as superelasticity depending on the temperature via optimizing the global potential, which additionally is used to derive yield criteria, boundary condition terms as well as evolution equations of the internal variables. Another key contribution is the use of the logarithmic strain in the dissipation potential, which ensures that inelastic strains vanish upon backward transformation.

The final publication deals with the challenges in manufacturing bistable shape memory nanoactuators based on thin film technology. Several experiments are used to evaluate various heat treatments for functional fatigue characteristics, transformation temperatures, electrical resistance and stress-strain behavior for varying temperatures. Additionally, the in the third article obtained material parameters are used in the model developed in the second article to simulate the striven for bistable shape memory alloy actuators.

Contents

Eidesstattliche Erklärung	i
1 Introduction	1
1.1 Motivation	1
1.2 State of the art	2
1.2.1 Modeling the incompressibility constraint for plasticity	2
1.2.2 Shape memory alloy modeling	4
1.3 Outline of the dissertation	8
1.3.1 Publication 1	8
1.3.2 Publication 2	8
1.3.3 Publication 3	9
2 Article 1: Numerical strategies for variational updates in large strain inelasticity with incompressibility constraint	10
2.1 Introduction	11
2.2 Modeling of finite deformation plasticity	12
2.2.1 Kinematics	12
2.2.2 Casting finite deformation plasticity into the generalized standard materials framework	13
2.2.3 Time discretization of the rate potential for isotropic hardening	16
2.2.4 Treatment of the incompressibility constraint	17
2.2.5 Elimination of singularities	19
2.2.6 Improving the approximation of the logarithm	20
2.2.7 Initial guesses for $\Delta\gamma$ and $\tilde{\mathbf{N}}^s$	21
2.2.8 Solving the local and global system of equations	22

2.3	Numerical results for isotropic hardening	22
2.3.1	Gauss point evaluations	22
2.3.2	3D necking of a circular bar	23
2.3.3	Comparison of the approximations of the logarithm	25
2.4	Incorporation of kinematic hardening	25
2.4.1	Expanding the potential for kinematic hardening	25
2.4.2	Results for kinematic hardening	26
2.5	Summary and outlook	29
2.6	Appendix A: Consistency of the dissipation potential	30
2.7	Appendix B: Algorithmic yield criterion	30
2.8	Appendix C: First derivatives of π_Δ	30
2.9	Appendix D: Second derivatives of π_Δ	32
2.10	Appendix E: Consistent tangent	34
3	Article 2: A thermomechanical finite strain shape memory alloy model and its application to bistable actuators	37
3.1	Introduction	38
3.2	Modeling of shape memory alloys	41
3.2.1	Kinematics	41
3.2.2	(Im-)Balance equations	42
3.2.3	Helmholtz free energy	43
3.2.4	Dissipation potential	46
3.2.5	Transformation / yield criteria and inelastic evolution equations	47
3.2.6	Rate potential	48
3.3	Numerical strategies	50
3.3.1	Inelastic volume preservation	51
3.3.2	Differentiability at $\Delta_r \mathbf{C}^i = 0$	51
3.3.3	Algorithmic yield criterion	52
3.3.4	Algorithmic transformation criteria	54
3.3.5	Initial guess for $\tilde{\mathbf{N}}^s$	54
3.3.6	Computing derivatives of $\Delta\alpha$ for $\Delta\gamma \rightarrow 0$	54
3.3.7	Regularization of ψ_h	55
3.3.8	Viscosity of the martensite volume fraction ξ	56
3.3.9	Thermomechanical coupling	56
3.3.10	Active set search	57
3.4	Numerical results	60
3.4.1	Gauss point evaluations	60
3.4.2	Plate with a hole	62
3.4.3	Finite element actuator model	64
3.5	Summary and outlook	69

3.6	Appendix A: Discussion of the dissipation potential ϕ_Δ	71
3.7	Appendix B: Proof of $\tilde{N}^{s, tr}$ minimizing π_Δ for $\Delta\gamma \rightarrow 0$	71
3.8	Appendix C: First derivatives of π_Δ	72
3.9	Appendix D: Second derivatives of π_Δ	74
3.10	Appendix E: Consistent tangent	78
4	Article 3: TiNiHf/SiO₂/Si shape memory film composites for bidirectional micro actuation	81
4.1	Introduction	82
4.2	Methods and materials	84
4.2.1	Preparation of freestanding TiNiHf films	84
4.2.2	Preparation of TiNiHf/Si and TiNiHf/SiO ₂ /Si bimorphs	85
4.2.3	Tensile testing	85
4.2.4	Differential scanning calorimetry	85
4.2.5	Electrical resistance measurements	86
4.2.6	X-Ray Diffraction (XRD)	86
4.2.7	Cantilever deflection measurements	86
4.2.8	Constitutive modeling	87
4.3	Results	88
4.3.1	Functional properties of freestanding TiNiHf films	88
4.3.2	Functional properties of TiNiHf/Si and TiNiHf/SiO ₂ /Si bimorphs	91
4.3.3	Bi-directional actuation based on TiNiHf/SiO ₂ /Si bimorph film composites	95
4.3.4	Bistable actuation based on PMMA/TiNiHf/Si trimorph film composites	96
4.4	Discussion	99
4.5	Conclusions	101
5	Conclusions and outlook	103
5.1	Conclusions	103
5.2	Outlook	104
5.2.1	One to one numerical comparison of the exponential map and projection approach	104
5.2.2	Inclusion of functional fatigue into the shape memory alloy model	104
5.2.3	Improve shape memory model for nanoscale simulations via a gradient extension	105
5.2.4	Better validation by conducting more shape memory alloy experiments	105
	Bibliography	106

CHAPTER 1

Introduction

1.1 Motivation

In the past decades, the development of computers and simulation tools enabled engineers of all branches to design better, more efficient products. For example, in civil engineering, the introduction of the finite element method (FEM) enabled the prediction of complex systems by solving differential equations which were not manageable before. This facilitated leaner, cost-effective designs as well as increased performance under heavy loads, such as earthquakes and wind loads in tall structures. Likewise, in mechanical engineering, the use of simulations to predict many relevant features like functionality, energy requirement or safety reduces the development cost, time as well as use of costly materials and energy. Therefore, simulations and computer aided design gained more and more share of the product design process to a point where FEM simulations are now mandatory in many engineering branches.

Subsequently, the FEM played a role during the onset of the so-called fourth industrial revolution, where now not only the final product, but rather the entire production process is controlled and modeled. Here, digital twins, which are the virtual counterpart to the physical product, are simulated over the entire production process and sometimes even the entire lifecycle of the product. Firstly, this allows for a very powerful optimization of the production process, which makes use of the entire data sampled on a specific work piece. Secondly, this enables manufacturers to have very customizable products without the additional cost of having to use several production lines or equipment for each and every product. In short, an integrated simulation into the production and design process is very desirable, as it improves the production quality while decreasing costs.

However, to maximize the benefit of simulations, one needs models, which can describe the entire history of the system under consideration fast, accurately and robustly. This dissertation is concerned with developing a material model for shape memory alloys under large deformations. Here, to accurately model the underlying physics, it is important to exactly satisfy the volumetric incompressibility constraint for inelastic deformations.

1.2 State of the art

Modeling materials with incompressibility constraint has a long history, and over the years, several important concepts and algorithms were established and since then adopted by many authors. To further motivate the research presented in this dissertation, the following sections provide an overview over the modeling of inelastic incompressibility as well as shape memory alloy modeling in general.

1.2.1 Modeling the incompressibility constraint for plasticity

In many metals, one often assumes that the inelastic deformations are incompressible. In the case of elasto-plasticity, this is deduced from the observation that the inelastic deformations arise from dislocations moving through the crystal lattice. Therefore, only changing the positions of these dislocations should not yield volumetric inelastic strains. When modeling inelasticity in metals (e.g. elasto-plasticity or the shape memory effect), one often splits the deformation in an elastic and an inelastic part. In the case of geometrically linear theories, this split of the strain $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_e + \boldsymbol{\varepsilon}_p$ is done additively into an elastic, unconstrained strain $\boldsymbol{\varepsilon}_e$ and a deviatoric plastic strain $\boldsymbol{\varepsilon}_p$, which is straight-forward and yields reasonable results for rather small deformations (see, e.g., Wriggers [146]). Here, the constraint restricts the plastic strains to be purely deviatoric, i.e., the plastic strain $\text{tr}(\boldsymbol{\varepsilon}_p) = 0$ is without trace. This constraint is easily satisfied, as restricting the plastic strain rate $\dot{\boldsymbol{\varepsilon}}_p$ to be deviatoric automatically yields deviatoric plastic strains $\boldsymbol{\varepsilon}_p$.

However, when modeling large strains and rotations, the more complex geometrically nonlinear theory is preferred, since it is more accurate in this case (see, e.g., Fish and Shek [25]). Here, mainly two different approaches have been applied, which both, when applied to infinitesimal strains and rotations, fall back to the aforementioned geometrically linear theory. One approach (see, e.g., Nemat-Nasser [86]), is to additively split the rate of deformation $\boldsymbol{d} = \boldsymbol{d}_e + \boldsymbol{d}_p$ into an elastic part \boldsymbol{d}_e and a plastic contribution \boldsymbol{d}_p . The advantage of this approach is its simplicity, as the extension of the models originating in the infinitesimal strain regime can easily be extended to this geometrically nonlinear theory. However, it is only exact when applying it to small elastic strains, which is a major drawback of this approach.

The other mainly employed approach goes back to the works of Eckart [18], Kröner [57] and Lee [61], which makes use of a multiplicative split of the deformation gradient $\boldsymbol{F} = \boldsymbol{F}^e \boldsymbol{F}^p$.

This approach has two major advantages over the additive decomposition. First, applying the multiplicative split yields accurate results in the case of finite elastic strains. Additionally, when combining this approach with a Helmholtz free energy density formulation which assumes an additive split of contributing energies, one can calculate the stresses in dependence on the elastic deformation only, which is computationally attractive. Then again, formulating models employing the multiplicative split is not as straight-forward as formulating models employing the additive split. Specifically, extending a model from the geometrically linear theory to the multiplicative split often requires careful modeling choices.

In this multiplicative split kinematic setting, constraining the plastic deformation to be volume preserving involves prescribing the plastic part of the deformation gradient to be unimodular, i.e., the determinant of the plastic part of the deformation gradient $\det(\mathbf{F}^P)$ must be one. Therefore, the left Cauchy-Green tensor $\mathbf{C}^P = \mathbf{F}^{P\top} \mathbf{F}^P$ is constrained to be unimodular as well. In contrast to the infinitesimal strain theory, where one can prescribe the plastic strain rate to be deviatoric in order to retain a deviatoric plastic strain, it is far more complex to satisfy the volume constraint for the finite deformation theory. This problem was, to the best of the author's knowledge, acknowledged first in the work of Nagtegaal et al. [85].

To solve this open question, many different attempts have been made to satisfy the unimodularity constraint for finite deformations. In general, they can be divided into two sets of algorithms: general integration schemes, which preserve the constraint approximately, and geometric integrators, which exactly preserve the volumetric constraint when integrating evolution equations. A good overview over geometrical integrators and why this is of high importance for solving differential equations with constraints is given in, e.g., Hairer [35], and for the case of multiplicative plasticity in Shutov and Kreißig [117].

Many publications working with geometric integrators make use of the exponential map for tensors. For example¹, [2, 15, 22, 79, 96, 120, 139] are early papers which consider the exponential map from various perspectives to satisfy the unimodularity constraint. Additionally, it has to be noted that many geometrical integrators do not keep the symmetry of the system, i.e., the symmetry of the symmetric internal variables, per se. This was addressed in Dettmer and Reese [16], where an exponential map algorithm, which in fact keeps the symmetry of the system, was developed. In Vladimirov et al. [133] as well as Reese and Christ [101], this enhancement of the exponential map was developed further for modeling springback in von Mises plasticity and shape memory alloys, respectively. Overall, the exponential map is a very elegant solution to the theoretical problem in keeping the volumetric constraint exactly satisfied.

However, many of the aforementioned authors also recognize the difficulties related to the application of this approach. The main problem lies in applying this scheme in a finite element analysis, which is the aim of many developed schemes. Here, one has to solve and linearize

¹In the following, some publications consider multi-surface plasticity, and others von Mises plasticity. However, here we are mainly interested in the approaches to satisfy the unimodularity constraint, which are often applicable to both cases.

the solution of an eigenvalue problem and in every iteration in each integration point, which is numerically elaborate. Additionally, the implementation is not clear cut, but complex (see, e.g., Moler and van Loan [81] and Ortiz et al. [89]). Therefore, there has been research into methods that circumvent the classical exponential mapping approach.

For instance, Miehe [77] proposed an algorithmic post-processing update, which exploits the Neo-Hookean elastic stress response, and is thus able to project the solution back onto the unimodular manifold. Other promising attempts included Lagrange multipliers or penalty methods to satisfy the volume constraint. For example, the work of Lühns et al. [68] makes use of an additional incompatible configuration, which is modeled by an additional independent variable. This independent variable is then enforced to be equal to the compatible field by a Lagrange multiplier formulation.

Another class of approaches is based on a projection ansatz, which is adapted from hyperelasticity (see, e.g., Flory and Volkenstein [26] and Simo et al. [121]). Here, an algorithm, which projects any tensor back onto the unimodular manifold, is used to intrinsically satisfy the incompressibility constraint. For plasticity, this approach was adopted by Helm [40], where the projection is combined with an Euler-backward integration for von Mises plasticity with kinematic hardening. This approach was later improved numerically and compared to the exponential map in Shutov and Kreißig [116] (see also Shutov and Kreißig [117] for an error analysis of this approach). A seemingly independently developed approach for multi-surface plasticity is constructed in Hurtado et al. [43], where it is shown that this approach vastly increases the numerical performance in comparison to the exponential map. Finally, Shutov [115] proposed a closed form solution for this approach, which again increases the efficiency and reliability of this projection ansatz.

The approach adopted in this dissertation makes use of a similar projection ansatz. However, it is embedded into the framework of generalized standard materials (see, e.g., Halphen and Nguyen [36], Hackl [34] and Miehe [78]), which yields several advantages. Firstly, by including the projection into the potential, one gains more control over the potential when compared to the exponential mapping approaches. Secondly, by including a time-discrete potential that considers kinematic hardening, the formulation is easier to analyze and work with. This, for example, yields an improvement to the kinematic hardening energy adopted in Chapter 2. Thirdly, having a potential formulation, which itself contains the constraints, makes it such that solutions minimizing this potential automatically satisfy the constraints. Therefore, minimizing this potential is conceptually easier than solving the original problem with the corresponding set of constraints.

1.2.2 Shape memory alloy modeling

Since the discovery of shape memory alloys, the number of research articles on them have been drastically increasing over decades [91]. This large interest is mainly driven by the two unique effects of shape memory alloys, namely the shape memory effect and superelasticity.

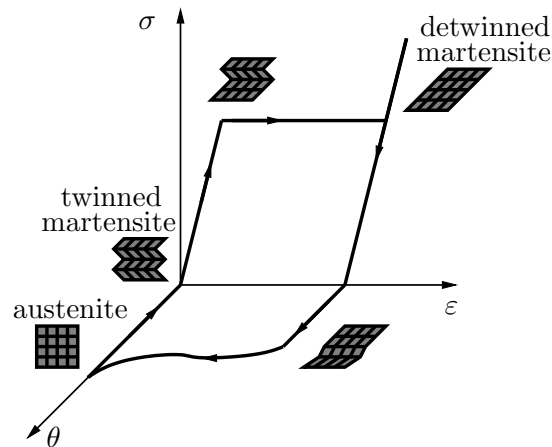


Fig. 1.1: Schematic stress-strain-temperature diagram showcasing the shape memory effect.

In general, the material's current temperature dictates which of these phenomena may occur. They are briefly summarized in the following section, after which an overview over the modeling of shape memory alloys is given.

Operating principles in shape memory alloys

The shape memory effect as well as superelasticity arise from the characteristic first order phase transformation between the monoclinic B19' martensite and face-centered cubic B2 austenite, which in contrast to, e.g., steel, can not only be caused by thermal, but also by mechanical loads in both directions. The experimental works by Otsuka and Waymann [93] as well as Otsuka and Ren [92] give a detailed overview over the topic of shape memory alloys in general, whereas the works of Zarinejad et al. [155], Karaca et al. [47], Karaca et al. [48] and Saghaian et al. [106] experimentally cover the shape memory alloy NiTiHf, which is the high temperature shape memory alloy modeled in this dissertation.

The aforementioned shape memory effect occurs when twinned martensite is loaded, which at sufficiently high loads starts to detwin into the preferentially-oriented phase (see Fig. 1.1). This detwinning process is assumed to be volume preserving, since reorienting the martensite variants should not change the volume that the martensite occupies.

After fully detwinning the martensite, it behaves elastic again. Then again, like in plastic materials, an inelastic strain remains upon unloading. However, when heating the shape memory alloy over the starting temperature of austenite transformation A_s and finally over the finish temperature of reverse transformation A_f , the shape memory alloy (SMA) transforms back to austenite and thus loses its inelastic strain. Upon cooling the SMA to the martensite transformation finish temperature M_f , twinned martensite is formed again, which returns the material to its initial state.

Hereby, the maximum recoverable strain is a very important feature, which differs for dif-

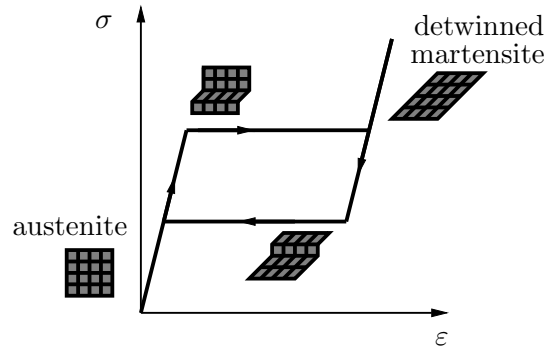


Fig. 1.2: Schematic stress-strain diagram showcasing superelasticity.

ferent alloys and compositions. The maximum recoverable strain for NiTi is about 10 %, which is due to its biocompatibility, resistance against corrosion, comparably good functional fatigue characteristics as well as ductility and stiffness [21, 91, 105] the most prominent shape memory alloy. In contrast, Sehitoglu et al. [111] show that the maximum recoverable strain in NiTiHf is almost 20 %. However, this comes at the cost of wider hysteresis (i.e., a larger difference in temperature is required to activate the shape memory effect), lower ductility at room temperature as well as worse cyclic stability (see, e.g., Noebe et al. [87] and Kockar et al. [52]).

The second commonly utilized effect in shape memory alloys is superelasticity. Here, the shape memory alloy is initially unloaded and in its austenite phase (see Fig. 1.2). After surpassing the elastic stage in the beginning, the material is increasingly transformed into detwinned martensite. This formation of favorable martensite variants allows to accommodate the prescribed deformation. In analogy to the behavior in the shape memory effect regime, after fully transforming the austenite into detwinned martensite, another elastic path occurs. Subsequently, when unloading before damage or plastic deformations occur, initially the material behaves elastic again. However, after the elastic unloading, a load driven phase transformation initiates, which transforms the detwinned martensite into its austenitic parent phase again. This reverts the strains accommodated by the first phase transformation. Finally, when the reverse transformation is finished, a final elastic unloading returns the material back into its initial state.

It has to be noted that both explanations of the effects and schematics are idealized and the underlying behavior is more complex. For example, many shape memory alloys transform into an additional, intermediate phase. For instance, NiTi has an intermediate rhombohedral crystal structure (see, e.g., Šittner et al. [122]), while NiTiHf exhibits an additional orthorhombic phase (see, e.g., Shen et al. [112]). While this is the case, due to their comparably small effect on the materials behavior, they are often neglected in many shape memory alloy models.

Overview of shape memory alloy modeling

While the effects leading to the shape memory alloy's unique behavior are well understood, the thermomechanical modeling of superelasticity as well as the shape memory effect is still an active and developing field of research. This originates in the severe nonlinearity of the shape memory alloy under complex thermomechanical loading, which is challenging to model accurately while staying computationally efficient. Due to the increasing number of applications, nowadays there exist many models that are tailored to a specific use case, where only a subset of the effects occurring in shape memory alloys are captured. For reviews on recent shape memory alloy models, the interested reader is referred to the works of Lester et al. [64] and Cisse et al. [14].

In general, the existing models for shape memory alloys can be roughly categorized into three classes of models: The first class comprises models based on statistical thermodynamics, which find the phase equilibrium by minimizing a three-well potential (see, e.g., Achenbach and Müller [1], Seelecke and Müller [110] or Govindjee et al. [32]). These models consider the microstructure of the material, which usually makes for great accuracy. However, it is therefore also challenging to use them in large scale structural simulations.

The second class is represented by the models based on micromechanics, which take into account the underlying mechanisms in single crystals. This category of models may then be extended to polycrystals via a utilization of homogenization schemes to obtain models for larger systems (see, among many, Patoor et al. [95], Lagoudas et al. [58], Frost et al. [28] and Stupkiewicz et al. [124]). Again, due to considering the underlying crystallographic phenomena, this class of models can describe the multitude of effects in shape memory alloys accurately, but also often comes with a high computational cost.

The last class of models, which is used in many structural simulations, is the class of phenomenological models. Since these models usually inherit only macroscopic parameters, the parameters can be obtained through comparatively simple experiments. In general, they can again be split into models considering the geometrically linear theory and models considering a geometrically nonlinear theory.

For small stretches and rotations, geometrically linear models are advantageous, since they usually have a lower computational cost, which arises from their lower geometrical complexity (see, e.g., the models in Auricchio et al. [7], Auricchio et al. [6] or Sedláč et al. [109]). However, when the stretches and rotations become large, a geometrically nonlinear theory is, due to its better accuracy, favorable (see, e.g., the models in Qidwai and Lagoudas [98], Müller and Bruhns [84], Reese and Christ [101] or Wang et al. [137]).

Lately, the modeling of functional fatigue, i.e., the loss in magnitude of the shape memory effect, shifting transformation stresses and temperatures as well as induced permanent strains, has gained a lot of attention in many publications. This allows for better model accuracy in cyclic loading and severe deformations. Many authors consider transformation induced plasticity (for a physical point of view, see Chowdhury and Sehitoglu [13]) as a key contribution

to the material's functional fatigue, which is then introduced into their models (see, among many, Hartl et al. [39], Xu et al. [148], Scalet et al. [108], Zhang et al. [157] and Woodworth et al. [145]).

Additionally, many of these models consider temperature as an additional field. However, to the best of the authors knowledge, there exists no model which considers finite strains, a thermomechanical coupling as well as thermal expansion and the volume change when undergoing phase transformation, which happens in some shape memory alloys (see, e.g., Potapov et al. [97]). However, all of these features are necessary when aiming to model the actuator concept published by Arivanandhan et al. [4]. This motivated the work on the shape memory alloy model shown and applied in Articles 2 and 3.

1.3 Outline of the dissertation

The present cumulative dissertation consists of three peer-reviewed journal articles and is structured as follows:

1.3.1 Publication 1

Chapter 2 is concerned with the development of a new projection scheme which accurately preserves the inelastic incompressibility. To investigate the details of this approach, it is implemented for plasticity instead of shape memory alloy models since plastic behavior is numerically and conceptually easier to model than the shape memory alloys complex behavior. In detail, a projection ansatz similar to the ones used by, e.g., Helm [40], Shutov and Kreißig [116] and Hurtado et al. [43], is embedded into the framework of generalized standard materials. Numerical singularities and problems are analyzed and resolved via additional contributions to the potentials. Furthermore, the symmetry of internal variables is intrinsically exactly preserved. Another key contribution is a new approach to approximate the tensor logarithm function, which employs the tensor invariants as additional variables to obtain better convergence with regard to the time step width. Additionally, the algorithmic tangent is derived analytically, which enables the efficient employment of the model into finite element schemes. Finally, the model is compared to results from the literature for isotropic and kinematic hardening.

1.3.2 Publication 2

Chapter 3 extends the small strain shape memory alloy model by Sedláč et al. [109] to the finite strain case using the projection ansatz developed in Chapter 2. This allows for a better prediction of shape memory alloys with high transformation strains. A key feature of this approach is, that it covers the shape memory effect as well as superelastic behavior depending on the temperature. Moreover, the model is enhanced to include thermal strains

due to phase transformation, which can occur in NiTiHf. Additionally, an ansatz for a mixed thermo-mechanical potential similar to Yang et al. [149] is presented to include the thermal problem into the generalized standard material formulation. Another essential contribution is the incorporation of a logarithmic strain, which ensures the disappearance of inelastic strains upon reverse transformation in the discretized form of the dissipation potential. Moreover, algorithmic considerations to resolve numerical instabilities as well as an active set search algorithm is presented, which increases the efficiency of the model. Finally, the model is applied to different boundary value problems.

1.3.3 Publication 3



Chapter 4 showcases a shape memory alloy actuator concept for bidirectional micro actuation. Here, special emphasis is on the analysis of thin film shape memory alloy properties on SiO_2 as well as Si substrates. Several experiments including X-ray diffraction, cantilever deflection measurements, electrical resistance measurements as well as tensile tests and scanning electron microscopy are conducted to show the influence of film thickness onto the shape memory alloy's key properties. Additionally, thermal fatigue tests show the fatigue characteristics, which are then optimized by varying the annealing process. Furthermore, the aforementioned experiments are used to identify the material parameters for the model presented in Chapter 3, which is extended for the hardening-like behavior found in shape memory alloy thin films. Moreover, this model is used to show the path towards novel bistable, bidirectional cooperative shape memory alloy micro- and nanoactuators by investigating favorable material and geometry parameters.

Finally, a conclusion and outlook into possible future extensions and open questions is given.

CHAPTER 2

Article 1: Numerical strategies for variational updates in large strain inelasticity with incompressibility constraint

This article was published as:

Sielenkämper, M. , Dittmann, J.  and Wulfinghoff, S. [2022], 'Numerical strategies for variational updates in large strain inelasticity with incompressibility constraint', *International Journal for Numerical Methods in Engineering* **123**(1), 245-267.

Own contributions to the following article:

- planning (large fraction)
- implementation of algorithm (large fraction)
- numerical simulations (large fraction)
- writing of the manuscript (large fraction)
- interpretation of the results (large fraction)

Abstract

In finite deformation inelasticity, one often has to deal with the incompressibility constraint. In the past, this was dealt with using, e.g., an exponential mapping approach, which yields exact volume preservation in plastic deformations. In this work however, the special-linear update approach by Hurtado et al. [43], which utilizes a projection method to fulfill the incompressibility constraint is used. The model is applied to isotropic plasticity by a novel approximation of the logarithm and treats kinematic hardening without losing the symmetry of the internal variables. The model results are compared to models utilizing an exponential mapping approach in numerical experiments.

Keywords: inelasticity, finite deformation plasticity, kinematic hardening, projection method, incompressibility constraint

2.1 Introduction

Nowadays, when modeling finite deformation plasticity or other inelastic deformations, the multiplicative split of the deformation gradient, i.e., $\mathbf{F} = \mathbf{F}^e \mathbf{F}^p$, introduced by Eckhart [18], Köner [57] and Lee [61] is widely accepted. Using this approach, one often has to deal with the incompressibility constraint, which is motivated by the observation that plastic deformations occur due to dislocation glide. In recent time, many finite deformation plasticity algorithms use an exponential mapping approach (see, e.g., [16, 76, 133, 139]) to cope with the incompressibility constraint. The advantage of using the exponential map approach is the intrinsic exact fulfillment of the incompressibility constraint. Moreover, the formulation is very neat and compact. However, it is not always obvious how to preserve the symmetry of the involved variables (see Dettmer and Reese [16]). Further, the exponential map comes with the disadvantage of having to perform a spectral decomposition, which is numerically elaborate. Additionally, while manageable, the linearization of the exponential map is not always straightforward (see, e.g., Ortiz et al. [89]). In this paper, the special-linear update by Hurtado et al. [43], which is an alternative to the exponential mapping approach is used. It also exactly satisfies the incompressibility constraint. For the kinematic hardening model, there are mainly two different types found in the literature. Some models are of the Chaboche-type, which makes use of evolution equations for the update of the back stress (see, e.g., [128, 151]). The other family of models makes use of a further multiplicative split of the plastic deformation gradient \mathbf{F}^p into an energetic and a dissipative part (see Lion [67] and Dettmer and Reese [16]), which is an extension of the Frederick-Armstrong model [5]. For the model presented in this paper, the latter is used.

This paper is structured as follows. First, the theory and numerical treatment of the model are discussed in Section 2.2. Here, the time-discrete counterpart for the isotropic hardening potential is given as well. In Section 2.3, results for the isotropic hardening model are shown.

Next, the model is extended for the incorporation of kinematic hardening in Section 2.4, where the results for kinematic hardening are also presented. Finally, a summary and outlook is given in Section 2.5.

2.2 Modeling of finite deformation plasticity

2.2.1 Kinematics

The finite deformation of a body can be described by using the gradient of deformation $\mathbf{x}(\mathbf{X}, t)$

$$\mathbf{F} = \text{Grad}(\mathbf{x}), \quad (2.1)$$

which relates line elements of a body in the reference configuration to the current configuration. Then, one often defines

$$J = \det(\mathbf{F}), \quad (2.2)$$

which describes the volumetric part of the deformation. In finite deformation plasticity, one often splits \mathbf{F} multiplicatively, i.e.,

$$\mathbf{F} = \mathbf{F}^e \mathbf{F}^p, \quad (2.3)$$

where \mathbf{F}^e is the elastic part and \mathbf{F}^p is the plastic part of the deformation gradient (see Kröner [57] and Lee [61]). To incorporate kinematic hardening, the additional multiplicative split of \mathbf{F}^p , employed by Lion [67] as well as Dettmer and Reese [16], is introduced as

$$\mathbf{F}^p = \mathbf{F}^{pe} \mathbf{F}^{pd}, \quad \det(\mathbf{F}^p) = 1, \quad \det(\mathbf{F}^{pe}) = \det(\mathbf{F}^{pd}) = 1, \quad (2.4)$$

where \mathbf{F}^{pe} is the energetic and \mathbf{F}^{pd} the dissipative part of \mathbf{F}^p . Note that \mathbf{F}^{pe} as well as \mathbf{F}^{pd} have to be both unimodular. Further, one commonly introduces the elastic left Cauchy-Green tensor \mathbf{b}^e and the plastic right Cauchy-Green tensor \mathbf{C}^p as

$$\begin{aligned} \mathbf{b}^e &= \mathbf{F}^e \mathbf{F}^{e\top} = \mathbf{F} \mathbf{C}^{p-1} \mathbf{F}^\top, \\ \mathbf{C}^p &= \mathbf{F}^{p\top} \mathbf{F}^p. \end{aligned} \quad (2.5)$$

Likewise, kinematic quantities for the dissipative and energetic parts are introduced as

$$\begin{aligned} \mathbf{C}^{pd} &= \mathbf{F}^{pd\top} \mathbf{F}^{pd}, \\ \mathbf{b}^{pe} &= \mathbf{F}^{pe} \mathbf{F}^{pe\top} = \mathbf{F}^p \mathbf{C}^{pd-1} \mathbf{F}^{p\top}. \end{aligned} \quad (2.6)$$

Moreover, one defines the plastic 'velocity gradient' \mathbf{L}^p as

$$\mathbf{L}^p = \dot{\mathbf{F}}^p \mathbf{F}^{p-1}, \quad (2.7)$$

where $\dot{\bullet}$ denotes a derivative with respect to time. Correspondingly, one defines the symmetric part of the plastic velocity gradient:

$$\mathbf{D}^p = \text{sym}(\mathbf{L}^p) = \frac{1}{2} \mathbf{F}^{p-\top} \dot{\mathbf{C}}^p \mathbf{F}^{p-1}. \quad (2.8)$$

2.2.2 Casting finite deformation plasticity into the generalized standard materials framework

To ensure thermodynamic consistency, the model is cast into the generalized standard materials (GSM) framework, which was first introduced by Halphen and Nguyen [36]. The Clausius-Duhem inequality for isothermal processes reads

$$\mathcal{D} = \boldsymbol{\tau} : \mathbf{d} - \dot{\psi} \geq 0, \quad (2.9)$$

where \mathcal{D} is the dissipation density, \mathbf{d} is the symmetric velocity gradient with respect to the current configuration and the Kirchhoff stress $\boldsymbol{\tau}$ is defined as $\boldsymbol{\tau} = \mathbf{F} \mathbf{S} \mathbf{F}^\top = J \boldsymbol{\sigma}$. Here, \mathbf{S} is the second Piola-Kirchhoff stress tensor and $\boldsymbol{\sigma}$ is the Cauchy stress tensor. Additionally, ψ is the Helmholtz free energy density, which is introduced as the sum of all energy storage functions and reads for this work

$$\psi = \psi_e(\mathbf{b}^e) + \psi_h(\alpha) + \psi_k(\mathbf{b}^{pe}), \quad (2.10)$$

where $\psi_e(\mathbf{b}^e)$ is the elastic energy density, ψ_h the stored isotropic hardening energy density due to isotropic hardening and $\psi_k(\mathbf{b}^{pe})$ the stored hardening energy due to kinematic hardening. Further, α is an isotropic internal hardening variable. For the elastic energy density, a Neo-Hookean elastic energy, which is only dependent on \mathbf{b}^e , in the form of

$$\psi_e(\mathbf{b}^e) = \frac{\lambda}{4} (J^{e2} - 1 - 2 \ln(J^e)) + \frac{\mu}{2} (\mathbf{I}_{\mathbf{b}^e} - 3 - 2 \ln(J^e)) \quad (2.11)$$

is chosen. Here, λ and μ are the Lamé parameters and $\mathbf{I}_{\mathbf{b}^e}$ is the first invariant of \mathbf{b}^e . For ψ_h , different choices, depending on the material that is to be modeled, can be made. However, we assume ψ_h to be an increasing function, i.e.,

$$\frac{\partial \psi_h}{\partial \alpha} \geq 0. \quad (2.12)$$

The kinematic hardening energy is assumed to be only dependent on $\mathbf{I}_{\mathbf{b}^{pe}}$, the first invariant of \mathbf{b}^{pe} , and reads

$$\begin{aligned} \psi_k(\mathbf{b}^{pe}) &= \frac{\mu^p}{2} \mathbf{I}_{\mathbf{b}^{pe}}, \\ \mathbf{I}_{\mathbf{b}^{pe}} &= \text{tr}(\mathbf{b}^{pe}) = \mathbf{C}^p : \mathbf{C}^{pd-1}. \end{aligned} \quad (2.13)$$

Further, to complete the requirements for the GSM-framework, the dissipation potential $\phi(\mathbf{D}^p, \dot{\alpha})$ is defined in similar fashion to Han and Reddy. [38] as

$$\phi(\mathbf{D}^p, \dot{\alpha}) = \begin{cases} \sqrt{\frac{2}{3}}\sigma_{y0}\|\mathbf{D}^p\| & \text{tr}(\mathbf{D}^p) = 0 \wedge \dot{\alpha} \geq \sqrt{\frac{2}{3}}\|\mathbf{D}^p\| \\ \infty & \text{else.} \end{cases} \quad (2.14)$$

Here, σ_{y0} is the initial yield stress. Note that the if-else construct ensures that $\text{tr}(\mathbf{D}^p) = 0$ is always true when minimizing the potential, and therefore guarantees volume preserving plastic deformations. Additionally, it ensures that the normalization condition $\dot{\alpha} \geq \sqrt{2/3}\|\mathbf{D}^p\|$ always holds. To obtain the yield criterion as well as the evolution of the internal variables, the dual potential $\phi^*(\boldsymbol{\Sigma}, q)$ is introduced as

$$\phi^*(\boldsymbol{\Sigma}, q) = \sup_{\mathbf{D}^p, \dot{\alpha}} (\boldsymbol{\Sigma} : \mathbf{D}^p + q\dot{\alpha} - \phi(\mathbf{D}^p, \dot{\alpha})) \quad (2.15)$$

$$= \sup_{\substack{\mathbf{D}^p = \mathbf{D}^{p'} \\ \dot{\alpha} \geq \sqrt{\frac{2}{3}}\|\mathbf{D}^p\|}} (\boldsymbol{\Sigma} : \mathbf{D}^p + q\dot{\alpha} - \sqrt{\frac{2}{3}}\sigma_{y0}\|\mathbf{D}^p\|), \quad (2.16)$$

utilizing a Legendre-Fenchel transformation. Here, $\boldsymbol{\Sigma}$ and q are the dual variables to \mathbf{D}^p and $\dot{\alpha}$, respectively. For $q \leq 0$, obviously

$$\dot{\alpha} = \sqrt{\frac{2}{3}}\|\mathbf{D}^p\| \quad (2.17)$$

always holds. Additionally, replacing $\boldsymbol{\Sigma}$ by its deviator $\boldsymbol{\Sigma}'$ drops the constraint $\mathbf{D}^p = \mathbf{D}^{p'}$, as follows

$$\begin{aligned} \phi^*(\boldsymbol{\Sigma}, q \leq 0) &= \sup_{\substack{\mathbf{D}^p \\ \dot{\alpha} \geq \sqrt{\frac{2}{3}}\|\mathbf{D}^p\|}} (\boldsymbol{\Sigma}' : \mathbf{D}^p - \sqrt{\frac{2}{3}}(\sigma_{y0} - q)\|\mathbf{D}^p\|) \\ &= \sup_{\gamma \geq 0} \sup_{\substack{\mathbf{N} \\ \|\mathbf{N}\|=1}} (\gamma \boldsymbol{\Sigma}' : \mathbf{N} - \sqrt{\frac{2}{3}}(\sigma_{y0} - q)\gamma) \\ &= \sup_{\gamma \geq 0} \gamma (\|\boldsymbol{\Sigma}'\| - \sqrt{\frac{2}{3}}(\sigma_{y0} - q)) \\ &= \sup_{\gamma \geq 0} \gamma f(\boldsymbol{\Sigma}, q \leq 0), \end{aligned} \quad (2.18)$$

where it was used that the supremum in \mathbf{N} is achieved for $\mathbf{N} \parallel \boldsymbol{\Sigma}'$, i.e.,

$$\mathbf{D}^p = \gamma \frac{\boldsymbol{\Sigma}'}{\|\boldsymbol{\Sigma}'\|}, \quad (2.19)$$

corresponding to the normality rule. Here, f is the yield criterion, which reads

$$f = \|\boldsymbol{\Sigma}'\| - \sqrt{\frac{2}{3}}(\sigma_{y0} - q). \quad (2.20)$$

Now, $\phi^*(\boldsymbol{\sigma}, q \leq 0)$ can be given in terms of f as

$$\phi^*(\boldsymbol{\sigma}, q \leq 0) = \begin{cases} 0 & f \leq 0 \\ \infty & f > 0. \end{cases} \quad (2.21)$$

Further, since ϕ and ϕ^* are convex, utilizing a Legendre-Fenchel transformation again, we obtain

$$\begin{aligned} \phi(\mathbf{D}^p, \dot{\alpha}) &= \sup_{\boldsymbol{\Sigma}, q \leq 0} (\boldsymbol{\sigma} : \mathbf{D}^p + q\dot{\alpha} - \phi^*(\boldsymbol{\Sigma}, q)) \\ &= \sup_{\boldsymbol{\Sigma}, q \leq 0} (\boldsymbol{\sigma} : \mathbf{D}^p + q\dot{\alpha} - \sup_{\boldsymbol{\Sigma}, q \leq 0} \gamma f). \end{aligned} \quad (2.22)$$

For $f > 0$, the term $-\sup \gamma f$ in Eq. (2.22) equals negative infinity. This definitely is not a solution of the first supremum problem in Eq. (2.22). Now, utilizing these findings, the Karush-Kuhn-Tucker conditions follow as

$$\gamma \geq 0, \quad f \leq 0, \quad \gamma f = 0. \quad (2.23)$$

With the dissipation potential ϕ and the free energy density ψ at hand, one now can formulate the rate potential as

$$\pi = \frac{\partial \psi}{\partial \mathbf{F}} : \dot{\mathbf{F}} + \frac{\partial \psi}{\partial \mathbf{C}^p} : \dot{\mathbf{C}}^p + \frac{\partial \psi}{\partial \alpha} \dot{\alpha} + \phi(\mathbf{D}^p, \dot{\alpha}), \quad (2.24)$$

where \mathbf{C}^p and α are the internal variables. Here, it can easily be verified that the second term of Eq. (2.24) can be reformulated to

$$\frac{\partial \psi}{\partial \mathbf{C}^p} : \dot{\mathbf{C}}^p = \underbrace{2\mathbf{F}^p \frac{\partial \psi}{\partial \mathbf{C}^p} \mathbf{F}^{pT}}_{=:-\boldsymbol{\Sigma}^{\text{eff}}} : \mathbf{D}^p. \quad (2.25)$$

Minimizing the rate potential from Eq. (2.24) in the internal variables while using the definition of $\boldsymbol{\Sigma}^{\text{eff}}$, one realizes that the minimization of π naturally involves the definition of the dual dissipation potential ϕ^* :

$$\begin{aligned} \inf_{\mathbf{D}^p, \dot{\alpha}} \pi &= \frac{\partial \psi}{\partial \mathbf{F}} : \dot{\mathbf{F}} - \sup_{\mathbf{D}^p, \dot{\alpha}} (\boldsymbol{\Sigma}^{\text{eff}} : \mathbf{D}^p - \underbrace{\frac{\partial \psi}{\partial \alpha} \dot{\alpha}}_{=:q_h \leq 0} - \phi(\mathbf{D}^p, \dot{\alpha})) \\ &= \frac{\partial \psi}{\partial \mathbf{F}} : \dot{\mathbf{F}} - \phi^*(\boldsymbol{\Sigma}^{\text{eff}}, q_h), \end{aligned} \quad (2.26)$$

where $q_h \leq 0$ follows from Eq. (2.12) and we assumed that $f(\Sigma^{\text{eff}}, q_h) \leq 0$. From Equations (2.17), (2.19) and (2.26) it follows that

$$\mathbf{D}^p = \gamma \frac{\Sigma^{\text{eff}'}}{\|\Sigma^{\text{eff}'}\|}, \quad \dot{\alpha} = \sqrt{\frac{2}{3}} \|\mathbf{D}^p\| = \sqrt{\frac{2}{3}} \gamma. \quad (2.27)$$

Additionally, note that

$$\Sigma^{\text{eff}} = \Sigma^e - \Sigma^b, \quad (2.28)$$

with the elastic Mandel stress tensor $\Sigma^e = \mathbf{C}^e \mathbf{S}^e$. Further, the elastic Second-Piola Kirchhoff stress \mathbf{S}^e and the Mandel back stress Σ^b are defined as

$$\mathbf{S}^e = 2 \frac{\partial \psi_e}{\partial \mathbf{C}^e}, \quad \Sigma^b = 2 \mathbf{b}^{\text{pe}} \frac{\partial \psi_k}{\partial \mathbf{b}^{\text{pe}}}. \quad (2.29)$$

With these definitions, one ends up with the yield criterion (compare with Equations (2.20) and (2.26)):

$$f = \|(\Sigma^e - \Sigma^b)'\| - \sqrt{\frac{2}{3}}(\sigma_{y0} - q) \leq 0. \quad (2.30)$$

For \mathbf{D}^{pd} , the Frederick-Armstrong hardening model assumes

$$\mathbf{D}^{\text{pd}} = \gamma \frac{b}{a} \Sigma^{b'}, \quad (2.31)$$

where a and b are material constants. The incorporation of kinematic hardening will be treated in Section 2.4.

2.2.3 Time discretization of the rate potential for isotropic hardening

This section only focuses of the handling of isotropic hardening, since the incorporation of kinematic hardening into the algorithmic counterpart of π needs some additional careful treatment, and is therefore done in Section 2.4. To solve the problem equations numerically, the time has to be discretized into time steps. This is done in a conventional manner, which is described in Appendix 2.6.

A possible transformation of the dissipation potential from Eq. (2.14) into the algorithmic counterpart ϕ_Δ reads

$$\phi_\Delta(\mathbf{C}^p, \alpha) = \begin{cases} \frac{1}{2} \sqrt{\frac{2}{3}} \sigma_{y0} \|\mathbf{c}^p - \mathbf{I}\| & \det(\mathbf{c}^p) = 1; \Delta\alpha \geq \sqrt{\frac{2}{3}} \frac{1}{2} \|\mathbf{c}^p - \mathbf{I}\| \\ \infty & \text{else,} \end{cases} \quad (2.32)$$

where \mathbf{c}^p can be interpreted as the incremental counterpart of \mathbf{C}^p , which is defined as

$$\mathbf{c}^p = \mathbf{U}_n^{p-1} \mathbf{C}^p \mathbf{U}_n^{p-1}. \quad (2.33)$$

Here, \mathbf{U}^p is the right plastic stretch tensor of the polar decomposition $\mathbf{F}^p = \mathbf{R}^p \mathbf{U}^p$, i.e., $\mathbf{U}^p = \sqrt{\mathbf{C}^p}$. The choice of using $1/2 \ln(\mathbf{C}^p)$ in the dissipation potential was introduced by Ortiz and Stainier [90] and $(\mathbf{C}^p - \mathbf{I})/2$ may be thought of as a linear approximation of the logarithm. For a unimodular \mathbf{C}^p and $\Delta t \rightarrow 0$, the algorithmic approximation becomes

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \phi_\Delta &= \sqrt{\frac{2}{3}} \sigma_{y0} \left\| \frac{1}{2} \mathbf{U}^{p-1} \dot{\mathbf{C}}^p \mathbf{U}^{p-1} \right\| \\ &= \sqrt{\frac{2}{3}} \sigma_{y0} \left\| \frac{1}{2} \mathbf{F}^{p-\top} \dot{\mathbf{C}}^p \mathbf{F}^{p-1} \right\| \\ &= \sqrt{\frac{2}{3}} \sigma_{y0} \|\mathbf{D}^p\|, \end{aligned} \quad (2.34)$$

which is obviously in line with Equations (2.64) and (2.14). Now, having obtained a consistent approximation for ϕ , one defines the algorithmic approximation π_Δ of π (compare Eq. (2.24)), which reads

$$\pi_\Delta = \psi - \psi_n + \phi_\Delta. \quad (2.35)$$

Further, the updates of internal variables may be obtained by solving the minimization problem

$$\inf_{\substack{\mathbf{C}^{p,\alpha} \\ \det(\mathbf{C}^p)=1}} = \pi(\mathbf{F}, \mathbf{C}^p, \alpha). \quad (2.36)$$

Similarly to the time-continuous case, the assumption of $q_h \geq 0$ (Eq. (2.12)) implies that the minimum is achieved for

$$\Delta\alpha = \sqrt{\frac{2}{3}} \, 1/2 \, \|\mathbf{C}^p - \mathbf{I}\|, \quad (2.37)$$

which we can directly insert in Eq. (2.36). Thus, we write for simplicity

$$\inf_{\substack{\mathbf{C}^p \\ \det(\mathbf{C}^p)=1}} \pi(\mathbf{F}, \mathbf{C}^p). \quad (2.38)$$

2.2.4 Treatment of the incompressibility constraint

In the sequel, the focus is on the minimization problems with incompressibility constraint, i.e.,

$$\inf_{\substack{\mathbf{C}^p \\ \det(\mathbf{C}^p)=1}} \pi_\Delta(\mathbf{F}, \mathbf{C}^p), \quad (2.39)$$

where we now employ the special-linear update from Hurtado et al. [43].

In order to obtain a constraint-free minimization problem, which is way easier to handle numerically, the unconstrained tensor $\hat{\mathbf{C}}^p$ is introduced via the nonlinear projection

$$\mathbf{C}^p(\hat{\mathbf{C}}^p) = (\det(\hat{\mathbf{C}}^p))^{-\frac{1}{3}} \hat{\mathbf{C}}^p. \quad (2.40)$$

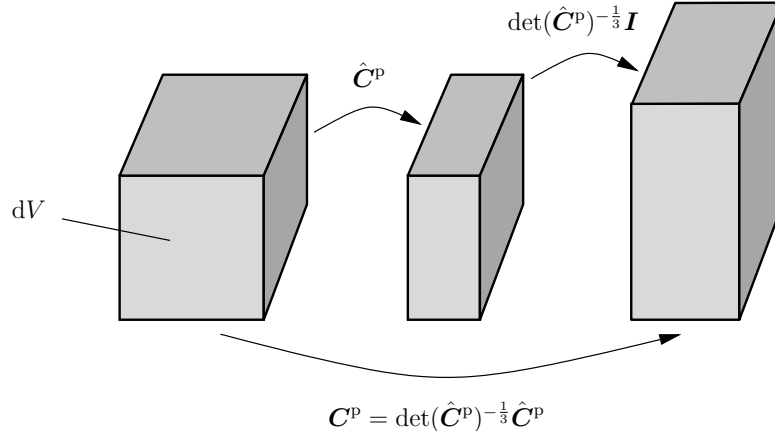


Fig. 2.1: Volume dV in the reference configuration is stretched by \hat{C}^p , followed by a volumetric stretch $\det(\hat{C}^p)^{-\frac{1}{3}} \mathbf{I}$, summarized by C^p .

This projection is depicted in Fig. 2.1, where one can see that the unconstrained tensor \hat{C}^p may change the shape and volume of the body. However, the term $\det(\hat{C}^p)^{-\frac{1}{3}} \mathbf{I}$ scales back any volumetric stretches that were included in \hat{C}^p . Therefore, this projection intrinsically ensures the unimodularity of C^p and allows to reformulate the minimization problem in Eq. (2.39) to

$$\inf_{\hat{C}^p} \pi_{\Delta}(\mathbf{F}, C^p(\hat{C}^p)). \quad (2.41)$$

Obviously, the constraint on C^p is gone. Unfortunately, the degree-1-homogeneous potential ϕ_{Δ} is not differentiable at $c^p = \mathbf{I}$, which makes the numerical solution difficult. For that reason, a further reparametrization of $\hat{c}^p = U_n^{p-1} \hat{C}^p U_n^{p-1}$ (compare Eq. (2.33)) is proposed as

$$c^p = U_n^{p-1} C^p U_n^{p-1} = \hat{\mathbb{M}}^{p-\frac{1}{3}} \hat{c}^p \quad (2.42)$$

$$\hat{c}^p = \mathbf{I} + 2\Delta\gamma \mathbf{N}^p \rightarrow \hat{C}^p \stackrel{(2.42)}{=} C_n^p + 2\Delta\gamma U_n^p \mathbf{N}^p U_n^p \quad (2.43)$$

where \mathbf{N}^p is a symmetric tensor with $\|\mathbf{N}^p\| = 1$. Employing this reparametrization into Eq. (2.32)₁ yields

$$\begin{aligned} \phi_{\Delta} &= \frac{1}{2} \sqrt{\frac{2}{3}} \sigma_{y0} \|c^p(\hat{c}^p) - \mathbf{I}\| \\ &= \frac{1}{2} \sqrt{\frac{2}{3}} \sigma_{y0} \|\hat{\mathbb{M}}^{p-\frac{1}{3}} (\mathbf{I} + 2\Delta\gamma \mathbf{N}^p) - \mathbf{I}\|. \end{aligned} \quad (2.44)$$

Inserting the reparametrization from Eq. (2.42) into Eq. (2.41) results in the minimization problem

$$\inf_{\substack{\Delta\gamma \geq 0, \mathbf{N}^p \\ \|\mathbf{N}^p\|=1}} \pi_{\Delta}(\mathbf{F}, C^p(U_n^{p-1}(\mathbf{I} + 2\Delta\gamma \mathbf{N}^p)U_n^{p-1})). \quad (2.45)$$

In order to eliminate the constraint $\|\mathbf{N}^p\| = 1$, one final reparametrization of \mathbf{N}^p is introduced as

$$\mathbf{N}^p = \frac{\mathbf{U}_n^{p-1} \tilde{\mathbf{N}}^s \mathbf{U}_n^{p-1}}{\|\mathbf{U}_n^{p-1} \tilde{\mathbf{N}}^s \mathbf{U}_n^{p-1}\|}, \quad (2.46)$$

where $\tilde{\mathbf{N}}^s$ is a symmetric but otherwise unconstrained tensor. Obviously, $\|\mathbf{N}^p\| = 1$ now intrinsically holds. Therefore, inserting this reparametrization into the minimization problem from Eq. (2.45) drops the constraint and yields

$$\inf_{\Delta\gamma \geq 0, \tilde{\mathbf{N}}^s} \pi_\Delta \left(\mathbf{F}, \mathbf{C}^p \left(\Delta\gamma, \tilde{\mathbf{N}}^s \right) \right). \quad (2.47)$$

Now, the minimization in Eq. (2.47) can be accomplished by solving the nonlinear equation set

$$\begin{aligned} \frac{\partial \pi_\Delta}{\partial \mathbf{z}} &= 0 \\ \Rightarrow \frac{\partial \pi_\Delta}{\partial \Delta\gamma} &= 0, \quad \frac{\partial \pi_\Delta}{\partial \tilde{\mathbf{N}}^s} = 0 \end{aligned} \quad (2.48)$$

with a Newton-scheme. Here, $\mathbf{z} = (\Delta\gamma, \tilde{\mathbf{N}}^s)$ is the set of parametrized internal variables. The entire first and second derivatives employed in the Newton-scheme are summarized in Appendix 2.8 and 2.9.

As the observant reader might have already seen, the system matrix $\frac{\partial^2 \pi_\Delta}{\partial \mathbf{z}^2}$ is singular. This is due to the minimization problem being independent of $\det(\hat{\mathbf{C}}^p)$ as well as $\|\tilde{\mathbf{N}}^s\|$. For $\det(\hat{\mathbf{C}}^p)$, this is depicted in Fig. 2.2, where two different $\hat{\mathbf{C}}_i^p$ with different determinants both yield the same tensor \mathbf{C}^p . Therefore, the minimization problem as it stands is not uniquely solvable. In addition, the algorithmic yield criterion is derived in Appendix 2.7.

2.2.5 Elimination of singularities

To eliminate the aforementioned singularities introduced through $\det(\hat{\mathbf{C}}^p)$ and $\|\mathbf{N}^p\|$, the regularization potential π_R is introduced as

$$\pi_R = \frac{1}{2} A ((\hat{\mathbf{I}}^p - 1)^2 + (\|\mathbf{U}_n^{p-1} \tilde{\mathbf{N}}^s \mathbf{U}_n^{p-1}\| - 1)^2), \quad (2.49)$$

where A is an arbitrary numerical constant. This energy can be interpreted as a spring attached to $\det(\hat{\mathbf{C}}^p)$ and $\|\mathbf{N}^p\|$, pulling on both to be equal to 1. Since π_Δ is otherwise free of $\det(\hat{\mathbf{C}}^p)$ and $\|\mathbf{N}^p\|$, adding this term to π_Δ ensures that in the converged solution state $\det(\hat{\mathbf{C}}^p) = 1$ as well as $\|\mathbf{N}^p\| = 1$ are *exactly* satisfied. Therefore, it is emphasized that the constant A has no influence on the solution for $\{\mathbf{C}^p, \alpha\}$, since $\pi_R = 0$ holds in the converged state.

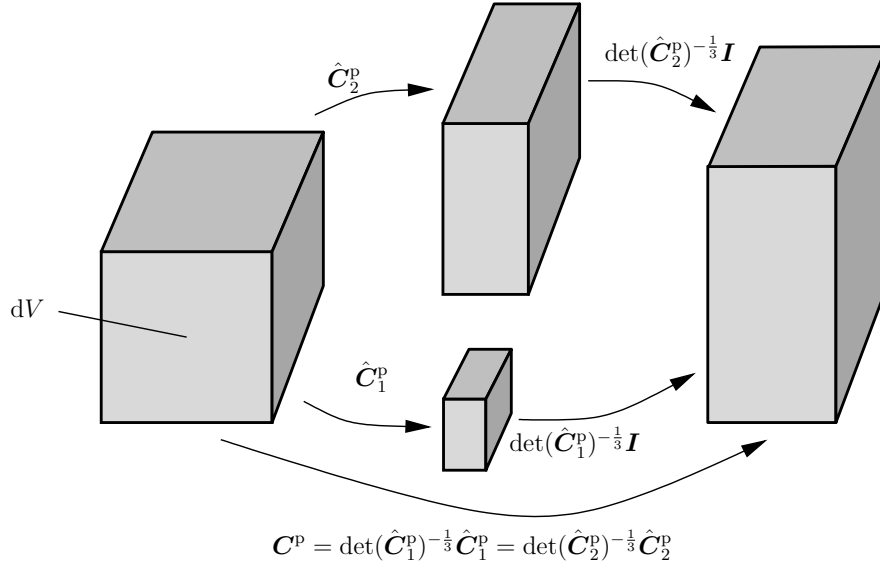


Fig. 2.2: Volume dV in the reference configuration is stretched by either $\det(\hat{C}_1^p)$ or $\det(\hat{C}_2^p)$, followed by a volumetric stretch $\det(\hat{C}_1^p)^{-\frac{1}{3}} \mathbf{I}$ and $\det(\hat{C}_2^p)^{-\frac{1}{3}} \mathbf{I}$, respectively. However, both yield the same C^p .

2.2.6 Improving the approximation of the logarithm

The approximation used in Eq. (2.32) is only first-order accurate. A higher accuracy can be achieved using the logarithm (see Subsection 2.2.3). This is depicted in Fig. 2.3, where the approximation of the logarithm is compared to the logarithm itself for the first two eigenvalues of Δc^p . Here, Δc^p is calculated as

$$\Delta c^p = c^p - \mathbf{I}. \quad (2.50)$$

Furthermore, the third eigenvalue is also accounted for using the constraint $\det(c^p) = 1$:

$$\rightarrow (1 + \lambda_1)(1 + \lambda_2)(1 + \lambda_3) = 1 \rightarrow \lambda_3 = \frac{1}{(1 + \lambda_1)(1 + \lambda_2)} - 1. \quad (2.51)$$

As one can see, the approximation is very close to the logarithm when Δc^p is small. However, as the eigenvalues of Δc^p get larger, which correspond to larger load-steps, the error drastically increases. Therefore, the approximation of the logarithm is improved using the first two invariants of Δc^p . For the improved approximation of the dissipation potential we use the ansatz

$$\phi_\Delta = \sqrt{\frac{2}{3}} \sigma_{y0} \left(\left\| \frac{1}{2} (c^p - \mathbf{I}) \right\| + a_1 I_{\Delta c^p} + a_2 II_{\Delta c^p} \right), \quad (2.52)$$

where a_1 and a_2 are constants. In the context of this work, they are optimized by a least-squares algorithm, minimizing the difference to the logarithmic approximation, as $a_1 = 3.0$ and $a_2 = 3.1$. For the least-squares algorithm, 21x21 equidistant support points over the range of $\lambda_1 \in [-0.25, 0.25]$ and $\lambda_2 \in [-0.25, 0.25]$ are used. Further, $I_{\Delta c^p}$ and $II_{\Delta c^p}$ are

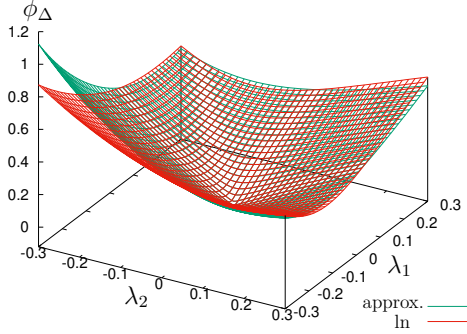


Fig. 2.3: Comparison of ϕ_Δ using the approximation (cf. Eq. (2.32)) and the logarithm.

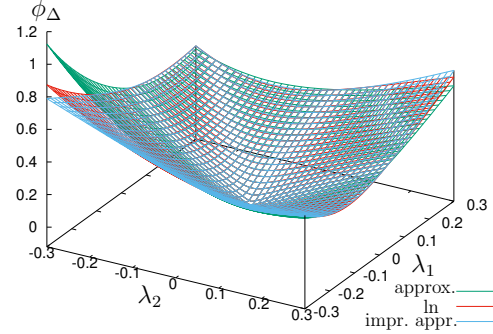


Fig. 2.4: Comparison of ϕ_Δ using the approximation (cf. Eq. (2.32)), the improved approximation (cf. Eq. (2.52)) and the logarithm.

the first and second invariants of $\Delta \mathbf{c}^p$, respectively. Here it is important to note, that this modification is time consistent, i.e., $\phi_\Delta \rightarrow 0$ for $\Delta t \rightarrow 0$. The improved approximation is compared to the old approximation and the logarithm in Fig. 2.4 for the eigenvalues of $\Delta \mathbf{c}^p$. Clearly the improved approximation is closer to the natural logarithm, even for larger eigenvalues of $\Delta \mathbf{c}^p$. Therefore, the improved approximation enables better results for larger time steps. This is shown in Subsection 2.3.3, where the results for the initial approximation are compared to the ones of the improved approximation.

2.2.7 Initial guesses for $\Delta \gamma$ and $\tilde{\mathbf{N}}^s$

In this work, a Newton-scheme is chosen to minimize the discretized potential. Therefore, good initial guesses for the unknowns are required to ensure that a solution can be found. For the initial guess of $\Delta \gamma$, the choice of $\Delta \gamma_0 = 0$ proves to be advantageous. Then, one can solve

$$\mathbf{N}_0^p = \arg \inf_{\substack{\mathbf{N}^p \\ \|\mathbf{N}^p\|=1}} \lim_{\Delta \gamma \rightarrow 0} \pi_\Delta \quad (2.53)$$

to obtain a good initial guess for $\tilde{\mathbf{N}}^s$ (compare Eq. (2.45)). This leads to

$$\mathbf{N}_0^p = \frac{\boldsymbol{\Sigma}_{\text{tr}}^{\text{eff}'} \mathbf{C}_n^p}{\left\| \boldsymbol{\Sigma}_{\text{tr}}^{\text{eff}'} \mathbf{C}_n^p \right\|}, \quad (2.54)$$

where $\boldsymbol{\Sigma}_{\text{tr}}^{\text{eff}'}$ is the trial value of $\boldsymbol{\Sigma}^{\text{eff}'}$ (see Eq. (2.28)), defined as

$$\boldsymbol{\Sigma}_{\text{tr}}^{\text{eff}'} = \mathbf{F}^{\text{etr}} \boldsymbol{\tau}^{\text{tr}} \mathbf{F}^{\text{etr}-1} - \boldsymbol{\Sigma}_n^{\text{b}'}. \quad (2.55)$$

Here, $\boldsymbol{\tau}^{\text{tr}}$ is the Kirchhoff stress evaluated with the trial value $\mathbf{b}^{\text{etr}} = \mathbf{F}^{\text{etr}\top} \mathbf{F}^{\text{etr}}$, where $\mathbf{F}^{\text{etr}} = \mathbf{F} \mathbf{F}_n^{\text{p}-1}$. With the initial guesses defined, the ingredients for the solution of the minimization problem are altogether.

2.2.8 Solving the local and global system of equations

The local minimization scheme is embedded in the global minimization problem, where one minimizes

$$\Pi_{\Delta} = \int_{\Omega_0} \pi_{\Delta} \, d\Omega - \int_{\partial\Omega_{0t}} \hat{\mathbf{t}} \cdot \mathbf{u} \, dS. \quad (2.56)$$

Here, $\hat{\mathbf{t}}$ is the traction vector acting on the surface $\partial\Omega_{0t} \subset \partial\Omega_0$ of the reference configuration Ω_0 and \mathbf{u} is the displacement vector. Additionally, body forces are neglected for simplicity. To minimize this global problem, the consistent tangent is required, which is given in Appendix 2.10. Additionally, the Newton scheme solving the local minimization problem is summarized in Algorithm 1.

Algorithm 1 Newton scheme solving the equation set in Eq. (2.48).

```

Compute trial variables  $\mathbf{F}^{\text{etr}}, \boldsymbol{\tau}^{\text{tr}}$ 
Compute yield criterion  $f$ 
if  $f^{\text{tr}} \geq 0$  then
     $\boldsymbol{\tau} = \boldsymbol{\tau}^{\text{tr}}, \mathbf{C}^{\text{p}} = \mathbf{C}_n^{\text{p}}, \alpha = \alpha_n$ 
    Compute algorithmic tangent  $\mathbb{C}^{\text{algo}}$ 
else
    while  $\left\| \frac{\partial \pi_{\Delta}}{\partial \mathbf{z}} \right\| \geq \text{tol}^{\text{NwtN}}$  do
        Compute  $\mathbf{N}^{\text{p}}, \hat{\mathbf{C}}^{\text{p}}, \mathbf{C}^{\text{p}}, \mathbf{b}^{\text{e}}, \Delta\alpha, \boldsymbol{\tau}, \mathbf{C}^{\text{pd}}$ 
        Compute Residual  $\frac{\partial \pi_{\Delta}}{\partial \mathbf{z}}$  and system matrix  $\frac{\partial^2 \pi_{\Delta}}{\partial \mathbf{z}^2}$ 
        Solve  $\frac{\partial^2 \pi_{\Delta}}{\partial \mathbf{z}^2} \Delta \mathbf{z} = -\frac{\partial \pi_{\Delta}}{\partial \mathbf{z}}$  for  $\Delta \mathbf{z}$ 
    end while
    Compute algorithmic tangent  $\mathbb{C}^{\text{algo}}$ 
end if

```

2.3 Numerical results for isotropic hardening

In this section, the presented model is tested for convergence with respect to step size Δt and compared to results from Simo [120].

2.3.1 Gauss point evaluations

To investigate the model implementation and convergence of the occurring stresses with regard to step size Δt , Gauss point evaluations are conducted. Therefore, the response of the material for a single Gauss-point in case of cyclic uniaxial tension and compression as well as

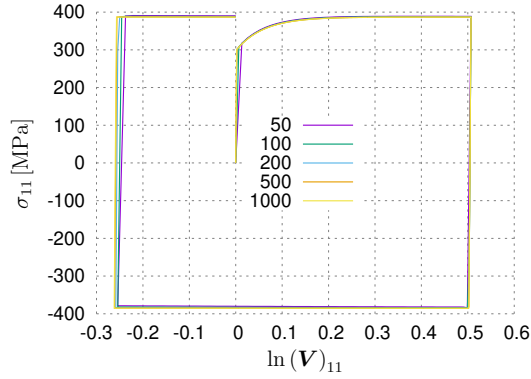


Fig. 2.5: Stress σ_{11} over strain $\ln(\mathbf{V})_{11}$ in a tensile test for different numbers of time steps.

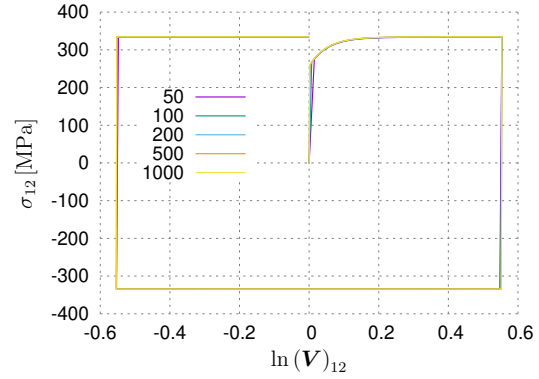


Fig. 2.6: Stress σ_{12} over strain $\ln(\mathbf{V})_{12}$ in a simple shear test for different numbers of time steps.

simple shear was investigated. The deformation was prescribed by incrementally increasing and decreasing \mathbf{F} while preserving the volume of the Gauss-point. Figures 2.5 and 2.6 show the comparison of stresses for different amounts of equidistant load steps for uniaxial tension and simple shear, respectively. Due to the lack of an analytical solution, the solutions can only be compared to a converged solution with many load steps. As one can see, the error in comparison to the converged solution obtained is very small. For tension and compression, even for only 50 steps, σ_{11} deviates by less than 0.5%. Here, it is noted that the largest difference lies in the elastic regime and only stems from the fact that the second evaluation is already far in the plastic regime. Further, for simple shear the error in the final shear stress is less than 1% for only 50 steps. Having evaluated the robustness and convergence behavior of the model, now a 3D problem is investigated.

2.3.2 3D necking of a circular bar

To validate the model results in 3D, the necking of a circular bar proposed by Simo [120] is compared to this model results.

A circular bar with radius $r_0 = 6.413\text{mm}$ of length $l = 53.334\text{mm}$ is subjected to a displacement controlled pure tension test using simple support boundary conditions. To prevent bifurcation, the cross-section is reduced in the middle of the bar, as described in [120]. Due to the obvious symmetries, only one eighth of the bar is analyzed, which is shown in Fig. 2.7. This boundary value problem as well as the additional 3D-Problem in Section 2.4 are solved using the finite element analysis program FEAP [127]. For this problem, the isotropic hardening energy ψ_h is chosen to be of the Voce-type as

$$\frac{\partial \psi_h}{\partial \alpha} = H\alpha + (\sigma_{y0} - \sigma_y)(1 - \exp -\beta\alpha), \quad (2.57)$$

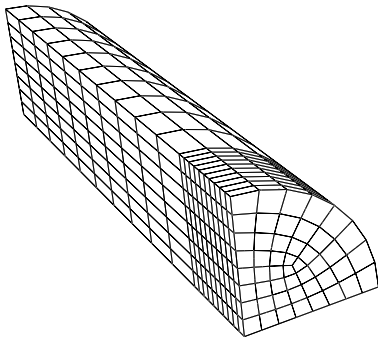


Fig. 2.7: Mesh of the circular bar with 960 Q1/P0 elements used in simulation.

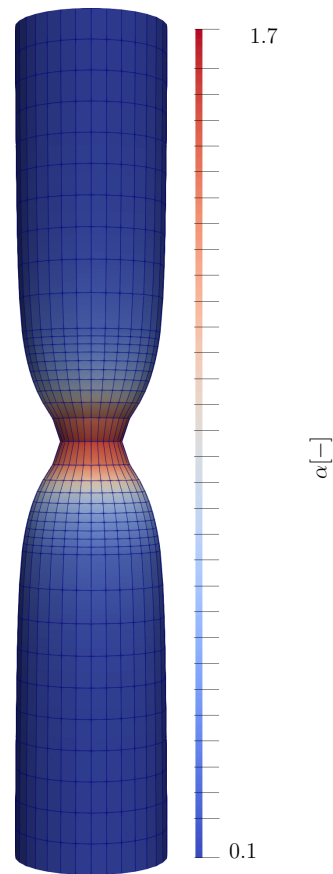


Fig. 2.8: Distribution of the equivalent plastic strain α over the deformed bar after loading.

Tab. 2.1: Material parameters for necking of a circular bar.

κ [MPa]	μ [MPa]	σ_y [MPa]	$\sigma_{y\infty}$ [MPa]	β	H [MPa]
164210	80193.8	450	715	16.93	129.24

Tab. 2.2: Numerical parameters for necking of a circular bar.

$\text{tol}^{\text{Nwt}} [-]$	A [MPa]	a_1 [MPa]	a_2 [MPa]
10^{-8}	10^4	3.0	3.1

like in Simo [120]. Likewise, the elastic and isotropic hardening material parameters were chosen accordingly, which are given in Table 2.1. Further, the numerical parameters are aggregated in Table 2.2. Here, tol^{Nwt} is the tolerance of the local Newton scheme presented in Algorithm 1. The deformed bar is shown in Fig 2.8. Now, the influence of the number of equidistant time steps is evaluated. The necking displacement is plotted against the elongation for a varying number of time steps in Fig. 2.9. Clearly, the final necking displacement converges to a value of roughly 3.8mm.

2.3.3 Comparison of the approximations of the logarithm

In this section, the results of the original first-order approximation of ϕ_Δ (Eq. (2.32)) and its improved counterpart (Eq. (2.52)), introduced in Section 2.2.6, are compared. To illustrate the difference in the convergence behavior, the same problem as in the previous section is run again with the approximation omitting the invariants. The results are shown in Fig. 2.10. While the results, in comparison to the ones in the previous section, clearly converge towards the same displacement, the results for smaller numbers of time steps clearly improved. Therefore it can be concluded that the improved approximation utilizing the invariances improved the model results by a large margin for not fully converged results with regard to time steps.

2.4 Incorporation of kinematic hardening

2.4.1 Expanding the potential for kinematic hardening

To expand the time discrete counterpart π_Δ for kinematic hardening, ψ_k is replaced by its algorithmic counterpart $\tilde{\psi}_k$, which is defined as

$$\tilde{\psi}_k = \mu^p \Delta \mathbf{C}^p : \mathbf{C}_{n+\frac{1}{2}}^{pd-1}. \quad (2.58)$$

Here, in comparison to ψ_k from Eq. (2.13), $\Delta \mathbf{C}^p$ is used instead of \mathbf{C}^p , because using \mathbf{C}^p leads to $\tilde{\psi}_k$ being not consistent with the time-continuous case. Additionally, it turned out

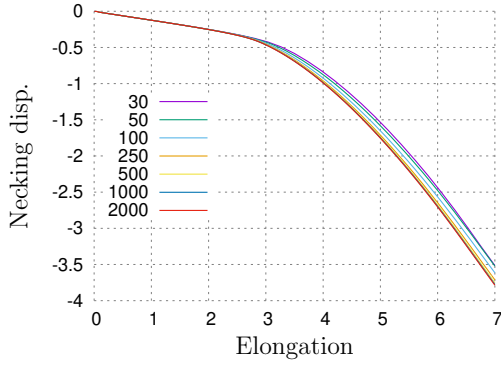


Fig. 2.9: Necking displacement versus elongation for varying amounts of time steps using the improved logarithms approximation.

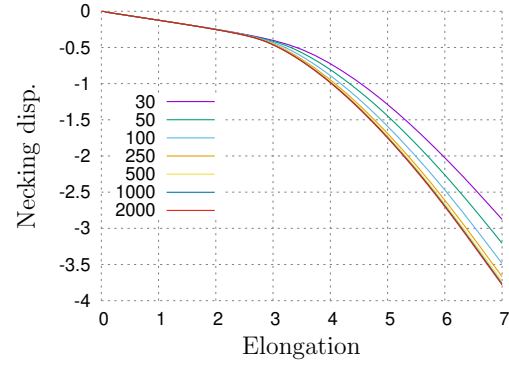


Fig. 2.10: Necking displacement versus elongation for varying amounts of time steps using the logarithm approximation without the invariants.

that using the midpoint evaluation

$$\mathbf{C}_{n+\frac{1}{2}}^{\text{pd}-1} = \frac{1}{2} \left(\mathbf{C}_n^{\text{pd}-1} + \mathbf{C}^{\text{pd}-1} \right) \quad (2.59)$$

instead of $\mathbf{C}^{\text{pd}-1}$ leads to more accurate results. Further, in line with Eq. (2.40), \mathbf{C}^{pd} is reparametrized to be unimodular:

$$\mathbf{C}^{\text{pd}} = (\det(\hat{\mathbf{C}}^{\text{pd}}))^{-\frac{1}{3}} \hat{\mathbf{C}}^{\text{pd}}. \quad (2.60)$$

Then again, since we assume a Frederick-Armstrong hardening law, $\hat{\mathbf{C}}^{\text{pd}}$ is defined depending on the plastic multiplier γ as

$$\hat{\mathbf{C}}^{\text{pd}} = \mathbf{C}_n^{\text{pd}} + 2\Delta\gamma \frac{b}{c} \Sigma_n^{b'}. \quad (2.61)$$

Incorporating $\tilde{\psi}_k$ into the discretized potential π_Δ yields

$$\pi_\Delta = \psi_e + \psi_h + \tilde{\psi}_k - (\psi_{e_n} + \psi_{h_n} + \tilde{\psi}_{k_n}) + \phi_\Delta. \quad (2.62)$$

In the following subsection, this potential is minimized with respect to $\Delta\gamma$ and $\tilde{\mathbf{N}}^s$, again.

2.4.2 Results for kinematic hardening

To validate the results of the kinematic hardening extension, the model results were compared to the results in Vladimirov et al. [133]. Therefore, the material parameters in Table 2.3 are adapted, which also partially originate from Lührs et al. [68]. Likewise, the isotropic hardening energy adopted from Vladimirov et al. [133] is of Voce-type and reads

$$\psi_h = Q \left(\alpha + \frac{\exp - \beta \alpha}{\beta} \right). \quad (2.63)$$

Tab. 2.3: Parameters for kinematic hardening from Lührs et al. [68] and Vladimirov et al. [133].

μ [MPa]	λ [MPa]	σ_y [MPa]	Q [MPa]	β	μ^p [MPa]	c [MPa]	b
80000	119999.67	300	400	2.5	950	1900	8.5

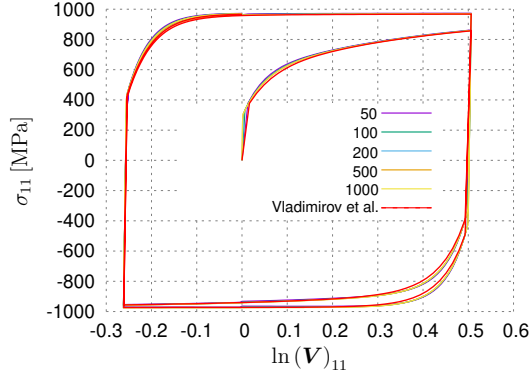


Fig. 2.11: Stress σ_{11} over strain $\ln(\mathbf{V})_{11}$ in a tensile test for different numbers of time steps in comparison to Vladimirov et al. [133].

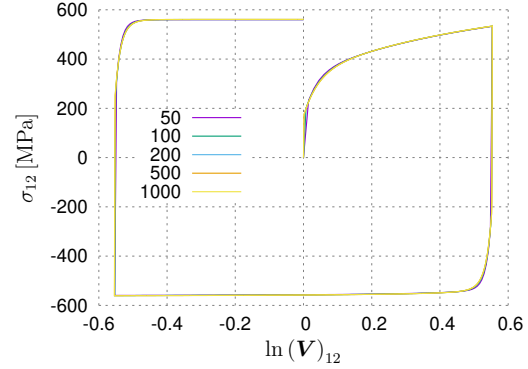


Fig. 2.12: Stress σ_{12} over strain $\ln(\mathbf{V})_{12}$ in a simple shear test for different numbers of time steps.

First, the stress-strain Gauss-point computations in tension-compression as well as simple shear are shown in Figures 2.11 and 2.12. Again, as in the isotropic hardening section, one can clearly see the results converging for an increasing number of time steps per load cycle. Further, since there is not an analytical solution, the results are compared to the results from Vladimirov et al. [133]. Here, one can see, that the error introduced by only using 50 time steps is very small in both tests. Additionally, the Bauschinger effect, which changes the yield stress depending on the direction of load, is clearly visible in both tests. Further, the results for the shear test compare well with Vladimirov et al. [133].

Next, the 3D finite element example from Vladimirov et al. [133] was used to validate the results of the model. A cube with a side length of 1mm is loaded by a combined tensile and shear loading. At the bottom of the cube, a boundary condition holds all degrees of freedom, while on the top a displacement boundary condition is prescribed in two steps for the displacements u_1 and u_2 as

$$\begin{aligned} \text{Step 1: } u_1 &= 1\text{mm}, & u_2 &= 0.5\text{mm}, \\ \text{Step 2: } u_1 &= -1\text{mm}, & u_2 &= 0.5\text{mm}. \end{aligned}$$

The cube is discretized by 4 different regular meshes with 8, 64, 512, 1728 and 4096 Q1P0 elements. Due to the obvious symmetries, only one quarter of the cube has to be modeled. The cube is shown in an almost undeformed state, after load *step 1* and after load *step 2* in Figure 2.13. Obviously, these large deformations are uncharacteristic for a metal cube.

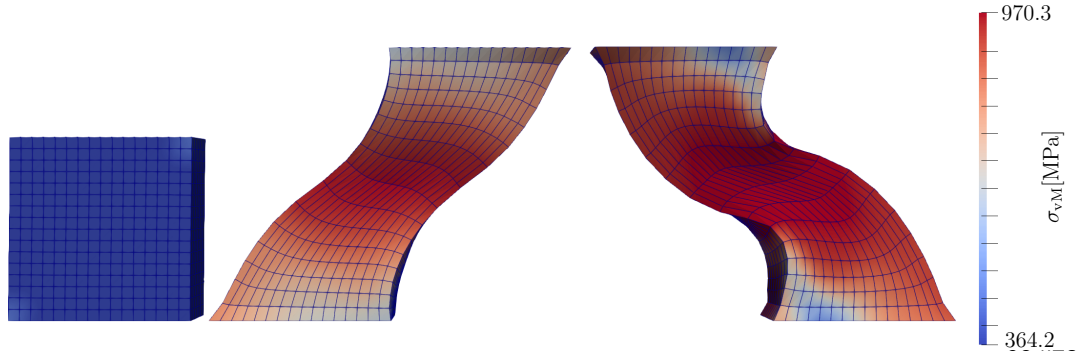


Fig. 2.13: Deformed cube from left to right: Almost undeformed, after load *step 1* and after load *step 2*.

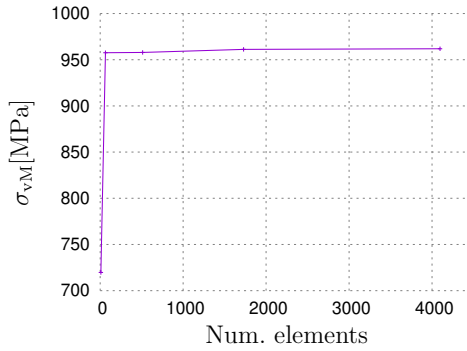


Fig. 2.14: Convergence of the stress in the middle of the cube after *step 1* with regards to the amount of finite elements.

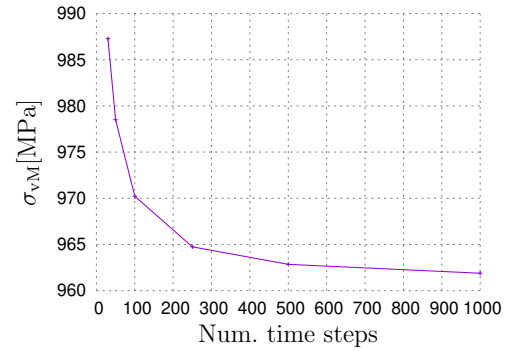


Fig. 2.15: Convergence of the stress in the middle of the cube after *step 1* with regards to the number of time steps.

However, as stated in Vladimirov et al. [133], it serves well as a test showing the model's capabilities in the finite strain regime. The convergence of the stresses in the middle of the cube after load *step 1* with regard to the amount of time steps is shown in Figure 2.14. Here, the finest discretization with 4096 finite elements is used. Clearly, one can see the stresses quickly converging towards 962 N/mm² for the finer meshes. Only for the mesh with just 8 elements the results are deviating far from the converged solution.

Next, the convergence of the stresses in the middle of the cube after load *step 1* of the cube with regard to the cube discretization is investigated. Therefore, the variously fine discretized meshes are all loaded in 1000 steps. The resulting stresses are shown in Figure 2.15. Again, the values converge for decreasing time steps. However, in contrast to Vladimirov et al. [133], the results converge strictly from above towards their final value. Overall, while being quite satisfying, the time convergence behavior of this model is not as good as that of the exponential mapping approach.

2.5 Summary and outlook

In this paper, a numerical approach for the solution of large deformation elastoplasticity is presented. In contrast to many recent works, it does not use the exponential map to fulfill the incompressibility constraint. The model is formulated in the generalized standard materials framework to ensure thermodynamic consistency. Further, the model was extended to kinematic hardening in Section 2.4. The model's capability of producing reasonable results in both isotropic and kinematic hardening is shown through numerical experiments. The time convergence behavior for isotropic hardening was improved using an ansatz involving the invariants of \mathbf{c}^P . This improvement is enabled by having the possibility for a flexible modification of the time discretized potentials, which opens the door for further improvements like the control of the convexity of the potential or potential-based line search algorithms. In the future, it should be investigated how to improve the accuracy for large time steps for kinematic hardening. In this case, the proposed model is still clearly inferior to the exponential map. However, it comes with the advantage of inherently having symmetric internal variables for kinematic hardening without any modifications to the approach. This is not the case for the exponential map, where it proved to be a tedious task, as can be seen from the publication history on that topic (e.g., Dettmer and Reese [16], Vladimirov et al. [133]). For these reasons, it remains very interesting to also employ this model in more complicated inelastic models, e.g., shape memory alloys, where the usage of the exponential map may also be involved.

Acknowledgment

Financial support of subproject A3 *Cooperative Actuator Systems for Nanomechanics and Nanophotonics: Coupled Simulation* of the Priority Programme SPP 2206 by the German Research Foundation (DFG) (Grant WU847/3-1) is gratefully acknowledged by M.S. and S.W.. Funding by the German Research Foundation (DFG) for the project P12 within the framework of the research training group GRK 2154 "Materials for Brain: Thin film functional materials for minimally invasive therapy of brain diseases" (Grant GRK 2154/1) is gratefully acknowledged by S.W. and J.D..

Data availability

Data, specifically the code used to generate the results, is openly available in a public repository that issues datasets with DOIs.

Appendix

2.6 Appendix A: Consistency of the dissipation potential

To solve the boundary value problem, one wants to find an algorithmic counterpart to ϕ , which is ϕ_Δ , where the rate-type quantities \mathbf{D}^p and $\dot{\alpha}$ are expressed in terms of the increments of the internal variables in a time step from t_n to t_{n+1} . The algorithmic approximation of ϕ must be consistent, i.e.,

$$\phi = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \phi_\Delta. \quad (2.64)$$

2.7 Appendix B: Algorithmic yield criterion

In order to decide whether a time step is elastic or plastic, we consider the trial state $\pi_\Delta(\mathbf{F}, \mathbf{C}_n^p)$. Then, we apply an infinitesimal increment

$$d\mathbf{C}^p = d\gamma \mathbf{U}_n^p \mathbf{N}^p \mathbf{U}_n^p \quad (d\gamma \geq 0) \quad (2.65)$$

and test if

$$\pi_\Delta(\mathbf{F}, \mathbf{C}_n^p + d\mathbf{C}^p) < \pi_\Delta(\mathbf{F}, \mathbf{C}_n^p). \quad (2.66)$$

If this is true for any deviatoric direction \mathbf{N}^p , the step is plastic, otherwise it is elastic. Equation (2.66) is equivalent to $f \leq 0$, i.e., it is sufficient to check the usual yield criterion. The proof is in analogy to the time-continuous case.

2.8 Appendix C: First derivatives of π_Δ

To minimize the discretized potential, one needs the first and second derivatives of the discretized potential with respect to the internal variables. To obtain a more compact notation, \mathbf{N}^s is introduced as

$$\mathbf{N}^s = \mathbf{C}_n^p \mathbf{N}^p. \quad (2.67)$$

Additionally, \mathbb{I}^s is the 4th order identity on symmetric matrices. Further, the total yield stress σ_y is given by

$$\sigma_y = \sigma_{y0} + \frac{\partial \psi_h}{\partial \alpha}. \quad (2.68)$$

Furthermore, the symmetrizing box product is defined via

$$\mathbf{A} \overset{s}{\square} \mathbf{B} : \mathbf{C} = \mathbf{A}_{\text{sym}}(\mathbf{C}) \mathbf{B}. \quad (2.69)$$

The first derivatives of the discretized potential w.r.t. the internal variables are given by

$$\begin{aligned}
 d\pi_\Delta = & \boldsymbol{\tau} : d\mathbf{d} \\
 & + \underbrace{\left(\frac{\partial(\psi_e + \tilde{\psi}_k)}{\partial \hat{\mathbf{C}}^p} : 2\mathbf{N}^s + \frac{\partial \tilde{\psi}_k}{\partial \hat{\mathbf{C}}^{pd}} : 2\frac{b}{a} \boldsymbol{\Sigma}_n^{b'} \mathbf{C}_n^{pd} + \sigma_y \frac{\partial \Delta \alpha}{\partial \Delta \gamma} + \frac{\partial \pi_R}{\partial \hat{\mathbb{M}}^p} \hat{\mathbb{M}}^p \hat{\mathbf{C}}^{p-1} : 2\mathbf{N}^s \right)}_{\frac{\partial \pi_\Delta}{\partial \Delta \gamma}} d\Delta \gamma \\
 & + \underbrace{\left(\mathbb{P}_{N^s}^\top : \left(2\Delta \gamma \frac{\partial(\psi_e + \tilde{\psi}_k)}{\partial \hat{\mathbf{C}}^p} + \sigma_y \frac{\partial \Delta \alpha}{\partial \mathbf{N}^s} + 2\Delta \gamma \frac{\partial \pi_R}{\partial \hat{\mathbb{M}}^p} \hat{\mathbb{M}}^p \hat{\mathbf{C}}^{p-1} \right) + \frac{\partial \pi_R}{\partial \tilde{\mathbf{N}}^s} \right)}_{\frac{\partial \pi_\Delta}{\partial \tilde{\mathbf{N}}^s}} : d\tilde{\mathbf{N}}^s.
 \end{aligned} \tag{2.70}$$

Here, the projector \mathbb{P}_{N^s} is defined as

$$\mathbb{P}_{N^s} = \frac{\partial \mathbf{N}^s}{\partial \tilde{\mathbf{N}}^s} = \frac{1}{\left\| \mathbf{U}_n^{p-1} \tilde{\mathbf{N}}^s \mathbf{U}_n^{p-1} \right\|} \left(\mathbb{I}^s - \mathbf{N}^s \otimes \mathbf{C}_n^{p-1} \mathbf{N}^s \mathbf{C}_n^{p-1} \right). \tag{2.71}$$

Likewise, the projectors \mathbb{P}_p and \mathbb{P}_{pd} are defined as

$$\mathbb{P}_p = \frac{\partial \mathbf{C}^p}{\partial \hat{\mathbf{C}}^p} = \hat{\mathbb{M}}^{p-\frac{1}{3}} \left(\mathbb{I}^s - \frac{1}{3} \hat{\mathbf{C}}^p \otimes \hat{\mathbf{C}}^{p-1} \right), \tag{2.72}$$

$$\mathbb{P}_{pd} = \frac{\partial \mathbf{C}^{pd}}{\partial \hat{\mathbf{C}}^{pd}} = \hat{\mathbb{M}}^{pd-\frac{1}{3}} \left(\mathbb{I}^s - \frac{1}{3} \hat{\mathbf{C}}^{pd} \otimes \hat{\mathbf{C}}^{pd-1} \right). \tag{2.73}$$

The other occurring derivatives are given as

$$\frac{\partial(\psi_e + \tilde{\psi}_k)}{\partial \hat{\mathbf{C}}^p} = \mathbb{P}_p^\top : \left(-\mathbf{C}^{p-1} \mathbf{F}^\top \frac{\partial \psi_e}{\partial \mathbf{b}^e} \mathbf{F} \mathbf{C}^{p-1} + \frac{\mu^p}{2} \left(\mathbf{C}_n^{pd-1} + \mathbf{C}^{pd-1} \right) \right) \tag{2.74}$$

$$\frac{\partial \psi_e}{\partial \mathbf{b}^e} = \left(\frac{\lambda}{4} (\det(\mathbf{b}^e) - 1) - \frac{\mu}{2} \right) \mathbf{b}^{e-1} + \frac{\mu}{2} \mathbf{I} \tag{2.75}$$

$$\frac{\partial \tilde{\psi}_k}{\partial \hat{\mathbf{C}}^{pd}} = -\mathbb{P}_{pd}^\top : \frac{1}{2} \mu^p \mathbf{C}^{pd-1} \Delta \mathbf{C}^p \mathbf{C}^{pd-1} \tag{2.76}$$

$$\begin{aligned}
 \mathbf{M}_{C^p} &:= \frac{\partial (\|\Delta \mathbf{C}^p\| + a_1 \mathbb{I}_{\Delta \mathbf{C}^p} + a_2 \mathbb{II}_{\Delta \mathbf{C}^p})}{\partial \mathbf{C}^p} \\
 &= \frac{\mathbf{C}_n^{p-1} \Delta \mathbf{C}^p \mathbf{C}_n^{p-1}}{\left\| \mathbf{U}_n^{p-1} \Delta \mathbf{C}^p \mathbf{U}_n^{p-1} \right\|} \\
 &\quad + a_1 \mathbf{C}_n^{p-1} + a_2 \left((\Delta \mathbf{C}^p : \mathbf{C}_n^{p-1}) \mathbf{C}_n^{p-1} - \mathbf{C}_n^{p-1} \Delta \mathbf{C}^p \mathbf{C}_n^{p-1} \right)
 \end{aligned} \tag{2.77}$$

$$\frac{\partial \Delta \alpha}{\partial \Delta \gamma} = \sqrt{\frac{2}{3}} \frac{1}{2} \mathbf{M}_{C^p} : \mathbb{P}_p : 2\mathbf{N}^s \tag{2.78}$$

$$\frac{\partial \pi_R}{\partial \hat{\mathbb{M}}^p} = A(\hat{\mathbb{M}}^p - 1) \tag{2.79}$$

$$\frac{\partial \Delta \alpha}{\partial \mathbf{N}^s} = \sqrt{\frac{2}{3}} 2\Delta \gamma \mathbb{P}_p^\top : \frac{1}{2} \mathbf{M}_{C^p} \tag{2.80}$$

$$\frac{\partial \pi_R}{\partial \tilde{\mathbf{N}}^s} = A(\|U_n^{p-1} \tilde{\mathbf{N}}^s U_n^{p-1}\| - 1) \left(\frac{C_n^{p-1} \tilde{\mathbf{N}}^s C_n^{p-1}}{\|U_n^{p-1} \tilde{\mathbf{N}}^s U_n^{p-1}\|} \right) \quad (2.81)$$

2.9 Appendix D: Second derivatives of π_Δ

The second derivatives of the discretized potential w.r.t. the internal variables are given by

$$\begin{aligned} \frac{\partial^2 \pi_\Delta}{\partial \Delta \gamma^2} = & 2\mathbf{N}^s : \frac{\partial^2(\psi_e + \tilde{\psi}_k)}{\partial \hat{\mathbf{C}}^{p^2}} : 2\mathbf{N}^s + 2\mathbf{N}^s : \frac{\partial^2 \pi_R}{\partial \hat{\mathbf{C}}^{p^2}} : 2\mathbf{N}^s \\ & + (2\frac{b}{a} \Sigma_n^{b'} C_n^{pd}) : \frac{\partial^2 \tilde{\psi}_k}{\partial \hat{\mathbf{C}}^{pd^2}} : (2\frac{b}{a} \Sigma_n^{b'} C_n^{pd}) + \sigma_y \frac{\partial^2 \Delta \alpha}{\partial \Delta \gamma^2} \\ & + 2 \cdot 2\mathbf{N}^s : \frac{\partial^2 \tilde{\psi}_k}{\partial \hat{\mathbf{C}}^p \partial \hat{\mathbf{C}}^{pd}} : (2\frac{b}{a} \Sigma_n^{b'} C_n^{pd}) + \frac{\partial^2 \psi_h}{\partial \alpha^2} \left(\frac{\partial \Delta \alpha}{\partial \Delta \gamma} \right)^2 \end{aligned} \quad (2.82)$$

$$\begin{aligned} \frac{\partial^2 \pi_\Delta}{\partial \Delta \gamma \partial \tilde{\mathbf{N}}^s} = & \mathbb{P}_{N^s}^\top : \left(2\frac{\partial(\psi_e + \tilde{\psi}_k)}{\partial \hat{\mathbf{C}}^p} + 2\Delta \gamma \frac{\partial^2(\psi_e + \tilde{\psi}_k)}{\partial \hat{\mathbf{C}}^{p^2}} : 2\mathbf{N}^s \right. \\ & + \sigma_y \frac{\partial^2 \Delta \alpha}{\partial \Delta \gamma \partial \tilde{\mathbf{N}}^s} + 2\frac{\partial \pi_R}{\partial \hat{\mathbf{C}}^p} + 2\Delta \gamma \frac{\partial^2 \pi_R}{\partial \hat{\mathbf{C}}^{p^2}} : 2\mathbf{N}^s \\ & \left. + 2\Delta \gamma \frac{\partial^2 \tilde{\psi}_k}{\partial \hat{\mathbf{C}}^p \partial \hat{\mathbf{C}}^{pd}} : (2\frac{b}{a} \Sigma_n^{b'} C_n^{pd}) + \frac{\partial^2 \psi_h}{\partial \alpha^2} \frac{\partial \Delta \alpha}{\partial \tilde{\mathbf{N}}^s} \frac{\partial \Delta \alpha}{\partial \Delta \gamma} \right), \end{aligned} \quad (2.83)$$

$$\begin{aligned} \frac{\partial^2 \pi_\Delta}{\partial \tilde{\mathbf{N}}^{s^2}} = & \mathbb{D}_{N^s}(\mathbf{A}) + \frac{\partial^2 \pi_R}{\partial \tilde{\mathbf{N}}^{s^2}} \\ & + \mathbb{P}_{N^s}^\top : \left((2\Delta \gamma)^2 \frac{\partial^2(\psi_e + \tilde{\psi}_k)}{\partial \hat{\mathbf{C}}^{p^2}} + \frac{\partial^2 \psi_h}{\partial \alpha^2} \frac{\partial \Delta \alpha}{\partial \tilde{\mathbf{N}}^s} \otimes \frac{\partial \Delta \alpha}{\partial \tilde{\mathbf{N}}^s} \right. \\ & \left. + \sigma_y \frac{\partial^2 \Delta \alpha}{\partial \tilde{\mathbf{N}}^{s^2}} + (2\Delta \gamma)^2 \frac{\partial^2 \pi_R}{\partial \hat{\mathbf{C}}^{p^2}} \right) : \mathbb{P}_{N^s} \end{aligned} \quad (2.84)$$

Here, \mathbb{D}_{N^s} is defined via

$$d(\mathbb{P}_{N^s}^\top : \mathbf{B}) = \mathbb{D}_{N^s}(\mathbf{B}) : d\tilde{\mathbf{N}}^s + \mathbb{P}_{N^s}^\top : d\mathbf{B} \quad (2.85)$$

and reads

$$\begin{aligned} \mathbb{D}_{N^s}(\mathbf{B}) = & -\frac{1}{\|U_n^{p-1} \tilde{\mathbf{N}}^s U_n^{p-1}\|^3} \left(\mathbf{f}(\mathbf{B}) \otimes C_n^{p-1} \tilde{\mathbf{N}}^s C_n^{p-1} \right. \\ & + C_n^{p-1} \tilde{\mathbf{N}}^s C_n^{p-1} \otimes \mathbf{f}(\mathbf{B}) \\ & - \frac{1}{\|U_n^{p-1} \tilde{\mathbf{N}}^s U_n^{p-1}\|^2} (\tilde{\mathbf{N}}^s : \mathbf{B}) C_n^{p-1} \tilde{\mathbf{N}}^s C_n^{p-1} \otimes C_n^{p-1} \tilde{\mathbf{N}}^s C_n^{p-1} \\ & \left. + (\tilde{\mathbf{N}}^s : \mathbf{B}) C_n^{p-1} \square C_n^{p-1} \right). \end{aligned} \quad (2.86)$$

where \mathbf{B} is an arbitrary second order tensor and $\mathbf{f}(\mathbf{B})$ is defined as

$$\mathbf{f}(\mathbf{B}) = \mathbf{B} - \frac{1}{\left\| \mathbf{U}_n^{p-1} \tilde{\mathbf{N}}^s \mathbf{U}_n^{p-1} \right\|^2} (\tilde{\mathbf{N}}^s : \mathbf{B}) \mathbf{C}_n^{p-1} \tilde{\mathbf{N}}^s \mathbf{C}_n^{p-1}. \quad (2.87)$$

Likewise, one can define \mathbb{D}_i and \mathbb{D}_{pd} via

$$d(\mathbb{P}_p^\top : \mathbf{B}) = \mathbb{D}_i(\mathbf{B}) : d\hat{\mathbf{C}}^p + \mathbb{P}_p^\top : d\mathbf{B} \quad (2.88)$$

$$\begin{aligned} \Rightarrow \mathbb{D}_i(\mathbf{B}) = & -\frac{1}{3} \hat{\mathbb{I}}^{p-\frac{1}{3}} \left(\left(\mathbf{B} - \frac{1}{3} (\hat{\mathbf{C}}^p : \mathbf{B}) \hat{\mathbf{C}}^{p-1} \right) \otimes \hat{\mathbf{C}}^{p-1} \right. \\ & + \hat{\mathbf{C}}^{p-1} \otimes \left(\mathbf{B} - \frac{1}{3} (\hat{\mathbf{C}}^p : \mathbf{B}) \hat{\mathbf{C}}^{p-1} \right) \\ & \left. - (\hat{\mathbf{C}}^p : \mathbf{B}) \left(\hat{\mathbf{C}}^{p-1} \hat{\square} \hat{\mathbf{C}}^{p-1} - \frac{1}{3} \hat{\mathbf{C}}^{p-1} \otimes \hat{\mathbf{C}}^{p-1} \right) \right) \end{aligned} \quad (2.89)$$

$$d(\mathbb{P}_{pd}^\top : \mathbf{B}) = \mathbb{D}_{pd}(\mathbf{B}) : d\hat{\mathbf{C}}^{pd} + \mathbb{P}_{pd}^\top : d\mathbf{B} \quad (2.90)$$

$$\Rightarrow \mathbb{D}_{pd}(\mathbf{B}) = \dots \quad (2.91)$$

Here, \mathbb{D}_{pd} is worked out in analogy to \mathbb{D}_i , which means just the indexes are changing. The remaining occurring derivatives are given as

$$\frac{\partial^2(\psi_e + \tilde{\psi}_k)}{\partial \hat{\mathbf{C}}^{p2}} = \mathbb{D}_i \left(\frac{\partial(\psi_e + \tilde{\psi}_k)}{\partial \mathbf{C}^p} \right) + \mathbb{P}_p^\top : \frac{\partial^2 \psi_e}{\partial \mathbf{C}^{p2}} : \mathbb{P}_p \quad (2.92)$$

$$\begin{aligned} \frac{\partial^2 \psi_e}{\partial \mathbf{C}^{p2}} = & \mathbf{C}^{p-1} \mathbf{F}^\top \hat{\square} \mathbf{F} \mathbf{C}^{p-1} : \frac{\partial^2 \psi_e}{\partial \mathbf{b}^{e2}} : \mathbf{C}^{p-1} \mathbf{F}^\top \hat{\square} \mathbf{F} \mathbf{C}^{p-1} \\ & - \mathbf{C}^{p-1} \hat{\square} \frac{\partial \psi_e}{\partial \mathbf{C}^p} - \frac{\partial \psi_e}{\partial \mathbf{C}^p} \hat{\square} \mathbf{C}^{p-1} \end{aligned} \quad (2.93)$$

$$\begin{aligned} \frac{\partial^2 \tilde{\psi}_k}{\partial \hat{\mathbf{C}}^{pd2}} = & \mathbb{D}_{pd} \left(\frac{\partial \tilde{\psi}_k}{\partial \mathbf{C}^{pd}} \right) \\ & + \mathbb{P}_{pd}^\top : \left(\mathbf{C}^{pd-1} \hat{\square} \frac{\partial \tilde{\psi}_k}{\partial \mathbf{C}^{pd}} + \frac{\partial \tilde{\psi}_k}{\partial \mathbf{C}^{pd}} \hat{\square} \mathbf{C}^{pd-1} \right) : \mathbb{P}_{pd} \end{aligned} \quad (2.94)$$

$$\frac{\partial^2 \tilde{\psi}_k}{\partial \hat{\mathbf{C}}^p \partial \hat{\mathbf{C}}^{pd}} = -\frac{\mu^p}{2} \mathbb{P}_p^\top : \left(\mathbf{C}^{pd-1} \hat{\square} \mathbf{C}^{pd-1} \right) : \mathbb{P}_{pd} \quad (2.95)$$

$$\frac{\partial^2 \pi_R}{\partial \hat{\mathbf{C}}^{p2}} = \left(\frac{\partial^2 \pi_R}{\partial \hat{\mathbb{I}}^{p2}} \hat{\mathbb{I}}^p + \frac{\partial \pi_R}{\partial \hat{\mathbb{I}}^p} \right) \hat{\mathbb{I}}^p \hat{\mathbf{C}}^{p-1} \otimes \hat{\mathbf{C}}^{p-1} - \frac{\partial \pi_R}{\partial \hat{\mathbb{I}}^p} \hat{\mathbb{I}}^p \hat{\mathbf{C}}^{p-1} \hat{\square} \hat{\mathbf{C}}^{p-1} \quad (2.96)$$

$$\frac{\partial^2 \Delta \alpha}{\partial \Delta \gamma^2} = \sqrt{\frac{2}{3}} 2N^s : \left(\mathbb{D}_i \left(\frac{1}{2} \mathbf{M}_{C^p} \right) + \mathbb{P}_p^\top : \frac{1}{2} \frac{\partial \mathbf{M}_{C^p}}{\partial \mathbf{C}^p} : \mathbb{P}_p \right) : 2N^s \quad (2.97)$$

$$\begin{aligned} \frac{\partial^2 \Delta \alpha}{\partial \Delta \gamma \partial \tilde{\mathbf{N}}^s} &= \sqrt{\frac{2}{3}} 2 \mathbb{P}_{N^s}^\top : \mathbb{P}_p^\top : \frac{1}{2} \mathbf{M}_{C^p} \\ &+ 2 \Delta \gamma \mathbb{P}_{N^s}^\top : \left(\mathbb{D}_i \left(\frac{1}{2} \mathbf{M}_{C^p} \right) + \mathbb{P}_p^\top : \frac{1}{2} \frac{\partial \mathbf{M}_{C^p}}{\partial \mathbf{C}^p} : \mathbb{P}_p \right) : 2 \mathbf{N}^s \end{aligned} \quad (2.98)$$

$$\begin{aligned} \frac{\partial^2 \Delta \alpha}{\partial \tilde{\mathbf{N}}^{s2}} &= \sqrt{\frac{2}{3}} 2 \Delta \gamma \mathbb{D}_{N^s} \left(\mathbb{P}_p^\top : \frac{1}{2} \mathbf{M}_{C^p} \right) \\ &+ (2 \Delta \gamma)^2 \mathbb{P}_{N^s}^\top : \left(\mathbb{D}_i \left(\frac{1}{2} \mathbf{M}_{C^p} \right) + \mathbb{P}_p^\top : \frac{1}{2} \frac{\partial \mathbf{M}_{C^p}}{\partial \mathbf{C}^p} : \mathbb{P}_p \right) \mathbb{P}_{N^s} \end{aligned} \quad (2.99)$$

$$\begin{aligned} \frac{\partial^2 \pi_R}{\partial \tilde{\mathbf{N}}^{s2}} &= A \left(\mathbf{C}_n^{p-1} \overset{s}{\square} \mathbf{C}_n^{p-1} \left(1 - \frac{1}{\| \mathbf{U}_n^{p-1} \tilde{\mathbf{N}}^s \mathbf{U}_n^{p-1} \|} \right) \right. \\ &\quad \left. + \frac{1}{\| \mathbf{U}_n^{p-1} \tilde{\mathbf{N}}^s \mathbf{U}_n^{p-1} \|} \mathbf{C}_n^{p-1} \mathbf{N}^s \mathbf{C}_n^{p-1} \otimes \mathbf{C}_n^{p-1} \mathbf{N}^s \mathbf{C}_n^{p-1} \right) \end{aligned} \quad (2.100)$$

$$\begin{aligned} \frac{\partial \mathbf{M}_{C^p}}{\partial \mathbf{C}^p} &= \frac{1}{\| \mathbf{U}_n^{p-1} \Delta \mathbf{C}^p \mathbf{U}_n^{p-1} \|} \left(\mathbf{C}_n^{p-1} \overset{s}{\square} \mathbf{C}_n^{p-1} - \mathbf{M}_{C^p} \otimes \mathbf{M}_{C^p} \right) \\ &+ a_2 \left(\mathbf{C}_n^{p-1} \otimes \mathbf{C}_n^{p-1} - \mathbf{C}_n^{p-1} \overset{s}{\square} \mathbf{C}_n^{p-1} \right) \end{aligned} \quad (2.101)$$

2.10 Appendix E: Consistent tangent

When solving boundary value problems using the finite element method, we make use of the consistent tangent operator $\underline{C}^{\text{algo}}$. This is derived in this model as follows. We start with the virtual work of the internal forces

$$\int_{V_0} \boldsymbol{\tau} : \mathbf{d}_\delta \, dV, \quad (2.102)$$

where \mathbf{d}_δ is defined as

$$\mathbf{d}_\delta = \text{sym} \left(\underbrace{\text{Grad}(\delta \mathbf{u}) \mathbf{F}^{-1}}_{\mathbf{l}_\delta} \right). \quad (2.103)$$

Now, calculating the differential, we get

$$\int_{V_0} d(\boldsymbol{\tau} : \mathbf{d}_\delta) \, dV. \quad (2.104)$$

Further, expressing the differential in terms of \mathbf{S} and \mathbf{F} , we get

$$d(\boldsymbol{\tau} : \mathbf{d}_\delta) = d(\boldsymbol{\tau} : \mathbf{l}_\delta) = d(\mathbf{F} \mathbf{S}) : \delta \mathbf{F}, \quad (2.105)$$

where the symmetry of $\boldsymbol{\tau}$ was exploited. Now, the differential is calculated to be

$$\begin{aligned} d(\mathbf{FS}) : \delta \mathbf{F} &= (d\mathbf{FS}) : \delta \mathbf{F} \\ &= (d\mathbf{F}\mathbf{F}^{-1}\mathbf{F}\mathbf{S}\mathbf{F}^{\top}(\delta\mathbf{F}\mathbf{F}^{-1})^{\top}) : \mathbf{I} + (\mathbf{F}d\mathbf{S}\mathbf{F}^{\top}) : \mathbf{l}_{\delta} \\ &= (\mathbf{l}_d\boldsymbol{\tau}\mathbf{l}_{\delta}) : \mathbf{I} + (\mathbb{C}^{\text{algo}} : \mathbf{d}_d) : \mathbf{l}_{\delta}. \end{aligned} \quad (2.106)$$

Here, the first term is the so-called geometric tangent and the second term arises from the elastic and inelastic contributions to the materials stiffness. Additionally, \mathbf{l}_d and \mathbf{d}_d are defined in analogy to \mathbf{l}_{δ} and \mathbf{d}_{δ} . Further, $\mathbb{C}^{\text{algo}} : \mathbf{d}_d$ can be split into the elastic and inelastic part:

$$\begin{aligned} \mathbb{C}^{\text{algo}} : \mathbf{d}_d &= \left(\mathbf{F} \overset{\text{s}}{\square} \mathbf{F}^{\top} : \frac{\partial \mathbf{S}}{\partial \mathbf{E}} : \mathbf{F}^{\top} \overset{\text{s}}{\square} \mathbf{F} \right) : \mathbf{d}_d + \frac{\partial(\mathbf{F}\mathbf{S}\mathbf{F}^{\top})}{\partial \mathbf{z}} : \frac{\partial \mathbf{z}}{\partial \mathbf{E}} : d\mathbf{E} \\ &= \mathbb{C}^e : \mathbf{d}_d + \frac{\partial \boldsymbol{\tau}}{\partial \mathbf{z}} \frac{\partial \mathbf{z}}{\partial \mathbf{E}} : \mathbf{F}^{\top} \mathbf{d}_d \mathbf{F}. \end{aligned} \quad (2.107)$$

Now, to evaluate $\partial \mathbf{z} / \partial \mathbf{E}$, we have a look at the local newton scheme, where

$$\frac{\partial \pi_{\Delta}}{\partial \mathbf{z}} = 0 \quad (2.108)$$

must hold in the converged state. Therefore, calculating the differential of Eq. (2.108), we obtain

$$\begin{aligned} 0 &= d\left(\frac{\partial \pi_{\Delta}}{\partial \mathbf{z}}\right) \\ &= \frac{\partial^2 \pi_{\Delta}}{\partial \mathbf{z}^2} d\mathbf{z} + \frac{\partial^2 \pi_{\Delta}}{\partial \mathbf{z} \partial \mathbf{E}} : d\mathbf{E} = 0 \\ &= \frac{\partial^2 \pi_{\Delta}}{\partial \mathbf{z}^2} d\mathbf{z} + \left(\frac{\partial \mathbf{S}}{\partial \mathbf{z}}\right)^{\top} : d\mathbf{E} = 0. \end{aligned} \quad (2.109)$$

Now, rearranging Eq. (2.109), we get

$$\begin{aligned} d\mathbf{z} &= - \left(\frac{\partial^2 \pi_{\Delta}}{\partial \mathbf{z}^2} \right)^{-1} \frac{\partial^2 \pi_{\Delta}}{\partial \mathbf{z} \partial \mathbf{E}} : d\mathbf{E} \\ &= - \frac{\partial \mathbf{z}}{\partial \mathbf{E}} : d\mathbf{E} \\ &= - \left(\frac{\partial^2 \pi_{\Delta}}{\partial \mathbf{z}^2} \right)^{-1} \left(\frac{\partial \mathbf{S}}{\partial \mathbf{z}} \right)^{\top} : \mathbf{F}^{\top} \overset{\text{s}}{\square} \mathbf{F} : \mathbf{d}_d \\ &= - \left(\frac{\partial^2 \pi_{\Delta}}{\partial \mathbf{z}^2} \right)^{-1} \left(\frac{\partial \boldsymbol{\tau}}{\partial \mathbf{z}} \right)^{\top} : \mathbf{d}_d. \end{aligned} \quad (2.110)$$

Inserting the relation for $\partial \mathbf{z}/\partial \mathbf{E}$ into Eq. (2.107), we finally obtain the algorithmic tangent:

$$\mathbb{C}^{\text{algo}} = \mathbb{C}^{\text{e}} - \frac{\partial \boldsymbol{\tau}}{\partial \mathbf{z}} \left(\frac{\partial^2 \pi_{\Delta}}{\partial \mathbf{z}^2} \right)^{-1} \left(\frac{\partial \boldsymbol{\tau}}{\partial \mathbf{z}} \right)^{\top}. \quad (2.111)$$

Here, the term $\partial \boldsymbol{\tau}/\partial \mathbf{z}$ is calculated via

$$\frac{\partial \boldsymbol{\tau}}{\partial \mathbf{z}} = \begin{pmatrix} \frac{\partial \boldsymbol{\tau}}{\partial \Delta \gamma} \\ \frac{\partial \boldsymbol{\tau}}{\partial \tilde{\mathbf{N}}^{\text{s}}} \end{pmatrix}, \quad (2.112)$$

where the occurring derivatives of $\boldsymbol{\tau}$ are given by

$$\frac{\partial \boldsymbol{\tau}}{\partial \Delta \gamma} = 2 \frac{\partial \boldsymbol{\tau}}{\partial \hat{\mathbf{C}}^{\text{p}}} : \mathbf{N}^{\text{s}}, \quad (2.113)$$


$$\frac{\partial \boldsymbol{\tau}}{\partial \tilde{\mathbf{N}}^{\text{s}}} = 2 \Delta \gamma \frac{\partial \boldsymbol{\tau}}{\partial \hat{\mathbf{C}}^{\text{p}}} : \mathbb{P}_{\mathbf{N}^{\text{s}}}^{\top}, \quad (2.114)$$

$$\begin{aligned} \frac{\partial \boldsymbol{\tau}}{\partial \hat{\mathbf{C}}^{\text{p}}} = & -2 \left(\mathbf{I} \stackrel{\text{s}}{\square} \frac{\partial \psi_{\text{e}}}{\partial \mathbf{b}^{\text{e}}} + \frac{\partial \psi_{\text{e}}}{\partial \mathbf{b}^{\text{e}}} \stackrel{\text{s}}{\square} \mathbf{I} + \left(\mathbf{I} \stackrel{\text{s}}{\square} \mathbf{b}^{\text{e}} + \mathbf{b}^{\text{e}} \stackrel{\text{s}}{\square} \mathbf{I} \right) : \frac{\partial^2 \psi_{\text{e}}}{\partial \mathbf{b}^{\text{e}2}} \right) \\ & : \mathbf{F} \stackrel{\text{s}}{\square} \mathbf{F}^{\top} : \mathbf{C}^{\text{p}-1} \stackrel{\text{s}}{\square} \mathbf{C}^{\text{p}-1} : \mathbb{P}_{\text{p}}. \end{aligned} \quad (2.115)$$

CHAPTER 3

Article 2: A thermomechanical finite strain shape memory alloy model and its application to bistable actuators

This article was published as:

Sielenkämper, M.  and Wulfinghoff, S. [2022], 'A thermomechanical finite strain shape memory alloy model and its application to bistable actuators', *Acta Mechanica* **233**, 3059-3094.

Own contributions to the following article:

- planning (large fraction)
- implementation of algorithm (large fraction)
- numerical simulations (large fraction)
- writing of the manuscript (large fraction)
- interpretation of the results (large fraction)

Abstract

This work presents a thermomechanical finite strain shape memory alloy model that utilizes a projection method to deal with the incompressibility constraint on inelastic strains. Due to its finite strain formulation, it is able to accurately predict the behavior of shape memory alloys with high transformation strains. The key feature of this model is the thermomechanical modeling of the shape memory effect and superelastic behavior by optimizing a global, incremental mixed thermomechanical potential, the variation of which yields the linear momentum balance, the energy balance, the evolution equations of the internal variables as well as boundary conditions of Neumann- and Robin-type. The proposed thermal strain model allows to properly capture transformation induced volume changes, which occur in some shape memory alloys. A finite strain dissipation potential is formulated, which incorporates the disappearance of inelastic strains upon austenite transformation. This important property is consistently transferred to the time-discrete potential using a logarithmic stain formulation. Yield and transformation criteria are derived from the dual dissipation potential. The implementation based on an active set search and the algorithmically consistent linearization are discussed in detail. The model is applied in three-dimensional simulations of a bistable actuator design to explore its capabilities.

Keywords: shape memory alloys, constitutive modeling, thermomechanical coupling, projection method, finite element method

3.1 Introduction

Since the discovery of their unique properties, shape memory alloys are used in many medical and engineering fields in various applications. Their frequent appearance stems from their unique features, like the shape memory effect and superelasticity. Both effects emerge from the characteristic first order phase transition from the austenite to the martensite state and vice versa. While the occurring crystallographic effects involving the detwinning of the martensite phase, which enables the shape memory effect, are well understood, the thermomechanical modeling of the occurring effects is not straightforward. In recent times, there has been a high effort to improve shape memory alloy models and to obtain fitting models for specific use cases (see, for reviews, Section 4 of Lester et al. [64] or Cisse et al. [14]). They can be roughly categorized into three classes: models based on statistical thermodynamics, models founded in micromechanics and phenomenological models.

The models based on statistical thermodynamics rely on finding the phase equilibrium through a minimization of a three well potential energy (e.g., Seelecke and Müller [110] or Govindjee et al. [32]). Since these models yield results that also consider the microstructure of the materials, they come mostly with a computational cost which is too high for large, structural simulations. Additionally, gathering required micromechanical material parameters is some-

times an elaborate task.

On the other hand, models based on micromechanics usually consider the mechanics of shape memory alloy (SMA) single-crystals. Many models are then extended into the regime of polycrystal modeling by usage of homogenization techniques (e.g., see Patoor et al. [95] and Lagoudas et al. [58] or the more recent models by Mirzaeifar et al. [80] and Yu et al. [152]). While these models consider the deep, underlying phenomena of shape memory alloys, this advantage again comes with a high computational cost. This makes it really challenging to use these models in large structural simulations of shape memory alloy actuators.

The third group of models is the class of phenomenological models. Usually, they come with the advantage of only having macroscopic material parameters, which mostly are obtained through tensile tests at different temperatures. In recent years, due to the plethora of shape memory alloys effects and applications, many new shape memory alloy models were published, which try to include more and more physical phenomenons. They can be divided into two subgroups: models which include a geometrically linear theory (see, e.g., Auricchio et al. [7], Auricchio et al. [6] and Sedláč et al. [109]), and models which include a geometrically nonlinear theory.

The two main ways to include a geometric nonlinearity in the shape memory alloy model is to either employ a multiplicative split of the deformation gradient going back to [18, 57, 61] or to make use of an additive split (see Nemat-Nasser [86]) of the rate of deformation, which are both well known from plasticity. While models based on the additive split are computationally enticing, they only allow for small strains, while still capturing large rotations well (see, e.g., Qidwai and Lagoudas [98], Müller and Bruhns [84] or Zhang and Baxevanis [156]). On the other hand, models employing a multiplicative split are computationally more elaborate, but can represent finite stretches well (see, e.g., Reese and Christ [101], Arghavani et al. [3] or Wang et al. [137]). Additionally, there exist many new models considering geometric nonlinearities, which are limited to superelasticity (see, e.g., Bellini et al. [9], Wang et al. [136] and Rezaee-Hajidehi et al. [103]). Furthermore, some models also consider transformation induced plasticity (see, e.g., Hartl et al. [39], Xu et al. [148] and its extension to partial phase transformations by Scalet et al. [108]).

Many of these models capture thermomechanics, martensite reorientation and detwinning as well as superelasticity at different temperatures or allow for different elastic properties of the materials while being numerically efficient and robust. However, when trying to model bistable shape memory actuators (see the actuator design in Arivanandhan [4]), one needs a numerically robust, fully thermomechanically coupled finite strain model that also models thermal expansion and volumetric effects during phase transition (see, e.g., Potapov et al. [97]). To our knowledge, there is no model in the literature fulfilling all of the aforementioned requirements, which led to the model described in this paper. For example, the model of Wang et al. [137] considers a thermomechanically coupled finite strain theory which is capable of modeling the shape memory effect as well as superelasticity, neglecting volumetric effects due to transformation as well as different expansion coefficients of the SMA phases.

The model at hand falls into the aforementioned category of phenomenological models. It is embedded into the generalized standard materials framework developed by Halphen and Nguyen [36], which was extended to thermomechanics by Yang et al. [149] and which allows to ensure thermodynamic consistency. The energies as well as the dissipation potential can be seen as an extension of Sedláč et al. [109] to the finite strain case. Because the satisfaction of the incompressibility of inelastic strains for finite strains is not as straightforward as for the small strain case, a projection method developed for plasticity is incorporated into the model (see Hurtado et al. [43] and Sielenkämper et al. [118]). Further, due to the character of the energies used in this model, special numerical treatment is necessary to solve the model equations using a Newton scheme. Since our aim is to model microactuators in which the R-Phase is not present, it is not incorporated into the model. Additionally, tension-compression anisotropy, which is an important effect in many shape memory alloys (for experimental publications see, e.g., Gall et al. [30] or Wang and Zhu [135], and for modeling approaches see, e.g., Zaki et al. [153] or Sedláč et al. [109]), is not included in the model.

In the past, numerous advancements to current microelectromechanical systems (MEMS) based on electrostatics, magnetism and electrothermal principals have been made. For example, Hoffmann et al. [41] proposed a microactuator based on electrothermal activation making use of bimetal effects. This, however, comes with the downside of low actuation frequencies and a high power consumption. Han et al. [37] proposed an electrostatic actuator based micro-switch for photonics. Devices based on electrostatics usually can be actuated with high frequencies and are adaptable to many applications. Devices using optomechanics were developed by, e.g., Eichenfield et al. [19], which come with a rather small tuning range, but allow for a very high operation speed [17]. Despite their unique advantages, current MEMS devices are challenging to use in downsized applications, where a high work output combined with a high power efficiency and bistability is crucial. The microactuator modeled in this paper is based on a concept published by Winzek et al. [140], which utilizes a high temperature shape memory alloy with a large thermal hysteresis. This concept may overcome the aforementioned weaknesses of other actuation principles, as shape memory alloys usually have a large work output density and favorable downscaling capabilities (see, e.g., Kohl [53]). Additionally, downscaling this design is expected to drastically increase the possible actuation frequency in comparison to other, electrothermally activated actuators due to the decreasing masses and increasing thermal gradients. One further key aspect is, that the proposed actuator design requires no power in the stable states.

The paper is structured as follows: First, the energies as well as the dissipation potential is derived. Then, in Section 3.3, numerical strategies necessary to solve the model equations are discussed. The numerical results are subsequently shown in Section 3.4 before a summary and outlook concludes the paper in Section 3.5.

Notation

Throughout this paper, a direct tensor notation is preferred. Scalars and scalar valued

functions are typeset by light-face italic characters, e.g., a or A . First and second-order tensors and tensor-valued functions are represented by bold-face italic letters, e.g., \mathbf{a} or \mathbf{A} . Further, blackboard bold-faced letters are used to denote fourth-order tensors, e.g., \mathbb{c} or \mathbb{C} . Additionally, the transpose of a second-order tensor is designated by \mathbf{A}^\top , while the major transpose of a fourth-order tensor is given by \mathbb{C}^\top . The symmetric and deviatoric part of a second-order tensor \mathbf{A} are denoted by $\text{sym}(\mathbf{A}) = \frac{1}{2}(\mathbf{A} + \mathbf{A}^\top)$ and $\mathbf{A}' = \mathbf{A} - \frac{1}{3}\text{tr}(\mathbf{A})\mathbf{I}$, respectively. Here, \mathbf{I} is the second-order identity tensor and $\text{tr}(\mathbf{A})$ denotes the trace of \mathbf{A} . A double contraction of two tensors \mathbf{A} and \mathbf{B} is denoted by $\mathbf{A} : \mathbf{B}$, while the dyadic product is denoted by $\mathbf{a} \otimes \mathbf{b}$. A determinant of a tensor \mathbf{A} is either designated by $\det(\mathbf{A})$ or by \mathbf{A} 's third invariant $\text{III}_\mathbf{A}$.

3.2 Modeling of shape memory alloys

3.2.1 Kinematics

The deformation gradient \mathbf{F} maps a line element from the reference configuration of a body with Volume V_0 into the current configuration of a body with Volume V and is defined as

$$\mathbf{F} = \text{Grad}(\mathbf{x}(\mathbf{X}, t)), \quad (3.1)$$

where $\text{Grad}(\bullet)$ refers to the gradient with respect to the reference configuration while \mathbf{X} and \mathbf{x} are the position vectors of a material point in the reference and current configuration, respectively. We consider a multiplicative split of the deformation gradient in the form (see Wang et al. [137])

$$\mathbf{F} = \mathbf{F}^e \mathbf{F}^i \mathbf{F}^\theta, \quad (3.2)$$

where \mathbf{F}^e is the elastic, \mathbf{F}^i the isochoric part of the deformation due to transformation¹ and \mathbf{F}^θ the part which describes the volume change due to thermal expansion and transformation² of the deformation gradient. This is fairly similar to the multiplicative split in plasticity going back to the works of Eckart [18], Kröner [57] and Lee [61]. Further, we define the determinant $J^\theta = J^\theta(\theta, \xi) = \det(\mathbf{F}^\theta)$ to be a function of the absolute temperature θ and the martensite volume fraction $\xi \in [0, 1]$. Since the thermal deformation is assumed to be volumetric, we can express \mathbf{F}^θ in terms of J^θ as

$$\mathbf{F}^\theta = (J^\theta)^{\frac{1}{3}} \mathbf{I}. \quad (3.3)$$

As is commonly done and will be useful later, we define the elastic and inelastic left Cauchy-Green tensors as

$$\mathbf{b}^e = \mathbf{F}^e \mathbf{F}^{e\top}, \quad \mathbf{b}^i = \mathbf{F}^i \mathbf{F}^{i\top}. \quad (3.4)$$

¹for brevity, \mathbf{F}^i is called the inelastic part of the deformation gradient throughout this paper.

²for brevity, \mathbf{F}^θ is called the thermal part of the deformation gradient throughout this paper.

Likewise, we define the inelastic right Cauchy-Green tensor and the inelastic Green-Lagrange strain

$$\mathbf{C}^i = \mathbf{F}^{iT} \mathbf{F}^i, \quad \mathbf{E}^i = \frac{1}{2}(\mathbf{C}^i - \mathbf{I}). \quad (3.5)$$

Motivated by the observation that the shape-memory effect is caused by an almost³ volume preserving transformation of the crystal lattice, we assume \mathbf{C}^i to be volume preserving. Therefore, because $\det(\mathbf{F}^i) = 1$ has to hold, the determinant of the deformation gradient is given by

$$J = \det(\mathbf{F}) = J^\theta J^e, \quad (3.6)$$

with $J^e = \det(\mathbf{F}^e)$. Additionally, we define the velocity gradient \mathbf{l} and its symmetric part \mathbf{d} as

$$\mathbf{l} = \dot{\mathbf{F}} \mathbf{F}^{-1}, \quad \mathbf{d} = \text{sym}(\mathbf{l}). \quad (3.7)$$

Finally, we define the inelastic 'velocity gradient' \mathbf{L}^i and its symmetric part \mathbf{D}^i in analogy to \mathbf{l} and \mathbf{d} by

$$\mathbf{L}^i = \dot{\mathbf{F}}^i \mathbf{F}^{i-1}, \quad \mathbf{D}^i = \text{sym}(\mathbf{L}^i). \quad (3.8)$$

3.2.2 (Im-)Balance equations

Momentum balances

The quasistatic linear momentum balance for a body with volume V in the current configuration is given by

$$\text{div}(\boldsymbol{\sigma}) + \rho \mathbf{b} = \mathbf{0} \quad \text{in } V, \quad (3.9)$$

where \mathbf{b} is the body force and ρ is the mass density. Further, Cauchy's lemma $\mathbf{t} = \boldsymbol{\sigma} \mathbf{n}$ as well as the angular momentum balance $\boldsymbol{\sigma} = \boldsymbol{\sigma}^T$ are assumed to hold.

Energy balance

Next, the balance of the energies is given by

$$\dot{u} = \boldsymbol{\tau} : \mathbf{d} - J \text{div}(\mathbf{q}) + w, \quad (3.10)$$

where u is the internal energy density per unit reference volume, $\boldsymbol{\tau}$ denotes the Kirchhoff stress tensor, \mathbf{q} is the heat flux vector in the current configuration and w represents the energy source term. Further, the divergence of the heat flux vector $\mathbf{Q} = J \mathbf{F}^{-1} \mathbf{q}$ with respect to the reference configuration is then given by $J \text{div}(\mathbf{q}) = \text{Div}(\mathbf{Q})$.

³Actually, for some shape memory alloys, there exists a difference in density between the martensite and austenite phase. For NiTiHf, see the article by Potapov et al. [97] and for modeling approaches for this effect in NiTi the work of Qidwai and Lagoudas [98]. We also model this phenomenon in Section 3.2.3.

Clausius-Planck inequality

To be thermodynamically consistent, the Clausius-Planck inequality

$$\mathcal{D} = \boldsymbol{\tau} : \mathbf{d} + \theta \dot{s} - \dot{u} \geq 0, \quad (3.11)$$

where \mathcal{D} is the mechanical dissipation density per unit reference volume and s is the entropy density per unit reference volume, has to be fulfilled at any time. Further, the Helmholtz free energy, which is defined by $\psi = u - \theta s$, is introduced. We assume ψ to be a function of the deformation gradient \mathbf{F} , the absolute temperature θ and the internal variables ξ and \mathbf{C}^i , i.e., $\psi = \psi(\mathbf{F}, \mathbf{C}^i, \xi, \theta)$. Furthermore, we make the common assumption that

$$\mathcal{D} \Big|_{\mathbf{C}^i, \xi} = 0, \quad (3.12)$$

i.e., there is no energy dissipation when the internal variables are virtually fixed. As Eq. (3.12) must hold for arbitrary processes, one can show the following standard results

$$\boldsymbol{\tau} = \frac{\partial \psi}{\partial \mathbf{F}} \mathbf{F}^\top, \quad s = -\frac{\partial \psi}{\partial \theta}. \quad (3.13)$$

Now, it is easy to show that

$$\mathcal{D} = \boldsymbol{\Sigma}^i : \mathbf{D}^i + q \dot{\xi} \geq 0 \quad (3.14)$$

with shorthand notations for the effective Mandel stress with respect to the intermediate configuration $\boldsymbol{\Sigma}^i = -2\mathbf{F}^i \left(\partial \psi / \partial \mathbf{C}^i \right) \mathbf{F}^{i\top}$ and the thermodynamic force $q = -\partial \psi / \partial \xi$ associated with the martensite volume fraction ξ .

3.2.3 Helmholtz free energy

The Helmholtz free energy density ψ for this model is assumed to be a sum of elastic, chemical and hardening-like contributions in the form

$$\psi = \psi_e(\mathbf{b}^e, \xi) + \psi_c(\xi, \theta) + \psi_h(\mathbf{C}^i, \xi). \quad (3.15)$$

Elastic energy

The elastic energy ψ_e is assumed to be isotropic and to follow a modified Neo-Hookean formulation:

$$\psi_e(\mathbf{b}^e, \xi) = \frac{\lambda(\xi)}{4} (J^{e2} - 1 - 2 \ln J^e) + \frac{\mu(\xi)}{2} (\text{tr}(\mathbf{b}^e) - 3 - 2 \ln J^e). \quad (3.16)$$

Here, $\lambda(\xi)$ and $\mu(\xi)$ are the Lamé parameters in dependence of the martensite volume fraction ξ . Further, we use a Reuss-like mixture rule to estimate the elastic constants, i.e.,

$$\mu(\xi) = \left(\frac{\xi}{\mu_M} + \frac{1-\xi}{\mu_A} \right)^{-1}, \quad \lambda(\xi) = \left(\frac{\xi}{\lambda_M} + \frac{1-\xi}{\lambda_A} \right)^{-1}. \quad (3.17)$$

Here, and subsequently, the indices \bullet_A and \bullet_M refer to the austenite and martensite phase, respectively. Manipulating Eq. (3.13), one can show that the ordinary form

$$\boldsymbol{\tau} = 2\mathbf{b}^e \frac{\partial \psi_e}{\partial \mathbf{b}^e} \quad (3.18)$$

holds for elastic isotropy. Using Eqns. (3.14) and (3.18), one can show that

$$\boldsymbol{\Sigma}^i = -2\mathbf{F}^i \frac{\partial \psi}{\partial \mathbf{C}^i} \mathbf{F}^{iT} = \boldsymbol{\Sigma}^e - \underbrace{2\mathbf{F}^i \frac{\partial \psi_h}{\partial \mathbf{C}^i} \mathbf{F}^{iT}}_{:=\boldsymbol{\Sigma}^h} \quad (3.19)$$

where $\boldsymbol{\Sigma}^e = \mathbf{C}^e \mathbf{F}^{e-1} \boldsymbol{\tau} \mathbf{F}^{e-T}$ is the Mandel stress with respect to the intermediate, elastically unloaded configuration, which is symmetric due to the assumption of elastic isotropy.

Chemical energy

For the chemical energy, we assume a standard relationship (see, e.g., Lexcellent et al. [65] or Panico and Brinson [94]):

$$\begin{aligned} \psi_c &= \underbrace{(u_0^A - \theta s_0^A)}_{\psi_0^A} - \underbrace{\xi (\Delta u^{AM} - \theta \Delta s^{AM})}_{\Delta \psi^{AM}} + c \left(\theta - \theta_0 - \theta \ln \frac{\theta}{\theta_0} \right) \\ &= u_0^A - \theta s_0^A + \xi (\theta - \theta_0) \Delta s^{AM} + c \left(\theta - \theta_0 - \theta \ln \frac{\theta}{\theta_0} \right) \end{aligned} \quad (3.20)$$

where c is the specific heat capacity and Δs^{AM} is the difference in specific entropy of the austenite and martensite phase: $\Delta s^{AM} = s_0^A - s_0^M$. Here, c is assumed constant (compare [109]). In Eq. (3.20), we made use of the definition of the equilibrium temperature of austenite and martensite, which is $\theta_0 = \Delta u^{AM} / \Delta s^{AM}$. Further, we assume for the equilibrium temperature $A_s > \theta_0 > M_s$, where A_s is the temperature where the reverse transformation starts and M_s is the starting temperature of the forward transformation.

Hardening energy

Since the inelastic strains vanish as $\xi \rightarrow 0$, the inelastic strains \mathbf{E}^i are assumed to satisfy the relation $\mathbf{E}^i = \xi \mathbf{E}^t$ (see Otsuka and Ren [92]), where \mathbf{E}^t is a measure for the effective transformation strain. Now, the hardening-like energy adapted from Sedláček et al. [109] for

finite strains is assumed to be given by

$$\psi_h = kE^{\text{int}}\xi \frac{\langle \mathbf{E}^t \rangle^2}{1 - \langle \mathbf{E}^t \rangle^4}, \quad \langle \mathbf{E}^t \rangle = \sqrt{\frac{2}{3}} \frac{\|\mathbf{E}^t\|}{k}. \quad (3.21)$$

In here, k is the maximum transformation strain, E^{int} is a hardening related parameter and $\langle \mathbf{E}^t \rangle$ is a modified von Mises equivalent strain. Obviously, $\psi_h \rightarrow \infty$ for $\langle \mathbf{E}^t \rangle \rightarrow 1$. This is desired, since it captures the martensite becoming fully detwinned, which is assumed to cost additional energy.

Thermal strains

Modeling thermal strains of multi-phase materials requires attention and special treatment. In this work, the determinant of the thermal part of the deformation gradient is connected to the coefficients of thermal expansion (CTEs) by

$$J^\theta = \left(1 + \varepsilon^\theta(\xi, \theta)\right)^3, \quad \varepsilon^\theta = \xi\alpha_M(\theta - \theta_{\text{refM}}) + (1 - \xi)\alpha_A(\theta - \theta_{\text{refA}}). \quad (3.22)$$

Here, $\varepsilon^\theta(\xi, \theta)$ is the thermal strain. Additionally, α_M and α_A are the CTEs of the martensite and austenite phase, respectively. Further, θ_{refM} and θ_{refA} are the reference temperatures for austenite and martensite. Here, we want to note that it is necessary to distinguish the two in order to properly represent the transformation-induced volume change which is present in some shape memory alloys (see below).

Example:

We assume a one-dimensional SMA rod with only one reference temperature. Further, we simplify the SMA model by assuming that martensite and austenite transformation occur at the distinct temperatures θ_M and θ_A , and not at temperature ranges (see Fig. 3.1). The thermal strains are then given by $\varepsilon^\theta = \alpha_{\text{AM}}(\theta - \theta_{\text{ref}})$, with $\alpha_{\text{AM}} = \xi\alpha_M + (1 - \xi)\alpha_A$. Furthermore, we assume that we know the shape of our rod at the thermal annealing temperature θ_{ref} (see Fig. 3.1). During forward transformation, the jump of ε^θ at θ_M can be analytically calculated and is given by

$$\Delta\varepsilon^\theta = (\theta_{\text{ref}} - \theta_M)(\alpha_A - \alpha_M). \quad (3.23)$$

Here, we can not only see that this jump is dependent on the reference temperature, but that it can be positive or negative, depending on the reference temperature θ_{ref} being larger or smaller than the forward transformation temperature θ_M .

Using two reference temperatures θ_{refM} and θ_{refA} , we can not only circumvent this problem, but also model the magnitude of this jump in compliance with, e.g., the experimental results of Potapov et al. [97], where they calculated the jump ΔV in volume from the lattice

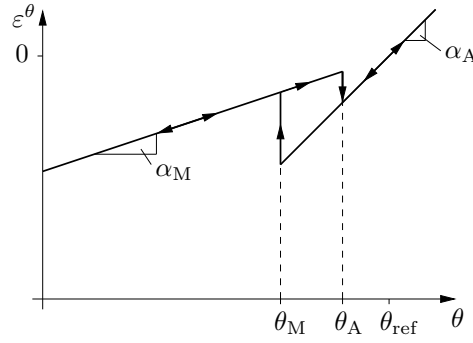


Fig. 3.1: Thermal strains when only considering one reference temperature θ_{ref} .

parameters for $\text{Ni}_{49.8}\text{Ti}_{35.2}\text{Hf}_{15}$ to be 0.47 %. However, we want to emphasize that this effect does not play a significant role in many other shape memory alloys.

3.2.4 Dissipation potential

The dissipative behavior of the shape memory alloy is assumed to be governed by the dissipation potential

$$\phi(\dot{\xi}, \mathbf{D}^i, \xi) = \begin{cases} \phi_M, & \text{if } \dot{\xi} \geq 0 \wedge \text{tr}(\mathbf{D}^i) = 0 \\ \phi_A, & \text{if } \dot{\xi} < 0 \wedge \text{tr}(\mathbf{D}^i) = 0 \\ \infty, & \text{else.} \end{cases} \quad (3.24)$$

Here, depending on the direction of transformation, either ϕ_M or ϕ_A are used as the active dissipation potential. As is commonly done for plasticity, the dissipation is assumed to be infinite if $\text{tr}(\mathbf{D}^i) \neq 0$, which guarantees volume-preserving inelastic deformations (compare Rockafellar [104]). Additionally, we define the dissipation potential for an increasing martensite volume fraction as

$$\phi_M = \underbrace{\dot{\xi} \Delta s^{\text{AM}}((\theta_0 - M_s) + \xi(M_s - M_f))}_{=: Q_M(\xi)} + \sqrt{\frac{2}{3}} \sigma^{\text{reo}} \|\mathbf{D}^i\|, \quad (3.25)$$

where M_f is the finish temperature of forward transformation and σ^{reo} is the reorientation stress. Likewise, the dissipation potential when going to a higher austenite volume fraction is defined as

$$\phi_A = \underbrace{\dot{\xi} \Delta s^{\text{AM}}((\theta_0 - A_f) + \xi(A_f - A_s))}_{=: Q_A(\xi)} + \sqrt{\frac{2}{3}} \sigma^{\text{reo}} \left(\left\| \frac{\dot{\xi}}{\xi} \boldsymbol{\epsilon}^i \right\| + \left\| \mathbf{D}^i - \frac{\dot{\xi}}{\xi} \boldsymbol{\epsilon}^i \right\| \right). \quad (3.26)$$

Here, A_f is the finish temperature of reverse transformation and $\boldsymbol{\epsilon}^i = \frac{1}{2} \ln \mathbf{b}^i$. The significance of using the logarithmic inelastic strain $\boldsymbol{\epsilon}^i$ as a strain measure for the dissipation is explained below. The dissipation potential represents a geometrically nonlinear generalization of the small strain potential proposed in Sedláč et al. [109], which is based on the works of Bernardini

and Pence [10], Panico and Brinson [94] and Moumni et al. [83].

3.2.5 Transformation / yield criteria and inelastic evolution equations

It is assumed that the evolution of \mathbf{C}^i and ξ follow from the minimization problem

$$\inf_{\mathbf{D}^i, \dot{\xi}} \dot{\psi} + \phi, \quad (3.27)$$

which is equivalent to the Legendre-Fenchel transformed problem (compare Equations (3.14),(3.19))

$$\sup_{\mathbf{D}^i, \dot{\xi}} \left(\boldsymbol{\Sigma}^e - \boldsymbol{\Sigma}^h \right) : \mathbf{D}^i + q\dot{\xi} - \phi(\dot{\xi}, \mathbf{D}^i, \xi) =: \phi^*(\boldsymbol{\Sigma}^i, q, \xi), \quad (3.28)$$

where ϕ^* is the dual dissipation potential of ϕ :

$$\phi^*(\boldsymbol{\Sigma}^i, q, \xi) = \sup_{\substack{\mathbf{D}^i, \dot{\xi} \\ \text{tr}(\mathbf{D}^i)=0}} \boldsymbol{\Sigma}^i : \mathbf{D}^i + q\dot{\xi} - \begin{cases} \phi_M, & \text{if } \dot{\xi} \geq 0 \\ \phi_A, & \text{if } \dot{\xi} < 0. \end{cases} \quad (3.29)$$

Due to the minimization in Eq. (3.27), there is a tendency that for $\dot{\xi} < 0$ we get $\mathbf{D}^i - (\dot{\xi}/\xi)\boldsymbol{\varepsilon}^i = \mathbf{0}$ as a result of the last term in Eq. (3.26). In that case $\mathbf{b}^i \rightarrow \mathbf{I}$ as $\xi \rightarrow 0$ in time (for a proof, see Appendix 3.6), i.e., for vanishing martensite content, the inelastic strain vanishes, which is a physical necessity. With the reparametrization $\mathbf{D}^t = \mathbf{D}^i - (\dot{\xi}/\xi)\boldsymbol{\varepsilon}^i$ (compare Eq. (3.26)) we find

$$\begin{aligned} & \phi^*(\boldsymbol{\Sigma}^i, q, \xi) \\ &= \sup \begin{cases} \sup_{\substack{\mathbf{D}^i, \dot{\xi} \geq 0 \\ \text{tr}(\mathbf{D}^i)=0}} \boldsymbol{\Sigma}^i : \mathbf{D}^i - \sqrt{\frac{2}{3}}\sigma^{\text{reo}} \|\mathbf{D}^i\| + (q - Q_M(\xi))\dot{\xi} \\ \sup_{\substack{\mathbf{D}^t, \dot{\xi} < 0 \\ \text{tr}(\mathbf{D}^t)=0}} \boldsymbol{\Sigma}^i : \mathbf{D}^t - \sqrt{\frac{2}{3}}\sigma^{\text{reo}} \|\mathbf{D}^t\| + \left(\frac{1}{\xi}\boldsymbol{\Sigma}^i : \boldsymbol{\varepsilon}^i + q - Q_A(\xi) + \sqrt{\frac{2}{3}}\sigma^{\text{reo}} \frac{\|\boldsymbol{\varepsilon}^i\|}{\xi}\right)\dot{\xi}, \end{cases} \end{aligned} \quad (3.30)$$

where Q_M and Q_A are defined in Eqns. (3.25) and (3.26) and capture the transformation hysteresis due to temperature change. Further reparametrizing $\mathbf{D}^{i/t} = \lambda \mathbf{N}$ with $\lambda \geq 0$ and the obvious solution

$$\mathbf{N} = \frac{\boldsymbol{\Sigma}^{i'}}{\|\boldsymbol{\Sigma}^{i'}\|}, \quad (3.31)$$

it follows that

$$\phi^*(\boldsymbol{\Sigma}^i, q, \xi) = \sup \begin{cases} \sup_{\lambda \geq 0, \dot{\xi} \geq 0} \lambda f(\boldsymbol{\Sigma}^i) + \dot{\xi} g_M(q) \\ \sup_{\lambda \geq 0, \dot{\xi} < 0} \lambda f(\boldsymbol{\Sigma}^i) + \dot{\xi} g_A(q, \boldsymbol{\Sigma}^i). \end{cases} = \begin{cases} 0, & \text{if } f \leq 0 \wedge g_M \leq 0 \wedge g_A \geq 0 \\ \infty, & \text{else} \end{cases} \quad (3.32)$$

Thus, we find the following transformation and yield criteria in the classical Karush-Kuhn-Tucker form as well as evolution equations:

$$\begin{aligned} f(\boldsymbol{\Sigma}^i) &= \|\boldsymbol{\Sigma}^{i'}\| - \sqrt{\frac{2}{3}}\sigma^{\text{reo}} \leq 0; & \lambda f &= 0; \quad \lambda \geq 0; & \mathbf{D}^i &= \lambda \frac{\partial f}{\partial \boldsymbol{\Sigma}^i} \\ g_M(q) &= q - Q_M(\xi) \leq 0; & \langle \dot{\xi} \rangle g_M &= 0; \quad \dot{\xi} > 0 & \text{possible if } g_M &= 0 \\ g_A(q, \boldsymbol{\Sigma}^i) &= \frac{1}{\xi} \boldsymbol{\Sigma}^i : \boldsymbol{\varepsilon}^i + q - Q_A(\xi) + \sqrt{\frac{2}{3}} \frac{\sigma^{\text{reo}}}{\xi} \|\boldsymbol{\varepsilon}^i\| \geq 0; & \langle \dot{\xi} \rangle_- g_A &= 0; \quad \dot{\xi} < 0 & \text{possible if } g_A &= 0 \end{aligned} \quad (3.33)$$

with the Macauley bracket $\langle \bullet \rangle = (\bullet + |\bullet|)/2$ and its modified form $\langle \bullet \rangle_- = (\bullet - |\bullet|)/2$. It is then straightforward to prove thermodynamic consistency (see Eq. (3.14)). The '0'-branch in Eq. (3.32) implies that (see Eq. (3.28))

$$\mathcal{D} = \boldsymbol{\Sigma}^i : \mathbf{D}^i + q\dot{\xi} = \phi. \quad (3.34)$$

Thus, we find with Eqns. (3.4) and (3.10) the following alternative form of the energy balance:

$$\theta \dot{s} = \phi - \text{Div}(\mathbf{Q}) + w. \quad (3.35)$$

3.2.6 Rate potential

To simplify the discussion, we start by the isothermal case, i.e., in a first step, the temperature θ is considered to be a given parameter. The rate potential and its time discretized form read

$$\pi = \dot{\psi} + \phi \simeq \frac{1}{\Delta t} \left(\psi(\mathbf{F}, \mathbf{C}^i, \xi, \theta) - \psi_n + \phi_\Delta \right). \quad (3.36)$$

Here, ψ_n refers to the Helmholtz free energy at the previous time, Δt is the time step from t_n to t_{n+1} ⁴ and ϕ_Δ is the time discretized version of the dissipation potential multiplied by Δt , defined by

$$\phi_\Delta = \begin{cases} \sigma^{\text{reo}} \Delta \alpha + Q(\mathbf{C}_n^i, \xi_{n+\frac{1}{2}}, \text{sg}(\Delta \xi)) \Delta \xi & \text{if } \text{III}^i = 1 \\ \infty, & \text{else,} \end{cases} \quad (3.37)$$

with $\text{sg}(\bullet)$ referring to the sign function and $\Delta \xi = \xi - \xi_n$. Additionally, III^i is the third invariant of \mathbf{C}^i , i.e.,

$$\text{III}^i = \det(\mathbf{C}^i). \quad (3.38)$$

⁴For brevity and simplicity, we dropped the index $n + 1$ whenever we deemed it to be not helpful for the presentation.

The constraint $\text{III}^i = 1$ in Eq. (3.37) is consistent with the requirement $\text{tr}(\mathbf{D}^i) = 0$ (see Eq. (3.24)). Additionally, we compute the effective inelastic strain increment $\Delta\alpha$ as

$$\Delta\alpha = \sqrt{\frac{2}{3}} \left\| \frac{1}{2} \mathbf{U}_n^{i-1} \Delta_r \mathbf{C}^i \mathbf{U}_n^{i-1} \right\| =: \sqrt{\frac{2}{3}} \left\| \frac{1}{2} \Delta_r \mathbf{C}^i \right\|_{\mathbf{U}_n^{i-1}} \quad (3.39)$$

where the inelastic right stretch is $\mathbf{U}^i = \sqrt{\mathbf{C}^i}$, $\Delta_r \mathbf{C}^i = \mathbf{C}^i - \mathbf{C}_r^i$ and the shorthand notation $\|\bullet\|_{\mathbf{U}_n^{i-1}} = \left\| \mathbf{U}_n^{i-1} \bullet \mathbf{U}_n^{i-1} \right\|$ is used. Further, we define \mathbf{C}_r^i as

$$\mathbf{C}_r^i = \left(\mathbf{C}_n^i \right)^{1 + \langle \frac{\Delta\xi}{\xi_n} \rangle_-}. \quad (3.40)$$

Now, for ξ going to zero in any given step, i.e., $\xi = \xi_n + \Delta\xi = 0 \Leftrightarrow \Delta\xi = -\xi_n \rightarrow \mathbf{C}_r^i = (\mathbf{C}_n^i)^0 = \mathbf{I}$, \mathbf{C}_r^i goes back to unity again. Thus, the time discrete potential (3.37) based on the definition (3.40) of \mathbf{C}_r^i is the key to ensure that the inelastic strain consistently disappears during the transformation from martensite to austenite. Further, the function $Q(\mathbf{C}_n^i, \xi_{n+\frac{1}{2}}, \text{sg}(\Delta\xi))$ is given by

$$Q(\mathbf{C}_n^i, \xi_{n+\frac{1}{2}}, \text{sg}(\Delta\xi)) = \begin{cases} \Delta s^{\text{AM}}((\theta_0 - M_s) + \xi_{n+\frac{1}{2}}(M_s - M_f)), & \text{if } \Delta\xi \geq 0 \\ \Delta s^{\text{AM}}((\theta_0 - A_f) + \xi_{n+\frac{1}{2}}(A_f - A_s)) - \sqrt{\frac{2}{3}} \frac{\sigma^{\text{reo}}}{\xi_n} \|\boldsymbol{\epsilon}_n^i\|, & \text{if } \Delta\xi < 0, \end{cases} \quad (3.41)$$

where $\xi_{n+\frac{1}{2}} = (\xi + \xi_n)/2$ is the midpoint evaluation of ξ , which is employed to obtain a reasonable transformation when the material is not stressed (see Frost et al. [27] for a similar concept). Moreover, we used \mathbf{C}_n^i and ξ_n for the reverse transformation in Eq. (3.41) to circumvent the eigenvalue problem as well as its linearization in every local Newton iteration. In this way, it suffices to solve the eigenvalue problem once per time step.

Using some further shorthand notations for constant terms in Q , we obtain

$$\begin{aligned} \phi_\Delta &= \sigma^{\text{reo}} \Delta\alpha \\ &+ \begin{cases} \overbrace{\Delta s^{\text{AM}}((\theta_0 - M_s) + \xi_n(M_s - M_f))}^{Q^{\text{M0}}} \Delta\xi + \frac{1}{2} \overbrace{\Delta s^{\text{AM}}(M_s - M_f)}^{H^{\text{M}}} \Delta\xi^2, & \Delta\xi \geq 0 \\ \overbrace{\left(\Delta s^{\text{AM}}((\theta_0 - A_f) + \xi_n(A_f - A_s)) - \sqrt{\frac{2}{3}} \frac{\sigma^{\text{reo}}}{\xi_n} \|\boldsymbol{\epsilon}_n^i\| \right)}^{Q^{\text{A0}}} \Delta\xi + \frac{1}{2} \overbrace{\Delta s^{\text{AM}}(A_f - A_s)}^{H^{\text{A}}} \Delta\xi^2, & \Delta\xi < 0 \end{cases} \\ &= \sigma^{\text{reo}} \Delta\alpha + \begin{cases} Q^{\text{M0}} \Delta\xi + \frac{1}{2} H^{\text{M}} (\Delta\xi)^2, & \text{if } \Delta\xi \geq 0 \\ Q^{\text{A0}} \Delta\xi + \frac{1}{2} H^{\text{A}} (\Delta\xi)^2, & \text{if } \Delta\xi < 0 \end{cases} \\ &= \sigma^{\text{reo}} \Delta\alpha + Q(\text{sg}(\Delta\xi)) \Delta\xi + \frac{1}{2} H(\text{sg}(\Delta\xi)) (\Delta\xi)^2, \end{aligned} \quad (3.42)$$

where H and Q summarize H^M , H^A , Q^{M0} and Q^{A0} , respectively. Further, $\det(\mathbf{C}_r^i) = \det(\mathbf{C}_n^i) = 1$ clearly holds when looking at the definition of \mathbf{C}_r^i in Eq. (3.40). Additionally, the consistency of ϕ_Δ with the time-continuous theory is trivially proven except for $\Delta\alpha$. Therefore, we have a look at the approximation of a^{1+x} at $x \ll 1$:

$$a^{1+x} \stackrel{|x| \ll 1}{\approx} a + xa \ln a \quad (3.43)$$

as $\ln a^x \approx a^x - 1$. Hence, we can approximate \mathbf{C}_r^i as

$$\mathbf{C}_r^i \stackrel{|\Delta\xi| \ll 1}{\approx} \mathbf{C}_n^i + \mathbf{C}_n^i \ln(\mathbf{C}_n^i) \left\langle \frac{\Delta\xi}{\xi_n} \right\rangle_- . \quad (3.44)$$

We can use this result for computation of the time-discrete derivative of α :

$$\begin{aligned} \frac{\Delta\alpha}{\Delta t} &\approx \sqrt{\frac{2}{3}} \left\| \frac{1}{2} \mathbf{U}_n^{i-1} \frac{1}{\Delta t} \Delta \mathbf{C}^i \mathbf{U}_n^{i-1} - \frac{1}{2} \ln(\mathbf{C}_n^i) \left\langle \frac{\Delta\xi}{\Delta t \xi_n} \right\rangle_- \right\| \\ &= \sqrt{\frac{2}{3}} \left\| \frac{1}{2} \mathbf{F}_n^{i-\top} \frac{1}{\Delta t} \Delta \mathbf{C}^i \mathbf{F}_n^{i-1} - \frac{1}{2\xi_n} \ln(\mathbf{b}_n^i) \left\langle \frac{\Delta\xi}{\Delta t} \right\rangle_- \right\|. \end{aligned} \quad (3.45)$$

Therefore, with the time step width going to zero, we obtain

$$\Delta t \rightarrow 0 \Rightarrow \sqrt{\frac{2}{3}} \left\| \mathbf{D}^i - \frac{\dot{\xi}}{2\xi} \ln(\mathbf{b}^i) \right\| = \dot{\alpha}, \quad (3.46)$$

which is consistent with the time-continuous Eq. (3.26). Furthermore, since $\mathbb{I}^i = \det(\mathbf{C}^i) = 1$, we get for the time derivative of \mathbb{I}^i :

$$\dot{\mathbb{I}}^i = \dot{\mathbb{I}}^i \mathbf{C}^{i-1} : \dot{\mathbf{C}}^i = 2\mathbb{I}^i \text{tr}(\mathbf{D}^i) = 0. \quad (3.47)$$

Hence, since ϕ_Δ is time-continuous, we can use the discretized potential to solve for \mathbf{C}^i and ξ :

$$(\mathbf{C}^i, \xi) = \arg \inf_{\mathbf{C}^i, \xi} \underbrace{(\psi - \psi_n + \phi_\Delta)}_{\pi_\Delta} = \arg \inf_{\mathbf{C}^i, \xi} \pi_\Delta(\mathbf{F}, \mathbf{C}^i, \xi, \theta). \quad (3.48)$$

3.3 Numerical strategies

The potential π_Δ , as we formulated it in Eq. (3.48) carries two major numerical difficulties. First, \mathbf{C}^i is constrained to be volume preserving, i.e., $\det(\mathbf{C}^i) = 1$, which is very important to be exactly satisfied to comply with physics and not accumulate errors. Secondly, π_Δ is not differentiable at $\Delta_r \mathbf{C}^i = 0$. The strategies employed to overcome these and other numerical difficulties as well as general numerical approaches are presented in this section.

3.3.1 Inelastic volume preservation

To deal with the constraint in \mathbf{C}^i , we employ a strategy using a projection of \mathbf{C}^i into the space of unimodular tensors, which was introduced by Hurtado et al. [43] for crystal plasticity. Our approach here closely follows the approach presented in Sielenkämper et al. [118].

First, we express \mathbf{C}^i in terms of the unconstrained inelastic auxiliary right Cauchy-Green tensor $\hat{\mathbf{C}}^i$:

$$\mathbf{C}^i = \hat{\mathbb{I}}\mathbb{I}^{i-\frac{1}{3}} \hat{\mathbf{C}}^i, \quad \hat{\mathbb{I}}\mathbb{I}^i = \det(\hat{\mathbf{C}}^i). \quad (3.49)$$

Thus, $\det(\mathbf{C}^i) = 1$ is automatically satisfied. This idea is borrowed from various formulations in hyperelasticity, where similar approaches are used to decouple volumetric and deviatoric deformations (see, e.g., Flory and Volkenstein [26] or Simo et al. [121]). Now, we replace \mathbf{C}^i by $\mathbf{C}^i(\hat{\mathbf{C}}^i)$ in the minimization problem in Eq. (3.48):

$$(\mathbf{C}^i, \xi) = \arg \inf_{\hat{\mathbf{C}}^i, \xi} \pi_{\Delta}(\mathbf{F}, \mathbf{C}^i(\hat{\mathbf{C}}^i), \xi, \theta). \quad (3.50)$$

However, while this removes the constraint from the minimization problem, π_{Δ} is now invariant with respect to changes of $\hat{\mathbb{I}}\mathbb{I}^i$, and is therefore not uniquely solvable. For this reason, we add the regularization energy

$$\psi_r = \frac{A}{2} (\hat{\mathbb{I}}\mathbb{I}^i - B)^2 \quad (3.51)$$

to π_{Δ} . We want to emphasize that this has *no* effect on the solution. This is due to the fact that $\hat{\mathbb{I}}\mathbb{I}^i = B$ is exactly satisfied after converging to a solution, since ψ_r is the only term in π_{Δ} that is dependent on $\hat{\mathbb{I}}\mathbb{I}^i$. Therefore, ψ_r is zero at any solution of the minimization problem. Further, the constants A and B can be chosen arbitrarily. They do not have to be chosen particularly large. In this paper, we chose $B = 1$, but any other positive value could be chosen and would lead to the exact same results. In this special case, $\mathbf{C}^i = \hat{\mathbf{C}}^i$ is exactly satisfied once the solution algorithm is converged. However, we want to note that this equality does not hold in the not yet converged state.

The following tensor, which is used in the residuals and stiffness terms later, is also known from hyperelastic models:

$$\mathbb{P}_i(\hat{\mathbf{C}}^i) = \hat{\mathbb{I}}\mathbb{I}^{i-\frac{1}{3}} \left(\mathbb{I}^s - \frac{1}{3} \hat{\mathbf{C}}^i \otimes \hat{\mathbf{C}}^{i-1} \right) = \frac{\partial \mathbf{C}^i}{\partial \hat{\mathbf{C}}^i}, \quad (3.52)$$

where \mathbb{I}^s is the fourth order identity on symmetric second order tensors.

3.3.2 Differentiability at $\Delta_r \mathbf{C}^i = 0$

The ansatz $\mathbf{C}^i = \hat{\mathbb{I}}\mathbb{I}^{i-\frac{1}{3}} \hat{\mathbf{C}}^i$ renders the solution \mathbf{C}^i a priori volume preserving, but it does not solve the lack of differentiability of ϕ_{Δ} for $\mathbf{C}^i = \mathbf{C}_r^i$. However, this is achieved by the

following reparametrization:

$$\hat{\mathbf{C}}^i(\Delta\gamma, \tilde{\mathbf{N}}^s) = \mathbf{C}_r^i + 2\Delta\gamma \frac{\tilde{\mathbf{N}}^s}{\|\tilde{\mathbf{N}}^s\|_{U_n^{i-1}}}, \quad \Delta\gamma \geq 0 \quad (3.53)$$

where $\tilde{\mathbf{N}}^s$ is an unconstrained, symmetric 2nd-order tensor. The final unconstrained minimization problem now reads

$$(\Delta\gamma, \tilde{\mathbf{N}}^s, \xi) = \arg \min_{\Delta\gamma > 0, \tilde{\mathbf{N}}^s, \xi} \pi_\Delta(\mathbf{F}, \mathbf{C}^i, (\hat{\mathbf{C}}^i(\Delta\gamma, \tilde{\mathbf{N}}^s)), \xi, \theta). \quad (3.54)$$

It is easy to show that π_Δ is invariant with respect to $\|\tilde{\mathbf{N}}^s\|_{U_n^{i-1}}$. In order to render the solution unique, we further modify ψ_r :

$$\psi_r = \frac{A}{2} \left((\hat{\Pi}^i - 1)^2 + \left(\|\tilde{\mathbf{N}}^s\|_{U_n^{i-1}} - 1 \right)^2 \right) + \frac{\epsilon}{2} (\tilde{\mathbf{N}}^s - \tilde{\mathbf{N}}^{s, \text{tr}})^2 \quad (3.55)$$

where the choice of A, with the same argument as in Subsection 3.3.1, still has no effect on the solution. Additionally, the last term with a very small and constant ϵ is explained in Subsection 3.3.5. Now, the minimization problem from Eq. (3.54) is computed by solving the stationary conditions:

$$\frac{\partial \pi_\Delta}{\partial \Delta\gamma} = 0, \quad \frac{\partial \pi_\Delta}{\partial \tilde{\mathbf{N}}^s} = \mathbf{0}, \quad \frac{\partial \pi_\Delta}{\partial \xi} = 0, \quad (3.56)$$

where in general only a subset of these three equations is involved, depending on the 'active' variables. The active set of variables is determined by the activation or yield criteria presented in the sequel (see Subsections 3.3.3 and 3.3.10).

3.3.3 Algorithmic yield criterion

In a given time step, \mathbf{C}^i will only change (i.e., $\mathbf{C}^i \neq \mathbf{C}_n^i \Leftrightarrow \Delta\gamma > 0$), if this decreases (minimizes) the potential. In order to decide whether the couple $(\Delta\gamma, \tilde{\mathbf{N}}^s)$ is activated for a given state $(\mathbf{F}, \mathbf{C}_r^i, \xi, \theta)$, we evaluate the algorithmic yield condition

$$f := -\frac{\partial}{\partial \Delta\gamma} \inf_{\tilde{\mathbf{N}}^s} \pi_\Delta(\mathbf{F}, \mathbf{C}^i(\hat{\mathbf{C}}^i(\Delta\gamma, \tilde{\mathbf{N}}^s)), \xi, \theta) \Big|_{\Delta\gamma \rightarrow 0}. \quad (3.57)$$

If $f > 0$, $\Delta\gamma$ is activated, i.e., $(\Delta\gamma, \tilde{\mathbf{N}}^s)$ are put into the active set. Otherwise, $\Delta\gamma = 0$ minimizes π_Δ , which corresponds to the case where $f \leq 0$. If $(\Delta\gamma, \tilde{\mathbf{N}}^s)$ are part of the active set, the related equations and concerning residuals are: $\partial \pi_\Delta / \partial \Delta\gamma \stackrel{!}{=} 0$ and $\partial \pi_\Delta / \partial \tilde{\mathbf{N}}^s \stackrel{!}{=} \mathbf{0}$. For a given \mathbf{F}, ξ and θ , we obtain the explicit form of f by variation of π : $\Delta\gamma = 0$ if

$\pi_\Delta(\mathbf{F}, \mathbf{C}^i, \xi, \theta)$ is minimized by $\mathbf{C}^i = \mathbf{C}_r^i$. In other words, $\Delta\gamma = 0$ is minimizer of π_Δ if

$$\delta\pi_\Delta \geq 0 \quad \forall \delta\mathbf{C}^i = \frac{\partial\mathbf{C}^i}{\partial\hat{\mathbf{C}}^i} : \underbrace{(2\delta\gamma\mathbf{N}^s)}_{\delta\hat{\mathbf{C}}^i}, \quad (3.58)$$

with shorthand notation

$$\mathbf{N}^s = \tilde{\mathbf{N}}^s / \|\tilde{\mathbf{N}}^s\|_{\mathbf{U}_n^{i-1}} \quad (3.59)$$

for an arbitrary $\tilde{\mathbf{N}}^s \neq 0$ and $\delta\gamma \geq 0$. Further, taking a closer look at the variation of π_Δ , we obtain (compare Appendix 3.7)

$$\delta\pi_\Delta = \frac{\partial\psi}{\partial\mathbf{C}^i} \Big|_{\mathbf{C}^i=\mathbf{C}_r^i} : \mathbb{P}_i(\mathbf{C}_r^i) : (2\delta\gamma\mathbf{N}^s) + \sqrt{\frac{2}{3}}\sigma^{\text{reo}}\delta\gamma \underbrace{\left\| \mathbf{U}_n^{i-1}\mathbf{N}^s\mathbf{U}_n^{i-1} \right\|}_{=1}. \quad (3.60)$$

Using the definition of the unimodular projector in Eq. (3.52), we obtain

$$\begin{aligned} \delta\pi_\Delta &\stackrel{(3.52)}{=} \left(\underbrace{\left(2\frac{\partial\psi}{\partial\mathbf{C}^i} \Big|_{\mathbf{C}^i=\mathbf{C}_r^i} - \frac{1}{3} \left(\mathbf{C}_r^i : 2\frac{\partial\psi}{\partial\mathbf{C}^i} \Big|_{\mathbf{C}^i=\mathbf{C}_r^i} \right) \mathbf{C}_r^{i-1} \right)}_{=: \text{DEV}_{\mathbf{C}_r^i} \left(2\frac{\partial\psi}{\partial\mathbf{C}^i} \Big|_{\mathbf{C}^i=\mathbf{C}_r^i} \right) = \mathbf{U}_r^{i-1} \left(2\mathbf{U}_r^i \frac{\partial\psi}{\partial\mathbf{C}^i} \Big|_{\mathbf{C}^i=\mathbf{C}_r^i} \mathbf{U}_r^i \right)' \mathbf{U}_r^{i-1}} : \mathbf{N}^s + \sqrt{\frac{2}{3}}\sigma^{\text{reo}} \right) \delta\gamma. \end{aligned} \quad (3.61)$$

With the definitions of \mathbf{N}^s in Eq. (3.59) and $\text{DEV}_{\mathbf{C}_r^i}(\bullet)$ in Eq. (3.61) as well as making use of the fact that \mathbf{U}_n^{i-1} and \mathbf{C}_r^{i-1} are coaxial, we can simplify Eq. (3.60) to

$$\delta\pi_\Delta = \left(\left(\mathbf{U}_n^i \text{DEV}_{\mathbf{C}_r^i} \left(2\frac{\partial\psi}{\partial\mathbf{C}^i} \Big|_{\mathbf{C}^i=\mathbf{C}_r^i} \right) \mathbf{U}_n^i \right) : \frac{\mathbf{U}_n^{i-1}\tilde{\mathbf{N}}^s\mathbf{U}_n^{i-1}}{\|\mathbf{U}_n^{i-1}\tilde{\mathbf{N}}^s\mathbf{U}_n^{i-1}\|} + \sqrt{\frac{2}{3}}\sigma^{\text{reo}} \right) \delta\gamma. \quad (3.62)$$

Now, since inequality (3.58) must hold for arbitrary symmetric $\tilde{\mathbf{N}}^s$, we get

$$\delta\pi_\Delta \geq - \left(\left\| \text{DEV}_{\mathbf{C}_r^i} \left(2\frac{\partial\psi}{\partial\mathbf{C}^i} \Big|_{\mathbf{C}^i=\mathbf{C}_r^i} \right) \right\|_{\mathbf{U}_n^i} - \sqrt{\frac{2}{3}}\sigma^{\text{reo}} \right) \delta\gamma \geq 0 \quad \forall \delta\gamma \geq 0. \quad (3.63)$$

Thus, the algorithmic yield criterion reads

$$f = \left\| \text{DEV}_{\mathbf{C}_r^i} \left(2\frac{\partial\psi}{\partial\mathbf{C}^i} \Big|_{\mathbf{C}^i=\mathbf{C}_r^i} \right) \right\|_{\mathbf{U}_n^i} - \sqrt{\frac{2}{3}}\sigma^{\text{reo}} \leq 0. \quad (3.64)$$

However, note that for $\xi = 0$, and thus $\mathbf{C}_r^i = \mathbf{C}_n^i$, we have a simplified yield criterion

$$\begin{aligned} f &= \left\| \mathbf{R}_n^i \left(2\mathbf{U}_n^i \frac{\partial\psi}{\partial\mathbf{C}^i} \Big|_{\mathbf{C}^i=\mathbf{C}_n^i} \mathbf{U}_n^i \right)' \mathbf{R}_n^{iT} \right\| - \sqrt{\frac{2}{3}}\sigma^{\text{reo}} \\ &\stackrel{(3.19)}{=} \left\| \left(\boldsymbol{\Sigma}^{\text{e,tr}} - \boldsymbol{\Sigma}^{\text{h,tr}} \right)' \right\| - \sqrt{\frac{2}{3}}\sigma^{\text{reo}}, \end{aligned} \quad (3.65)$$

where the superscript 'tr' denotes the trial state and $\mathbf{R}^i = \mathbf{F}^i \mathbf{U}^{i-1}$. This result again shows the consistency with the time-continuous theory (see Eq. (3.33)).

3.3.4 Algorithmic transformation criteria

In analogy to Subsection 3.3.3, the activation criteria for forward ('M', $\Delta\xi \geq 0$) and reverse ('A', $\Delta\xi < 0$) transformations read

$$g_M = -\frac{\partial}{\partial \xi} \pi_\Delta \Big|_{\Delta\xi \rightarrow 0^+}, \quad g_A = -\frac{\partial}{\partial \xi} \pi_\Delta \Big|_{\Delta\xi \rightarrow 0^-}, \quad (3.66)$$

where 0^+ and 0^- denote the right- and left-hand limit (recall that $\text{sg}(\xi)$ is an argument of Q in Eq. (3.41)).

3.3.5 Initial guess for $\tilde{\mathbf{N}}^s$

In general, the direction of inelastic flow $\tilde{\mathbf{N}}^s$ is determined by the minimization of π_Δ through a Newton scheme, which requires a reasonable initial guess when $\Delta\gamma$ is activated. Coincidentally, an analytical solution for $\Delta\gamma \rightarrow 0$ exists for $\tilde{\mathbf{N}}^s$:

$$\tilde{\mathbf{N}}^{s,\text{tr}} = \frac{\mathbf{C}_n^i \text{DEV}_{\mathbf{C}_r^i} \left(-2 \frac{\partial \psi}{\partial \mathbf{C}^i} \Big|_{\mathbf{C}^i = \mathbf{C}_r^i} \right) \mathbf{C}_n^i}{\left\| \text{DEV}_{\mathbf{C}_r^i} \left(-2 \frac{\partial \psi}{\partial \mathbf{C}^i} \Big|_{\mathbf{C}^i = \mathbf{C}_r^i} \right) \right\|_{\mathbf{U}_n^i}}. \quad (3.67)$$

The proof is given in Appendix 3.7.

It is noted that for $\Delta\gamma = 0$, it follows that $\mathbf{C}^i = \mathbf{C}_r^i$, which is independent of $\tilde{\mathbf{N}}^s$. That means that *any* choice of $\tilde{\mathbf{N}}^s$ minimizes π_Δ , i.e., there is no unique solution. For very small $\Delta\gamma$, the sensitivity of π_Δ with respect to $\tilde{\mathbf{N}}^s$ is also small, which can lead to a bad condition of the nonlinear equation system, which needs to be solved to minimize π_Δ . To stabilize the solution process, the last term in Eq. (3.55) is added to π , since for $\Delta\gamma \rightarrow 0$ it is known that $\tilde{\mathbf{N}}^s = \tilde{\mathbf{N}}^{s,\text{tr}}$ minimizes π_Δ . In theory, ϵ can be chosen arbitrarily small, in practice a finite value is necessary due to the limited numerical accuracy. Thus, the last term in Eq. (3.55) is used to 'guide' the algorithm towards the right solution when it is no longer able to find it by itself.

3.3.6 Computing derivatives of $\Delta\alpha$ for $\Delta\gamma \rightarrow 0$

The derivatives of $\Delta\alpha$ are numerically tough to obtain for $\Delta\gamma \rightarrow 0$, due to an almost zero denominator of $\tilde{\mathbf{M}}_{\mathbf{C}^i}$ (see Eq. (3.112) and preceding equations). To overcome this issue, we derived the derivatives of $\Delta\gamma$ with respect to ξ , $\tilde{\mathbf{N}}^s$ and $\Delta\gamma$ separately for $\Delta\gamma \rightarrow 0$, which are then numerically feasible to obtain. However, this seems to be a rather theoretical as we never observed the case that $\Delta\gamma$ was too close to zero to obtain the derivatives using

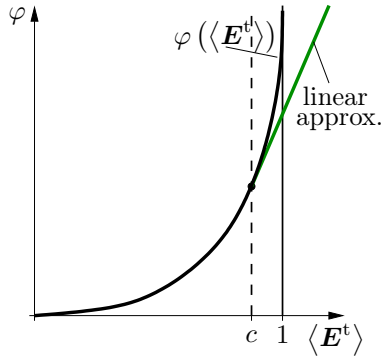


Fig. 3.2: Linear approximation of ψ_h .

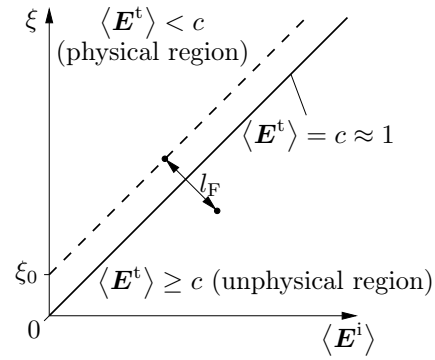


Fig. 3.3: Penalty function penalizing non-physical states.

the usual way shown in Appendix 3.8. Therefore, and for brevity, we do not discuss them here.

3.3.7 Regularization of ψ_h

Numerically, ψ_h is challenging, since with $\langle \mathbf{E}^t \rangle \rightarrow 1$, ψ_h and $\partial\psi_h/\partial\langle \mathbf{E}^t \rangle$ go to infinity. This is very demanding, because values of $\langle \mathbf{E}^t \rangle \geq 1$ can occur during the iterative solution process of the Newton scheme, leading to a loss of convergence. The term (compare to Eq. (3.21))

$$\varphi(\langle \mathbf{E}^t \rangle) = \frac{\langle \mathbf{E}^t \rangle^2}{1 - \langle \mathbf{E}^t \rangle^4} \quad (3.68)$$

is illustrated in Fig. 3.2 in black. To stabilize the solution process, we partially replace the function φ by a linear approximation in the region where $\langle \mathbf{E}^t \rangle > c$, where c is close to 1 (see Fig. 3.2 in green). This regularization approach is adopted from crystal plasticity, where it successfully improved the numerical treatment of the power law in Wulfinghoff and Böhlke [147]. In order to prevent the solution $\langle \mathbf{E}^t \rangle$ from taking nonphysical solutions beyond 1, we add a penalty-type energy as illustrated in Fig. 3.3. We penalize all states $(\langle \mathbf{E}^i \rangle, \xi)$ (see Fig. 3.3) below the dashed half-line starting at ξ_0 by the energy

$$\psi_p = \begin{cases} \frac{1}{2} H_p l_F^2 & \xi < \frac{\langle \mathbf{E}^i \rangle}{c} + \xi_0 \\ 0 & \text{else} \end{cases} \quad (3.69)$$

where l_F is the minimum distance of the point $(\langle \mathbf{E}^i \rangle, \xi)$ from the half-line and H_p is a large penalty parameter.

3.3.8 Viscosity of the martensite volume fraction ξ

To further improve the numerical robustness, we added an artificial viscosity-like term for the martensite volume fraction. Since it is a dissipative term, we add this term to the discretized dissipation potential:

$$\phi_{\Delta\eta} = \frac{1}{2\Delta t} \eta_\xi \Delta \xi^2, \quad (3.70)$$

where η_ξ is a small, positive constant.

3.3.9 Thermomechanical coupling

Like the displacements, the temperature is an unknown for our model. Both fields are coupled through the thermomechanical energy and through the dissipative terms, which lead to a heating of the material. Additionally, the temperature changes the material behavior and leads to thermal strains. Both fields are coupled using a potential-based monolithic approach similar to the one presented in Yang et al. [149]. Further, we assume Fourier's heat conduction law with heat conductivity κ . The thermomechanic quasistatic time-discrete potential then reads:

$$\pi_\Delta = \psi - (u_n - \theta s_n) + \frac{\theta}{\theta_n} \phi_\Delta - \frac{\Delta t}{2\theta_n} \kappa \|\text{Grad}(\theta)\|^2 + \Delta t \frac{\theta}{\theta_n^2} \kappa \|\text{Grad}(\theta_n)\|^2, \quad (3.71)$$

where we neglect the influence of body forces. The integral form Π_Δ of the local potential π_Δ is then given by

$$\Pi_\Delta = \int_{V_0} \pi_\Delta \, dV - \int_{\partial V_{0t}} \bar{\mathbf{t}} \cdot \mathbf{u} \, dS - \int_{V_{0Q}} \frac{\Delta t}{\theta_n} \bar{Q} \theta \, dS + \int_{V_0} \frac{\Delta t}{\theta_n} \theta w \, dV, \quad (3.72)$$

where w denotes the heat source and $\bar{\mathbf{t}}$ as well as \bar{Q} denote prescribed tractions and normal heat flux at the Neumann-type boundaries ∂V_{0t} and ∂V_{0Q} , respectively. Now, the classical weak form of the linear momentum balance is obtained by variation of Eq. (3.71) with respect to \mathbf{u} :

$$\delta_{\mathbf{u}} \Pi_\Delta = \int_{V_0} \boldsymbol{\tau} : \mathbf{d}_\delta \, dV - \int_{\partial V_{0t}} \bar{\mathbf{t}} \cdot \delta \mathbf{u} \, dS \stackrel{!}{=} 0, \quad (3.73)$$

with $\mathbf{d}_\delta = \text{sym}(\text{Grad}(\delta \mathbf{u}) \mathbf{F}^{-1})$. Likewise, we obtain the weak form of the energy balance by variation of Eq. (3.71) with respect to θ :

$$\begin{aligned} \delta_\theta \Pi_\Delta = \int_{V_0} \left(\left(\frac{\partial \psi}{\partial \theta} + s_n + \frac{1}{\theta_n} \phi_\Delta \right) \delta \theta - \frac{\kappa \Delta t}{\theta_n} \text{Grad}(\theta) \cdot \text{Grad}(\delta \theta) + \Delta t \frac{\delta \theta}{\theta_n^2} \kappa \|\text{Grad}(\theta_n)\|^2 \right. \\ \left. + \frac{\Delta t}{\theta_n} w \right) \delta \theta \, dV - \int_{V_{0Q}} \frac{\Delta t}{\theta_n} \bar{Q} \delta \theta \, dS \stackrel{!}{=} 0. \end{aligned} \quad (3.74)$$

Subsequently, we find $\mathbf{u}, \theta, \xi, \mathbf{C}^i$ by finding the solution to the saddle point problem

$$(\mathbf{u}, \theta, \xi, \mathbf{C}^i) = \arg \inf_{\mathbf{u} \in \kappa_{\mathbf{u}}} \sup_{\theta \in \kappa_{\theta}} \inf_{\mathbf{z}} \Pi_{\Delta}, \quad (3.75)$$

where $\kappa_{\mathbf{u}} = \{\mathbf{u} : \mathbf{u} = \bar{\mathbf{u}} \text{ on } \partial V_{0u}\}$ is the set of admissible displacements satisfying the Dirichlet boundary conditions imposed on the boundary ∂V_{0u} and $\kappa_{\theta} = \{\theta : \theta = \bar{\theta} \text{ on } \partial V_{0\theta}\}$ is the set of admissible temperatures satisfying the Dirichlet boundary conditions imposed on the boundary $\partial V_{0\theta}$. Additionally, \mathbf{z} is the vector of internal variables, i.e., $\mathbf{z} = (\xi, \mathbf{C}^i)$.

Manipulating Eq. (3.74) using the definition $s = -\partial\psi/\partial\theta$ as well as applying Gauss theorem, we obtain

$$\begin{aligned} \delta_{\theta} \Pi_{\Delta} = & \int_{V_0} \left(-\Delta s + \frac{1}{\theta_n} \phi_{\Delta} + \frac{\kappa \Delta t}{\theta_n} \text{Div}(\text{Grad}(\theta)) - \frac{\kappa \Delta t}{\theta_n^2} \text{Grad}(\theta) \cdot \text{Grad}(\theta_n) + \frac{\Delta t}{\theta_n} w \right. \\ & \left. + \frac{\kappa \Delta t}{\theta_n^2} \|\text{Grad}(\theta_n)\|^2 \right) \delta\theta \, dV + \int_{\partial V_{0Q}} -\Delta t \kappa \frac{\text{Grad}(\theta) \cdot \mathbf{N}}{\theta_n} \delta\theta \, dS - \int_{\partial V_{0Q}} \frac{\Delta t}{\theta_n} \bar{Q} \delta\theta \, dS = 0, \end{aligned} \quad (3.76)$$

where \mathbf{N} is the external normal on the boundary ∂V_0 in the reference configuration. To show consistency with the time-continuous theory, we take a look at the integrand over the volume integral in Eq. (3.76). By multiplying with $\theta_n/\Delta t$, we get

$$\theta_n \frac{\Delta s}{\Delta t} = \frac{1}{\Delta t} \phi_{\Delta} - \text{Div}(-\kappa \text{Grad}(\theta)) + w + \frac{\kappa}{\theta_n} \text{Grad}(\theta_n) \cdot (\text{Grad}(\theta_n) - \text{Grad}(\theta)). \quad (3.77)$$

Here, we want to note that this requires κ to be independent of the temperature θ . However, one possibility is to consider $\kappa = \kappa(\theta_n)$, which could circumvent this limitation.

Now, for $\Delta t \rightarrow 0$, we get (also see Eq. (3.35))

$$\theta \dot{s} = \phi - \text{Div}(-\kappa \text{Grad}(\theta)) + w. \quad (3.78)$$

Hence, we arrived at the energy balance (Eq. (3.35)) with Fourier's law $\mathbf{Q} = -\kappa \text{Grad}(\theta)$, which proves the consistency with the time-continuous theory. Further, the surface integrals in Eq. (3.76) imply the Neumann boundary condition $\bar{Q} = -\kappa \text{Grad}(\theta) \cdot \mathbf{N}$ on ∂V_{0Q} . Alternatively, one can include Robin-type boundary conditions into the model by replacing the integrand $\Delta t \bar{Q} \theta / \theta_n$ in Eq. (3.72) by $1/2 \Delta t h (\theta - \theta_s)^2 / \theta_n$, where h is the heat convection coefficient and θ_s is the temperature of the surrounding medium. In that case, the variation of Π_{Δ} yields the boundary condition $-\kappa \text{Grad}(\theta) \cdot \mathbf{N} = h(\theta - \theta_s)$ on ∂V_{0Q} .

3.3.10 Active set search

When solving the set of Equations (3.56), one has to decide which variables will evolve using the activation criteria given in Eqns. (3.64) and (3.66). They will then be put into the

active array of variables \mathcal{A} . If then, at a later state, an activation criterion is inactive, the variable is taken from \mathcal{A} again. This is done using the active set search algorithm outlined in Algorithm 2. The algorithm is structured as follows: First, if $\mathcal{A} = \emptyset$ or no variable from \mathcal{A}

Algorithm 2 Active set algorithm solving the minimization problem in Eq. (3.56).

```

loop
   $j \leftarrow j + 1$ 
  if no rescaling done in last step or  $\mathcal{A}^j = \emptyset$  then
    determine  $\mathcal{A}^{j+1}$  (see Algorithm 3)
  end if
  if  $\mathcal{A}^{j+1} = \mathcal{A}^j$  then
    compute  $\mathbf{c}^a, \frac{\partial \pi_\Delta}{\partial \theta}, \frac{\partial^2 \pi_\Delta}{\partial \theta^2}, \frac{\partial \tau}{\partial \theta}$  Eqns. (3.166), (3.172), (3.173)
    save history variables
    exit
  end if
   $\xi_0^* = \xi^j, \Delta \gamma_0^* = \Delta \gamma^j, \tilde{\mathbf{N}}_0^{s*} = \tilde{\mathbf{N}}^{sj}$ 
  loop
     $k \leftarrow k + 1$ 
    get  $\frac{\partial \pi_\Delta}{\partial \mathcal{A}^{j+1}}, \frac{\partial^2 \pi_\Delta}{\partial \mathcal{A}^{j+1,2}}$  based on  $\xi_{k+1}^*, \Delta \gamma_{k+1}^*$  and  $\tilde{\mathbf{N}}_{k+1}^{s*}$  Eqns. (3.87), (3.120)-(3.125)
    if  $k > \text{maxIter}$  or  $\|\partial \pi_\Delta / \partial \mathcal{A}^{j+1}\| < \text{tolNwtn}$  then
      exit
    else
      calculate  $\Delta \xi_{k+1}^*, \Delta \Delta \gamma_{k+1}^*, \Delta \tilde{\mathbf{N}}_{k+1}^{s*}$  if in  $\mathcal{A}^{j+1}$ 
      if  $\xi^j + \Delta \xi_{k+1}^* < 0$  then
        rescale  $\Delta \xi_{k+1}^*, \Delta \Delta \gamma_{k+1}^*$  and  $\Delta \tilde{\mathbf{N}}_{k+1}^{s*}$  s.t.  $\xi_{k+1}^* = \xi_{\min}^{\text{hard}}$ 
      end if
      update  $\xi_{k+1}^* = \xi_k^* + \Delta \xi_{k+1}^*, \Delta \gamma_{k+1}^* = \Delta \gamma_k^* + \Delta \Delta \gamma_{k+1}^*, \tilde{\mathbf{N}}_{k+1}^{s*} = \tilde{\mathbf{N}}_k^{s*} + \Delta \tilde{\mathbf{N}}_{k+1}^{s*}$ 
      according to  $\mathcal{A}^{j+1}$ 
    end if
  end loop
  if  $\Delta \xi_{k+1}^*$  goes in dir. not matching  $g_A/g_M$  then
    rescale  $\Delta \xi_{k+1}^*, \Delta \Delta \gamma_{k+1}^*$  and  $\Delta \tilde{\mathbf{N}}_{k+1}^{s*}$ 
    xiJustDeactivated  $\leftarrow$  true
  else
    xiJustDeactivated  $\leftarrow$  false
  end if
   $\xi^{j+1} = \xi^j + \Delta \xi_{k+1}^*, \Delta \gamma^{j+1} = \Delta \gamma^j + \Delta \Delta \gamma_{k+1}^*$  and  $\tilde{\mathbf{N}}^{sj+1} = \tilde{\mathbf{N}}^{sj} + \Delta \tilde{\mathbf{N}}_{k+1}^{s*}$ 
end loop

```

was rescaled (see the end of this subsection for an explanation of rescaling) during the last iteration, we need to evaluate the activation criteria. With them at hand, we determine our new set \mathcal{A}^{j+1} . The details for determining which variable to activate are given in Algorithm 3. Subsequently, we check if the active set changed from the last iteration. If this is not the case, the solution from the last iteration is confirmed as solution of the equation system. In that case, we compute the algorithmic tangent (see Appendix 3.10 for details), save the history variables and exit the material routine. Otherwise, we solve the minimization problem for

Algorithm 3 Algorithm deciding the priority when updating \mathcal{A} .

compute f Eq. (3.65)

if $\xi \notin \mathcal{A}^j$ **then**

 compute g_A **and** g_M Eq. (3.66)

end if

if $\mathcal{A}^j = \emptyset$ **then**

if $g_A < 0$ **and** **not** xiJustDeactivated **then**

 add ξ to \mathcal{A}^{j+1}

else if $g_M > 0$ **and** **not** xiJustDeactivated **then**

 add ξ to \mathcal{A}^{j+1}

else if $f > 0$ **then**

 add $\Delta\gamma$ to \mathcal{A}^{j+1}

end if

else if $\xi \in \mathcal{A}^j$ **and** $\Delta\gamma \notin \mathcal{A}^j$ **then**

if $f > 0$ **then**

 add $\Delta\gamma$ to \mathcal{A}^{j+1}

 compute new $\tilde{N}^{s, tr}$ Eq. (3.67)

end if

else if $\Delta\gamma \in \mathcal{A}^j$ **and** $\xi \notin \mathcal{A}^j$ **and** **not** xiJustDeactivated **then**

if $g_A < 0$ **then**

 add ξ to \mathcal{A}^{j+1}

else if $g_M > 0$ **then**

 add ξ to \mathcal{A}^{j+1}

end if

end if

the current active set \mathcal{A}^{j+1} . This is done using a Newton scheme. First, we compute our residual and stiffness concerning active set \mathcal{A}^{j+1} , i.e., only derivatives with respect to the active set are computed. If the maximum iterations are exceeded or the norm of the residual is lower than the tolerance, we exit the loop. Otherwise, we check if updating ξ^{j+1} would result in a negative ξ when applying the Newton step. If this is the case, we rescale, i.e., multiply the increments by a scalar such that ξ_{k+1}^* is set to ξ_{\min}^{hard} , which is 5×10^{-5} in this work. Subsequently, we compute $\Delta \xi_{k+1}^*$, $\Delta \Delta \gamma_{j+1}^*$ and $\Delta \tilde{\mathbf{N}}_{j+1}^{s*}$. In fact, we do not update ξ^{j+1} , $\Delta \gamma^{j+1}$ or $\tilde{\mathbf{N}}^{sj+1}$ yet. We update them only after exiting the Newton scheme and being sure that ξ incremented in a direction matching g_A and g_M , which were calculated before the Newton loop. If, due to a bad starting solution, this is not the case, we rescale the Newton step such that $\xi^{j+1} = \xi_n$. If this happens, we make sure that in the next iteration of the active set algorithm, we won't activate ξ again.

3.4 Numerical results

In this section, the previously presented model is tested for different examples. First, to test the model's time convergence behavior and to show the superelastic as well as martensite reorientation behavior at high and low temperatures, respectively, thermomechanical Gauss point evaluations are conducted. Finally, a full actuator model is simulated. To cope with the thin structures occurring in the actuator, the SMA model at hand is embedded into a hexahedral element formulation with reduced integration and hourglass stabilization for the displacement, while the heat conduction terms are fully integrated.

3.4.1 Gauss point evaluations

For the Gauss point evaluations, thermal expansion as well as the transformation induced volume change, which is discussed in Section 3.2.3, are neglected for simplicity. The simulations are conducted using a reduced integration hexahedral element with hourglass stabilization, which is embedded into the finite element program FEAP [127]. The material constants used in the Gauss-point are given in Table 3.1. Here, it is noted that θ_{refA} and θ_{refM} are chosen such that $\Delta V = 0.47\%$ (see Eq. (3.23)), matching the findings reported in Potapov et al. [97]. The numerical parameters are summarized in Table 3.2. Figure 3.4 shows a tensile test at a temperature of 270 °C for 49 and 40000 time steps. Clearly, convergence with regard to time step width is not an issue, as both time step widths yield accurate results.

To demonstrate that the model results indeed do not depend on the numerical parameters, we compare results for two different sets of numerical parameters in a tensile test. The first set of numerical parameters is the one given in Table 3.2, which is used in the remainder of this paper. The second set is defined in Table 3.3 just for this comparison. The results are shown in Figure 3.5. Clearly, the numerical parameters do not have a noticeable effect on the results.

Tab. 3.1: Material parameters.

Elastic constants						
λ_{A} [GPa]	λ_{M} [GPa]	μ_{A} [GPa]	μ_{M} [GPa]			
23.076	13.461	34.615	20.192			
Dissipation potential parameters						
θ_0 [°C]	M_{s} [°C]	M_{f} [°C]	A_{s} [°C]	A_{f} [°C]	σ^{reo} [MPa]	
80	70	50	120	140	160	
Chemical and hardening parameters						
Δs^{AM} [$\frac{\text{MPa}}{\text{K}}$]	E^{int} [MPa]	k [-]				
0.383	30	0.12				
Thermal constants						
α_{M} [$0 \times 10^{-6} \frac{1}{\text{K}}$]	α_{A} [$0 \times 10^{-6} \frac{1}{\text{K}}$]	c [$\frac{\text{mJ}}{\text{mm}^3 \text{K}}$]	κ [$\frac{\text{W}}{\text{m}}$]	θ_{refM} [°C]	θ_{refA} [°C]	
6.6	11.0	2.9	8.6	588.28	500.0	

Tab. 3.2: Numerical parameter set 1 for simulation.

Numerical parameters:						
c [-]	H_p [MPa]	A [MPa]	B [-]	ξ_0 [-]	ϵ [MPa]	η_ξ [$\frac{\text{MPa}}{\text{s}}$]
0.99	10^8	10^5	1	10^{-4}	10^{-4}	0.01

Tab. 3.3: Numerical parameter set 2, only for comparison in Fig. 3.5.

Numerical parameters:						
c [-]	H_p [MPa]	A [MPa]	B [-]	ξ_0 [-]	ϵ [MPa]	η_ξ [$\frac{\text{MPa}}{\text{s}}$]
0.98	10^7	10^6	1	10^{-3}	10^{-3}	10^{-3}

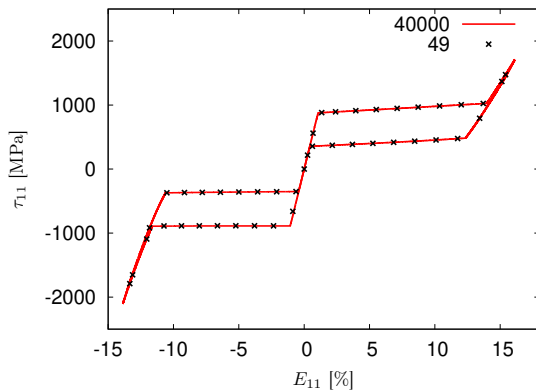


Fig. 3.4: Tensile test for 49 and 40000 time steps. No error due to large load steps is visible.

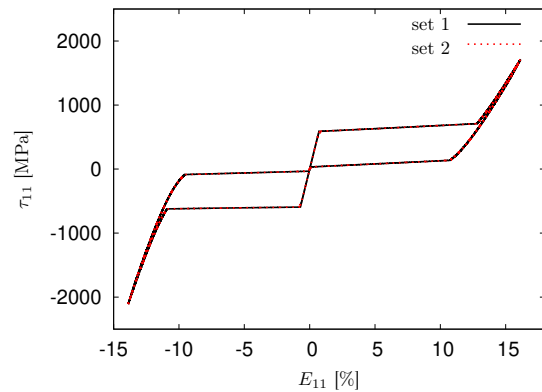


Fig. 3.5: Investigation of the numerical parameters influence (set 1 is given in Table 3.2, set 2 in Table 3.3).

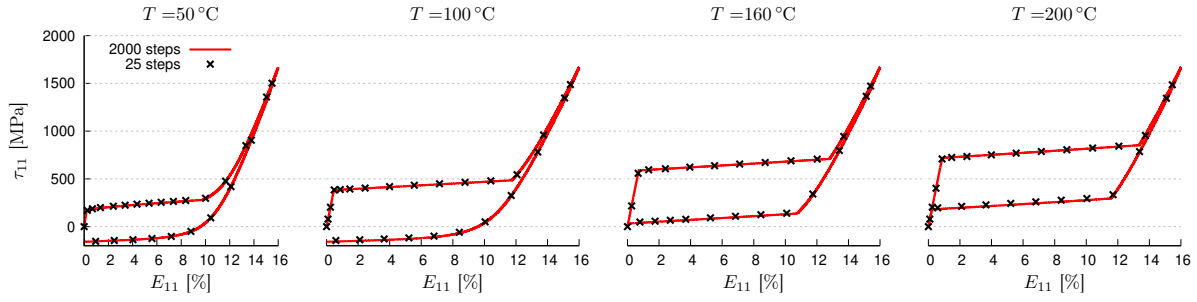


Fig. 3.6: Tensile test at various temperatures.

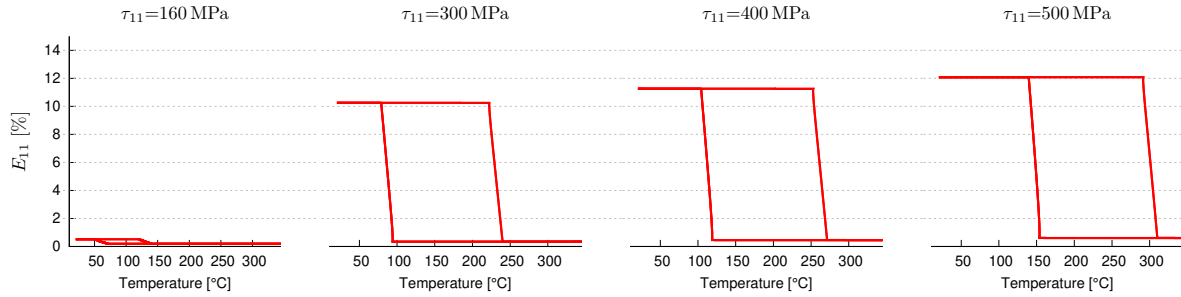


Fig. 3.7: Thermal cycling tests with different stresses (160 MPa, 300 MPa, 400 MPa and 500 MPa).

Next, we modeled tensile tests at 50 °C, 100 °C, 160 °C and 200 °C. The resulting stress-strain curves are shown in Fig. 3.6. For low temperatures, i.e. 50 °C and 100 °C, the model predicts a remaining martensite reorientation. On the other hand, for rather high temperatures, the model captures superelastic material behavior when unloading. However, we want to emphasize that these results do not consider any damage or plastic deformations which might already occur.

Finally, the model captures the shape memory effect well. This is shown in Fig. 3.7, where we applied different stresses (160 MPa, 300 MPa, 400 MPa and 500 MPa) and then started a thermal cycle. For 160 MPa, one obtains only twinned martensite, i.e., $\mathbf{E}^t = \mathbf{0}$ in the context of this work. Therefore, one only gets a small hysteresis due to the different elastic constants of austenite and martensite. For the larger prestresses, the austenite is transformed into detwinned martensite, i.e., $\langle \mathbf{E}^t \rangle \neq 0$. This leads to the typical shape memory effect.

3.4.2 Plate with a hole

In this example, we simulate a plate with a cylindrical hole (see Fig. 3.8), which is at first loaded by a traction $\bar{\mathbf{t}}$, then unloaded and subsequently heated, which lets it recover the initial shape. The material and numerical parameters are unchanged from Subsection 3.4.1, except the reference temperatures are now $\theta_{\text{refA}} = 80$ °C and $\theta_{\text{refM}} = 87.66$ °C. Due to the symmetry, using appropriate symmetry conditions, only one-eighth of the entire plate is simulated.

During the loading by the traction, the temperature at the upper and lower end is held

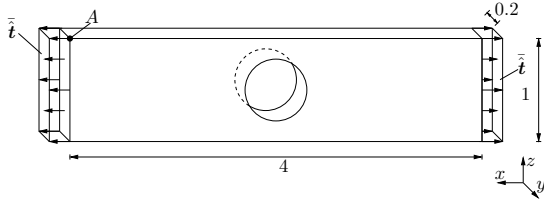


Fig. 3.8: Sketch of the plate with a hole subjected to tension (dimensions in mm).

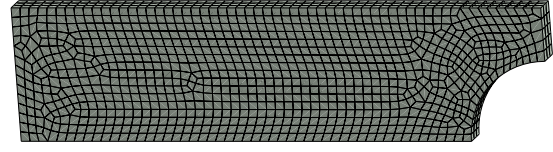


Fig. 3.9: Mesh of the plate with a hole with 2907 elements.

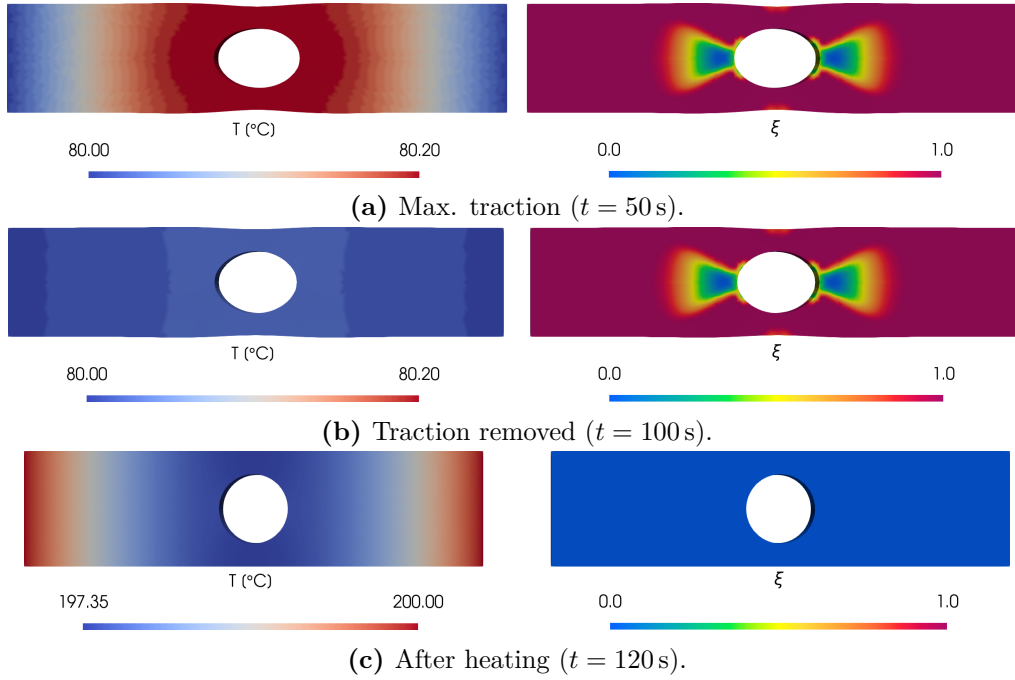


Fig. 3.10: Temperature and martensite volume fraction at different loading stages (see text).

constant at the initial temperature of $\bar{\theta} = 80^\circ\text{C}$ ($M_s < \bar{\theta} < A_s$). Additionally, in the initial state, the plate is fully austenite. First, the traction \bar{t} is increased linearly to a maximum of 330 MPa in longitudinal direction over the duration of 50 s. Subsequently, the traction \bar{t} is decreased to zero over the duration of 50 s. Finally, the temperature at both ends of the plate is increased to 200°C over the course of 20 s.

The mesh used in the eighth of the plate consists of 2907 uniformly distributed elements and is shown in Fig. 3.9. For the entire simulation, 293 time steps were required, taking a total CPU-time of 2524 s.

At first, during the loading, almost the entire plate is transformed to martensite. The plate at maximum traction is shown in Figure 3.10a. Here, the E_{xx} strain at the edge of the hole reaches roughly up to 18% (see Fig. 3.11). Additionally, the plate heats up slightly due to the latent heat and mechanical dissipation (Fig. 3.10a). However, most of the heat is conducted out of the plate at the temperature Dirichlet boundaries due to the small size of the plate

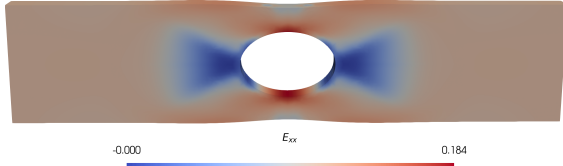


Fig. 3.11: Strain E_{xx} at max. traction ($t = 50$ s).

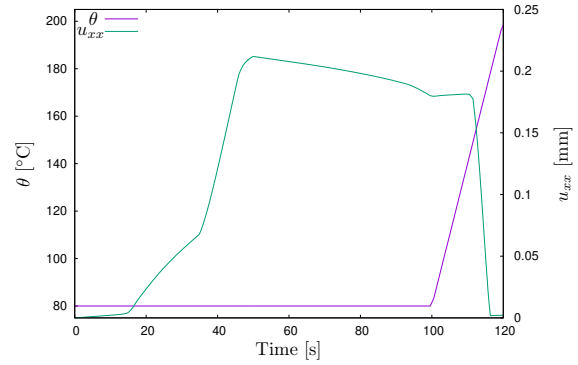


Fig. 3.12: Temperature and displacement at point A (see Fig. 3.8) over time.

and the long simulated time⁵. Subsequently, after unloading, the plate does not transform back, which is shown in Fig 3.10b. Finally, when increasing the temperature at both ends, the martensite is transformed back to austenite, leading to a recovery of the initial shape (see Fig. 3.10c). Further, Fig. 3.12 shows the temperature and displacement of point A (see Fig. 3.8) over time. Here, the increase in slope of the displacement at roughly 15 s is the point where we have a forward transformation in the vicinity of the hole. Then, the second change in slope of the displacement at roughly 36 s is the rest of the plate starting to undergo forward transformation. At 50 s, when releasing the traction, the displacements decrease. Then again, due to thermal expansion, the displacement increases slightly after increasing the temperature at the ends at 100 s. However, when the temperature reaches the backward transformation temperatures, the material almost recovers the initial shape. The non-zero displacement at $t = 120$ s stems from thermal expansion and the ΔV effect.

Overall, the examples have been chosen such that the stresses in the vicinity of the hole reach unphysical values, at which a real material is expected to show irrecoverable strains or even fracture. Thus, this simulation rather serves as an example showing that the implementation is able to find a solution, even under severe loads.

3.4.3 Finite element actuator model

In this section, we model a bistable shape memory microactuator using in the finite element program FEAP [127]. The actuator concept was published by Winzek et al. [140] for large structures and is built with three layers: a bottom layer of molybdenum, a middle layer of NiTiHf and a top layer of polymethyl methacrylate (PMMA) (Winzek et al. [140] used PMMA layers on both sides). For simplicity, a hexahedral element formulation with reduced integration and hourglass stabilization for the displacements is used for all materials. The material and numerical parameters for the shape memory alloy remain unchanged (see

⁵A separate simulation with "zero heat flux" boundary conditions led to a temperature increase of approximately 50 °C.

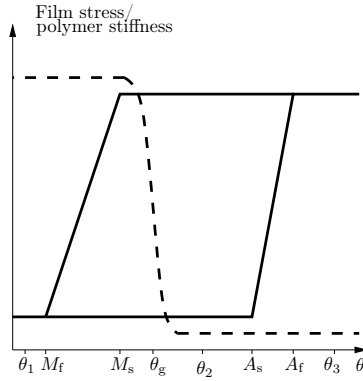


Fig. 3.13: Schematic of the interlaced hysteresis of the NiTiHf-molybdenum films and the polymer. Dashed line for the polymer stiffness, solid line for the SMA film stress.

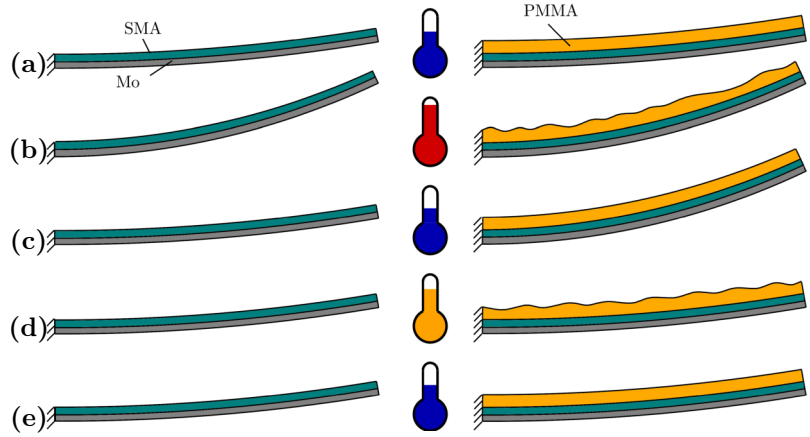


Fig. 3.14: A bimorph with SMA and molybdenum (left) in comparison to the bistable actuator with an additional layer of PMMA (right). The temperatures are indicated by different thermometer colorings.

Tables 1 and 2).

Actuating principle

The actuator works through interlacing of the large hysteresis of NiTiHf with the polymers hysteresis. This is shown in Fig. 3.13. It is actuated by joule heating to specific temperatures and cooling back to room temperature. Depending on the heat cycle, the martensite state or austenite state is held in place by the polymer. This is understood best when looking at the following example, which is depicted in Fig. 3.14, where we neglected the CTE of PMMA for simplicity. Here, we compare a bimorph of SMA and molybdenum on the left to the proposed actuator on the right. When heating from ambient temperature (blue, Fig. 3.14a) to a temperature above A_f (red), the glass transition temperature is reached, which drastically decreases the stiffness of the PMMA. Subsequently, the SMA in the bimorph as well as the bistable actuator reach reverse transformation temperatures, which makes them bend up (Fig. 3.14b). Now, upon cooling, the actuator reaches θ_g , which 'freezes' its current shape. Thus, when subsequently reaching forward transformation temperatures, unlike the bimorph, it will not revert into the original shape at room temperature (Fig. 3.14c). Now, we can

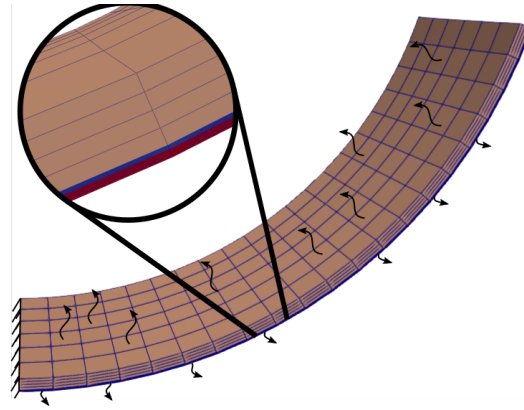


Fig. 3.15: Zoom in into the layers of the actuator (molybdenum in red, NiTiHf in blue and PMMA in orange).

heat the actuator to a temperature above θ_g but below A_s (orange in Fig. 3.14d) to soften the polymer, which makes the actuator adopt the deformation of the bimorph. When now cooling back to room temperature again, the actuator is in its initial stable state again and the actuation cycle could be started again. The two stable states at room temperature are given by Fig. 3.14c and e.

For the actuation behavior, it is important to have a homogeneous temperature profile in the actuator. Therefore, a wing-shape actuator developed in Arivanandhan et al. [4] is used to obtain rather homogeneous temperatures in the double beam cantilever. To rapidly optimize the actuators properties, it turned out that building and modeling only the beam section of the actuator is advantageous. Therefore, we simulated just the beam part with suitable boundary conditions instead of simulating the entire wing structure including the attached wafer material. This thin film actuator with a zoom into the layers is shown in Fig. 3.15, where the Dirichlet boundary condition for the displacements and the Robin boundary condition for the temperature are indicated as well.

Polymer and molybdenum model

Since the interest in modeling the actuator lies rather in the states at θ_1 , θ_2 and θ_3 , and less in the states between them, we chose a very simple polymer model. It is governed by a thermally coupled viscoelastic Maxwell model for finite strains (Young's modulus and Poisson's ratio are 500 MPa and 0.4, respectively), where the viscosity is high (10^7 MPa s) at low temperatures and low (1 MPa s) at high temperatures. Therefore, it has almost no stiffness at high temperatures while it behaves almost elastically at low ones. Additionally, the glass transition temperature is given by $\theta_g = 77^\circ\text{C}$. The molybdenum is modeled with a thermally coupled Neo-Hookean elastic model. For the molybdenum, Young's modulus, Poisson's ratio, the thermal expansion coefficient as well as the reference temperature are $E = 65 \times 10^3$ MPa, $\nu = 0.31$, $\alpha = 5 \times 10^{-6}$ K and $\theta_{\text{ref}} = 500^\circ\text{C}$, respectively.

Mesh convergence

The actuator considered is shown in Figure 3.15. It is clamped on the left side. It has a length of 10mm, a width of 5mm. The layer thicknesses of the molybdenum and TiNiHf are for this work 20μm and 10μm, respectively. The polymer layer thickness is 160μm for now, before different layer thicknesses are compared in Section 3.4.3. For the initial conditions, we assume zero displacements as well as a temperature of 500 °C, i.e., the temperature at which the actuator is thermally annealed in a flat state. Furthermore, we assume the material to be in its austenite state at $t = 0$, which directly implies that initially $\mathbf{C}^i = \mathbf{I}$. Subsequently, it is cooled down to room temperature at 20 °C, which bends the actuator due to the mismatch in the coefficients of thermal expansions and difference in cell volume between the austenite and martensite phase.

The heating cycle is realized through applying a heat source in the Mo and NiTiHf material. The heat source magnitude is modeled by a sine function, which is cut off when below 0. It is given by

$$w_h = \max(2250(\sin(0.03t) - 0.6), 0) \frac{\text{mW}}{\text{mm}^3}, \quad w_l = \max(0.65 \cdot 2250(\sin(0.03t) - 0.6), 0) \frac{\text{mW}}{\text{mm}^3},$$

for the higher and lower heat cycles, respectively.

At the top and at the bottom face of the actuator a Robin boundary condition is applied, which cools the thin film to room temperature over time. Due to their small areas, the heat transfer at the lateral faces is neglected. The surrounding air's temperature is 20 °C while the convective heat transfer coefficient is assumed to be $70 \frac{\text{W}}{\text{m}^2 \text{K}}$ in accordance with Kohl et al. [54].

For these thin structures, a sufficiently fine mesh is crucial to obtain converged results. Therefore, we conducted convergence studies with regard to the necessary elements in each layer and the amount of elements needed over the length and width of the actuator. First, it turned out that due to the bending deformation of the thin film, it is sufficient to only use one element over the width of the actuator. Then, the convergence with respect to the amount of elements over the length is tested using four elements over the thickness for each material layer. The results are shown in Fig. 3.16, where the stroke and temperature of the SMA at the tip is plotted over time. Here, only the temperature for 20 elements over the length is shown, since there was virtually no difference in temperature for the different discretizations. Further, we concluded that using 20 elements over the length leads to an acceptable error in stroke while the main features of the actuator are conveyed well. Additionally, roughly 160 time steps were used for one actuation cycle, e.g., from 0s to 180s in Fig. 3.16. Next, the convergence with regards to the elements used over the depth for each material was studied. The results are depicted in Figure 3.17. Here, for the number of elements over the SMA thickness, one element is already enough to obtain results with an acceptable error. For the molybdenum and polymer layer, one needs at least two elements to obtain converged results. To illustrate the importance of modeling the ΔV effect, Fig. 3.18 shows the stroke and tem-

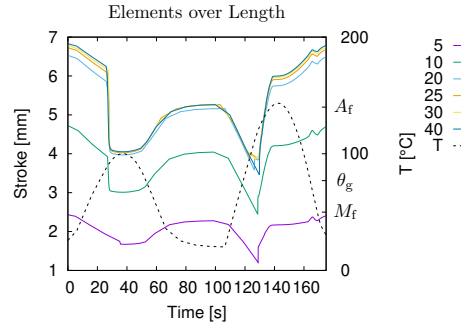


Fig. 3.16: Stroke (left y-axis) and temperature (right y-axis, dashed in black) over time for different amounts of elements over the length of the actuator.

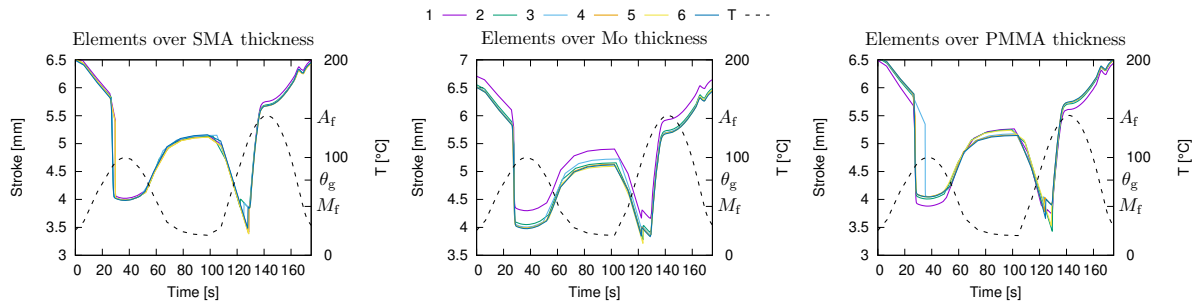


Fig. 3.17: Stroke (left y-axis) and temperature (right y-axis, dashed in black) over time for different amounts of elements over the thickness of each material layer.

perature over time for ΔV being 0 and 0.49%. Here, one can see that the volume change contributes to large parts of the achieved stroke between the two stable states at room temperature.

Influence of the polymer thickness

The polymer layer thickness plays a key role for this actuator design – if it is too thin, the high temperature shape can not be held at room temperature by the polymer layer. On the other hand, it should not be too thick, since body forces and manufacturing problems come up in that case. Further, since the polymer insulates the actuator thermally to one side, thinner polymer layers lead to the possibility of faster actuation cycles. Therefore, we tested several polymer layer thicknesses, the results are shown in Fig. 3.19. First, at ① (see Fig. 3.19), the actuator is heated. At first, the larger CTE of the polymer makes the actuator bend down. Then, when reaching θ_g , the polymer softens up and the actuator relaxes. Subsequently, at ②, the actuator is cooled down to room temperature by the surrounding air. The polymer hardens again and the device reaches stable configuration ④ at room temperature. Now, when heating again, the actuator bends down due to the large CTE of the PMMA at ③. As soon as the temperature reaches θ_g , the polymer softens up. Afterwards, at ④ the shape memory alloy reaches A_f and the martensite is transformed back to austenite, which makes

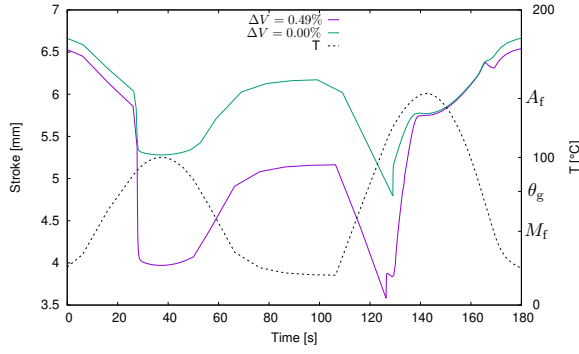


Fig. 3.18: Stroke and temperature over time for $\Delta V = 0.49\%$ and $\Delta V = 0.0\%$.

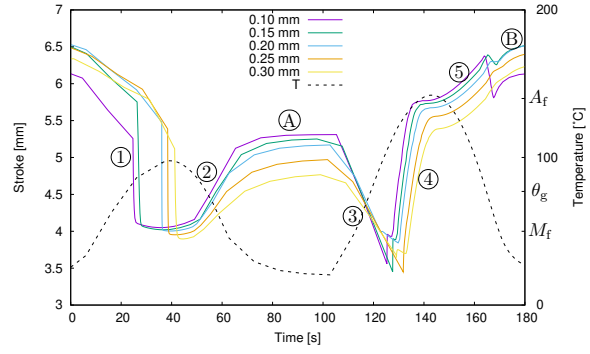


Fig. 3.19: Stroke (left y-axis) and temperature (right y-axis, dashed in black) over time for different PMMA layer thicknesses.

the actuator deflect up. Finally, when we remove the source at ⑤, the polymer hardens before M_f is reached. Thus, depending on the polymer thickness, the polymer may hold the actuators shape or release some of it's stroke. Finally, the actuator reaches it's second stable state ⑥ at room temperature, which is also it's initial state.

For many actuator designs, one wants to maximize the achievable stroke while keeping the device as small and power efficient as possible. Therefore, we try to maximize the difference in stroke between states ④ and ⑥. In turn, one must choose a sufficient polymer thickness, such that the actuator does not release too much stroke at ⑤. For example, 0.1 mm of PMMA is not thick enough to hold the shape, while 0.3 mm holds the shape perfectly (see Fig. 3.19 at ⑤). Depending on the actuation frequency in mind and necessary achievable stroke, either value in between might be a suitable choice.

3.5 Summary and outlook

In this paper, a new thermomechanical shape memory alloy model for finite strains is presented. It uses a projection method to fulfill the incompressibility constraint for the inelastic stretches. Further, the model is realized in the generalized standard material formulation being extended to thermomechanics. The optimization of the global, incremental mixed thermomechanical potential by variation yields the mechanical balance principles as well as the evolution equations of the internal variables and boundary condition integrals. The presented model employs a thermal strain formulation for the shape memory alloy which allows to describe the transformation induced volume changes found in some shape memory alloys. Using a logarithmic strain formulation, a finite strain dissipation potential incorporates vanishing inelastic strains upon austenite transformation in a manner consistent with the time-continuous case. Additionally, yield and transformation criteria as well as the algorithmically consistent tangent for the coupled problem are given and discussed. Due to numerical difficulties, a regularization of the hardening energy term is implemented.

The numerical results show, that the model is capable of producing reasonable results in tensile tests. The dependency on the temperature at which the tensile tests are carried out is captured accurately. Furthermore, it enables the solution of thin film problems.

Using the model, we are now able to estimate viable layer thicknesses and device sizes as well as fitting joule heating parameters. Additionally, it turned out that only applying polymer to one side increases the maximum actuation frequency while also introducing more deflection between the stable states and requiring more power. Moreover, we found that the ΔV effect as well as the thermal expansion of the SMA, due to the design being very sensitive to volume changes, is fairly important for the actuators stroke and therefore needs to be modeled accurately as well. Furthermore, the model results show that the bistability at room temperature, which is enabled by the polymer locking in the deformation, is achievable for the shown actuator concept.

In the future, it remains interesting to find ways of solving the model problem without any penalty terms, which are not reasoned physically and are numerically challenging to deal with. A possible solution to this problem could be to make use of the Fischer-Burmeister [24] complementary function. In fact, this has been done in Auricchio et al. [6] for shape memory alloys and in Brepols et al. [11] for damage-plasticity with great success. Furthermore, there is still space for improvements to the algorithms convergence behavior which could increase the speed of the proposed material model, especially under severe loading conditions. Additionally, an inclusion of functional fatigue properties, e.g., a shifting of the TiNiHf transformation temperatures over several transformation cycles into the model would enable to predict the long-term behavior of the actuators.

Acknowledgment

Financial support of subproject A3 *Cooperative Actuator Systems for Nanomechanics and Nanophotonics: Coupled Simulation* of the Priority Programme SPP 2206 by the German Research Foundation (DFG) (Grant WU847/3-1) is gratefully acknowledged by the authors.

Statements and Declarations

The authors have no competing interests to declare that are relevant to the content of this article.

Appendix

3.6 Appendix A: Discussion of the dissipation potential ϕ_Δ

In this section, we have a close look at the term

$$\left\| \mathbf{D}^i - \frac{\dot{\xi}}{\xi} \boldsymbol{\varepsilon}^i \right\| = \left\| \mathbf{D}^i - \frac{\dot{\xi}}{2\xi} \ln \mathbf{b}^i \right\| \quad \text{for } \dot{\xi} \leq 0 \quad (3.79)$$

from Eq. (3.26). If this term is zero, then $\mathbf{b}^i \rightarrow \mathbf{I}$ for $\xi \rightarrow 0$.

Proof. Consider $\xi(t) = \xi_0 (1 - \frac{t}{T})$ and $\mathbf{D}^i = \frac{\dot{\xi}}{\xi_0} \boldsymbol{\varepsilon}_0^i = -\frac{1}{2T} \ln(\mathbf{b}_0^i)$.

Then

$$\mathbf{b}^i(t) = \exp(2t\mathbf{D}^i)\mathbf{b}_0^i = \exp\left(-\frac{t}{T} \ln(\mathbf{b}_0^i)\mathbf{b}_0^i\right) = (\mathbf{b}_0^i)^{1-\frac{t}{T}} \rightarrow \mathbf{b}^i(T) = \mathbf{I}. \quad (3.80)$$

While this only shows that \mathbf{b}^i goes to \mathbf{I} for one special $\xi(t)$, one could easily rescale the time to obtain arbitrary $\xi(t)$, for which $\mathbf{b}^i \rightarrow \mathbf{I}$ with $\xi \rightarrow 0$. \square

3.7 Appendix B: Proof of $\tilde{\mathbf{N}}^{\text{s, tr}}$ minimizing π_Δ for $\Delta\gamma \rightarrow 0$

Proof. We are looking for $\tilde{\mathbf{N}}^{\text{s}}$ minimizing π_Δ for $\Delta\gamma \rightarrow 0$:

$$\delta\Pi = \frac{\partial\psi}{\partial\mathbf{C}^i} \Big|_{\mathbf{C}^i=\mathbf{C}_r^i} : \delta\mathbf{C}^i = \frac{\partial\psi}{\partial\mathbf{C}^i} \Big|_{\mathbf{C}^i=\mathbf{C}_r^i} : \mathbb{P}_i : (2\Delta\gamma\delta\mathbf{N}^{\text{s}}) \stackrel{!}{=} 0, \quad (3.81)$$

where the variation is done exclusively with respect to $\tilde{\mathbf{N}}^{\text{s}}$. Here, the dissipation potential ϕ_Δ is neglected because

$$\delta\phi_\Delta = \delta \left(\sqrt{\frac{2}{3}}\sigma^{\text{reo}} \left\| \frac{1}{2}\Delta_r \mathbf{C}^i \right\|_{\mathbf{U}_n^{i-1}} \right) = \frac{1}{2}\delta \left(\sqrt{\frac{2}{3}}\sigma^{\text{reo}} \left\| \hat{\mathbb{M}}^{i-\frac{1}{3}} (\mathbf{C}_r^i + 2\Delta\gamma\mathbf{N}^{\text{s}}) - \mathbf{C}_r^i \right\|_{\mathbf{U}_n^{i-1}} \right). \quad (3.82)$$

Since π_Δ is constructed such that it is minimized by $\hat{\mathbb{M}}^i = 1$, we constrain $\tilde{\mathbf{N}}^{\text{s}}$ to be in the subspace of symmetric 2nd-order tensors satisfying $\det(\hat{\mathbf{C}}^i) = \det(\mathbf{C}_r^i + 2\Delta\gamma\mathbf{N}^{\text{s}}) = 1$ (i.e., we only search this subspace for the minimizer). Hence, $\hat{\mathbb{M}}^i = 1$ and

$$\delta\phi_\Delta = \delta \left(\sqrt{\frac{2}{3}}\sigma^{\text{reo}} \underbrace{\left\| \mathbf{U}^{i-1} \mathbf{N}^{\text{s}} \mathbf{U}^{i-1} \right\|_1}_{1} \delta\Delta\gamma \right), \quad (3.83)$$

which is independent of $\tilde{\mathbf{N}}^s$.

With the definition of the tensor

$$\mathbb{P}_n = \frac{\partial \mathbf{N}^s}{\partial \tilde{\mathbf{N}}^s} = \frac{1}{\|\tilde{\mathbf{N}}^s\|_{\mathbf{U}_n^{i-1}}} \left(\mathbb{I}^s - \mathbf{N}^s \otimes \mathbf{C}_n^{i-1} \mathbf{N}^s \mathbf{C}_n^{i-1} \right) \quad (3.84)$$

Eq. (3.81) becomes

$$\delta \pi = \Delta \gamma \mathbf{U}_n^i \left(\mathbb{P}_i^\top (\mathbf{C}_r^i) : \left(2 \frac{\partial \psi}{\partial \mathbf{C}^i} \right) \Big|_{\mathbf{C}^i = \mathbf{C}_r^i} \right) \mathbf{U}_n^i : \delta \underbrace{\frac{\mathbf{U}_n^{i-1} \tilde{\mathbf{N}}^s \mathbf{U}_n^{i-1}}{\|\mathbf{U}_n^{i-1} \tilde{\mathbf{N}}^s \mathbf{U}_n^{i-1}\|}}_{=: \mathbf{N}^i} \stackrel{!}{=} 0. \quad (3.85)$$

Since \mathbf{N}^i is normalized (i.e., its norm is one), the solution reads

$$\mathbf{N}^i \stackrel{(3.61)}{=} \frac{\mathbf{U}_n^i \text{DEV}_{\mathbf{C}_r^i} \left(-2 \frac{\partial \psi}{\partial \mathbf{C}^i} \Big|_{\mathbf{C}^i = \mathbf{C}_r^i} \right) \mathbf{U}_n^i}{\left\| \text{DEV}_{\mathbf{C}_r^i} \left(-2 \frac{\partial \psi}{\partial \mathbf{C}^i} \Big|_{\mathbf{C}^i = \mathbf{C}_r^i} \right) \right\|_{\mathbf{U}_n^i}}, \quad (3.86)$$

which leads to $\tilde{\mathbf{N}}^{s, \text{tr}}$ in Eq. (3.67). \square

3.8 Appendix C: First derivatives of π_Δ

To minimize the potential using the active set algorithm including the Newton scheme, we need to obtain first and second order derivatives of the potential with respect to the set of internal variables and temperature. The differential of the discretized potential with respect to the internal variables (\mathbf{u} and θ are fixed) is given by

$$\begin{aligned} d\pi_\Delta &\stackrel{(3.53), (3.59)}{=} + \left(\frac{\partial(\psi_e + \psi_h + \psi_r + \psi_p)}{\partial \hat{\mathbf{C}}^i} : \frac{\partial \mathbf{C}_r^i}{\partial \Delta \xi} + \sigma^{\text{reo}} \frac{\partial \Delta \alpha}{\partial \Delta \xi} + \frac{\partial(\psi_e + \psi_h + \psi_c + \psi_p)}{\partial \Delta \xi} \right. \\ &\quad \left. + Q(\xi_n, \Delta \xi) + H(\text{sg}(\Delta \xi)) \Delta \xi \right) d\Delta \xi \\ &\quad + \left(\frac{\partial(\psi_e + \psi_h + \psi_r + \psi_p)}{\partial \hat{\mathbf{C}}^i} : (2\mathbf{N}^s) + \sigma^{\text{reo}} \frac{\partial \Delta \alpha}{\partial \Delta \gamma} \right) d\Delta \gamma \\ &\quad + \left(\mathbb{P}_n^\top : \left(2\Delta \gamma \frac{\partial(\psi_e + \psi_h + \psi_r + \psi_p)}{\partial \hat{\mathbf{C}}^i} \right) + \sigma^{\text{reo}} \frac{\partial \Delta \alpha}{\partial \tilde{\mathbf{N}}^s} + \frac{\partial \psi_r}{\partial \tilde{\mathbf{N}}^s} \right) d\tilde{\mathbf{N}}^s, \end{aligned} \quad (3.87)$$

where, for simplicity, the factor θ/θ_n is neglected in front of σ^{reo} , Q and H . Here, the tensor \mathbb{P}_n is given by (see Eq. (3.59))

$$\mathbb{P}_n = \frac{\partial \mathbf{N}^s}{\partial \tilde{\mathbf{N}}^s} = \frac{1}{\|\mathbf{U}_n^{i-1} \tilde{\mathbf{N}}^s \mathbf{U}_n^{i-1}\|} \left(\mathbb{I}^s - \mathbf{N}^s \otimes \mathbf{C}_n^{i-1} \mathbf{N}^s \mathbf{C}_n^{i-1} \right). \quad (3.88)$$

Further, the derivative of \mathbf{C}_r^i with respect to $\Delta\xi$ is (see Eq. (3.40))

$$\frac{\partial \mathbf{C}_r^i}{\partial \Delta\xi} = \frac{H_-(\Delta\xi)}{\xi_n} \ln(\mathbf{C}_n^i) \mathbf{C}_r^i, \quad (3.89)$$

where $H_-(\Delta\xi)$ is the double negative Heaviside function: $H_-(\Delta\xi) = -H(-\Delta\xi)$. The remaining occurring derivatives are given by

$$\frac{\partial(\psi_e + \psi_h + \psi_p)}{\partial \hat{\mathbf{C}}^i} \stackrel{(3.50)}{=} \mathbf{P}_i^\top : \frac{\partial(\psi_e + \psi_h + \psi_p)}{\partial \mathbf{C}^i} \quad (3.90)$$

$$\frac{\partial \psi_e}{\partial \mathbf{C}^i} = -\mathbf{C}^{i-1} \mathbf{F}^\top \frac{\partial \psi_e}{\partial \mathbf{b}^e} \mathbf{F} \mathbf{C}^{i-1} \quad (3.91)$$

$$\frac{\partial \psi_e}{\partial \mathbf{b}^e} \stackrel{(3.49)}{=} \left(\frac{\lambda(\xi)}{4} (\det(\mathbf{b}^e) - 1) - \frac{\mu(\xi)}{2} \right) \mathbf{b}^{e-1} + \frac{\mu(\xi)}{2} \mathbf{I} \quad (3.92)$$

$$\frac{\partial \psi_e}{\partial \xi} = \frac{1}{4} \frac{\partial \lambda}{\partial \xi} (\det(\mathbf{b}^e) - 1 - 2 \ln J^e) + \frac{1}{2} \frac{\partial \mu}{\partial \xi} (\text{tr}(\mathbf{b}^e) - 3 - 2 \ln J^e) \quad (3.93)$$

$$\frac{\partial \lambda}{\partial \xi} = - \left(\frac{1-\xi}{\lambda_A} + \frac{\xi}{\lambda_M} \right)^{-2} \left(\frac{-1}{\lambda_A} + \frac{1}{\lambda_M} \right) \quad (3.94)$$

$$\frac{\partial \mu}{\partial \xi} = - \left(\frac{1-\xi}{\mu_A} + \frac{\xi}{\mu_M} \right)^{-2} \left(\frac{-1}{\mu_A} + \frac{1}{\mu_M} \right) \quad (3.95)$$

$$\frac{\partial \psi_c}{\partial \xi} \stackrel{(3.20)}{=} (\theta - \theta_0) \Delta s^{\text{AM}} \quad (3.96)$$

$$\frac{\partial \psi_h}{\partial \mathbf{C}^i} \stackrel{(3.21)}{=} \frac{\partial \psi_h}{\partial \langle \mathbf{E}^t \rangle} \frac{\partial \langle \mathbf{E}^t \rangle}{\partial \langle \mathbf{E}^i \rangle} \frac{\partial \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^i} : \frac{\partial \mathbf{E}^i}{\partial \mathbf{C}^i} \quad (3.97)$$

$$= \frac{1}{2} \frac{1}{\xi} \frac{\partial \psi_h}{\partial \langle \mathbf{E}^t \rangle} \frac{\partial \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^i} \quad (3.98)$$

$$\frac{\partial \psi_h}{\partial \langle \mathbf{E}^t \rangle} \stackrel{(3.68)}{=} \begin{cases} k E^{\text{int}} \xi \left(\frac{2 \langle \mathbf{E}^t \rangle + 2 \langle \mathbf{E}^t \rangle^5}{1 - \langle \mathbf{E}^t \rangle^4} \right) & \langle \mathbf{E}^t \rangle \leq c \\ k E^{\text{int}} \xi a_1 & \langle \mathbf{E}^t \rangle > c \end{cases} \quad (3.99)$$

$$\frac{\partial \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^i} = \frac{1}{k} \sqrt{\frac{2}{3}} \frac{\mathbf{E}^i}{\|\mathbf{E}^i\|} \quad (3.100)$$

$$\frac{\partial \psi_h}{\partial \xi} = \frac{\partial \psi_h}{\partial \langle \mathbf{E}^t \rangle} \frac{\partial \langle \mathbf{E}^t \rangle}{\partial \xi} + \begin{cases} k E^{\text{int}} \frac{\langle \mathbf{E}^t \rangle^2}{1 - \langle \mathbf{E}^t \rangle^4} & \langle \mathbf{E}^t \rangle \leq c \\ k E^{\text{int}} (a_1 \langle \mathbf{E}^t \rangle + a_0) & \langle \mathbf{E}^t \rangle > c \end{cases} \quad (3.101)$$

$$\frac{\partial \langle \mathbf{E}^t \rangle}{\partial \xi} = \frac{-\langle \mathbf{E}^i \rangle}{\xi^2} \quad (3.102)$$

$$\frac{\partial \psi_r}{\partial \hat{\mathbf{C}}^i} \stackrel{(3.55)}{=} A(\hat{\mathbf{M}}^i - 1) \hat{\mathbf{M}}^i \hat{\mathbf{C}}^{i-1} \quad (3.103)$$

$$\frac{\partial \psi_r}{\partial \tilde{\mathbf{N}}^s} \stackrel{(3.59), (3.55)}{=} A \left(\left\| \tilde{\mathbf{N}}^s \right\|_{U_n^{i-1}} - 1 \right) \left(\frac{\mathbf{C}_n^{i-1} \tilde{\mathbf{N}}^s \mathbf{C}_n^{i-1}}{\left\| \tilde{\mathbf{N}}^s \right\|_{U_n^{i-1}}} \right) \quad (3.104)$$

$$\frac{\partial \psi_p}{\partial \mathbf{C}^i} \stackrel{(3.69)}{=} \frac{\partial \psi_p}{\partial \langle \mathbf{E}^t \rangle} : \frac{\partial \langle \mathbf{E}^t \rangle}{\partial \langle \mathbf{E}^i \rangle} : \frac{\partial \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^i} : \frac{\partial \mathbf{E}^i}{\partial \mathbf{C}^i} \quad (3.105)$$

$$= \frac{1}{2} \frac{1}{\xi} \frac{\partial \psi_p}{\partial \langle \mathbf{E}^t \rangle} : \frac{\partial \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^i} \quad (3.106)$$

$$\frac{\partial \psi_p}{\partial \langle \mathbf{E}^t \rangle} = \begin{cases} H_p l_F \frac{\xi}{\sqrt{c^2+1}} & \xi \geq \xi_0 - c \langle \mathbf{E}^i \rangle \\ H_p \xi^2 \langle \mathbf{E}^t \rangle & \text{else} \end{cases} \quad (3.107)$$

$$\frac{\partial \psi_p}{\partial \xi} = \begin{cases} H_p l_F \frac{\langle \mathbf{E}^t \rangle - c}{\sqrt{c^2+1}} & \xi \geq \xi_0 - c \langle \mathbf{E}^i \rangle \\ H_p (\xi - \xi_0 + \xi \langle \mathbf{E}^t \rangle^2) & \text{else} \end{cases} \quad (3.108)$$

$$\frac{\partial \Delta \alpha}{\partial \Delta \xi} = \sqrt{\frac{2}{3}} \frac{1}{2} \tilde{\mathbf{M}}_{C^i} : (\mathbb{P}_i - \mathbb{I}^s) : \frac{\partial \mathbf{C}_r^i}{\partial \Delta \xi} \quad (3.109)$$

$$\frac{\partial \Delta \alpha}{\partial \Delta \gamma} = \sqrt{\frac{2}{3}} \tilde{\mathbf{M}}_{C^i} : \mathbb{P}_i : \mathbf{N}^s \quad (3.110)$$

$$\frac{\partial \Delta \alpha}{\partial \tilde{\mathbf{N}}^s} = \sqrt{\frac{2}{3}} \Delta \gamma \mathbb{P}_{N^s}^\top : \mathbb{P}_i^\top : \tilde{\mathbf{M}}_{C^i} \quad (3.111)$$

$$\tilde{\mathbf{M}}_{C^i} = \frac{\partial \|\Delta_r \mathbf{C}^i\|_{U_n^{i-1}}}{\partial \mathbf{C}^i} = \frac{\mathbf{C}_n^{i-1} \Delta_r \mathbf{C}^i \mathbf{C}_n^{i-1}}{\|\Delta_r \mathbf{C}^i\|_{U_n^{i-1}}} \quad (3.112)$$

$$(3.113)$$

Here, a_1 is the slope of the linear approximation in Fig. 3.2 at $\langle \mathbf{E}^t \rangle = ck$ and a_0 is the value of the function φ at that same point. Further, the terms related to ψ_p are set to zero if $\xi > \langle \mathbf{E}^i \rangle / c + \xi_0$.

3.9 Appendix D: Second derivatives of π_Δ

The second derivatives of the potential π_Δ are necessary to compute the algorithmic tangent. They are briefly given in this appendix. First, we define the symmetrizing box product as

$$\mathbf{A} \overset{s}{\square} \mathbf{B} : \mathbf{C} = \mathbf{A} \text{sym}(\mathbf{C}) \mathbf{B}, \quad (3.114)$$

where \mathbf{A} and \mathbf{B} are arbitrary second order tensors. Now, we compute the differential of the \mathbb{P}_{N^s} operator:

$$d(\mathbb{P}_{N^s}^\top : \mathbf{B}) = \mathbb{D}_{N^s}(\mathbf{B}) : d\tilde{\mathbf{N}}^s + \mathbb{P}_{N^s}^\top : d\mathbf{B} \quad (3.115)$$

where \mathbf{B} is an arbitrary symmetric second order tensor and

$$\begin{aligned} \mathbb{D}_{N^s}(\mathbf{B}) = & - \frac{1}{\|\tilde{\mathbf{N}}^s\|_{U_n^{i-1}}^3} \left(\mathbf{f}(\mathbf{B}) \otimes \mathbf{C}_n^{i-1} \tilde{\mathbf{N}}^s \mathbf{C}_n^{i-1} + \mathbf{C}_n^{i-1} \tilde{\mathbf{N}}^s \mathbf{C}_n^{i-1} \otimes \mathbf{f}(\mathbf{B}) \right. \\ & \left. - \frac{1}{\|\tilde{\mathbf{N}}^s\|_{U_n^{i-1}}^2} (\tilde{\mathbf{N}}^s : \mathbf{B}) \mathbf{C}_n^{i-1} \tilde{\mathbf{N}}^s \mathbf{C}_n^{i-1} \otimes \mathbf{C}_n^{i-1} \tilde{\mathbf{N}}^s \mathbf{C}_n^{i-1} + (\tilde{\mathbf{N}}^s : \mathbf{B}) \mathbf{C}_n^{i-1} \overset{s}{\square} \mathbf{C}_n^{i-1} \right). \end{aligned} \quad (3.116)$$

Here, \mathbf{B} is defined as

$$\mathbf{f}(\mathbf{B}) = \mathbf{B} - \frac{1}{\|\tilde{\mathbf{N}}^s\|_{U_n^{i-1}}^2} (\tilde{\mathbf{N}}^s : \mathbf{B}) \mathbf{C}_n^{i-1} \tilde{\mathbf{N}}^s \mathbf{C}_n^{i-1}. \quad (3.117)$$

Similarly, we define the derivative of \mathbb{P}_i via

$$\mathrm{d}(\mathbb{P}_i^\top : \mathbf{B}) = \mathbb{D}_i(\mathbf{B}) : \mathrm{d}\hat{\mathbf{C}}^i + \mathbb{P}_i^\top : \mathrm{d}\mathbf{B}, \quad (3.118)$$

with

$$\begin{aligned} \mathbb{D}_i(\mathbf{B}) = & -\frac{1}{3} \hat{\mathbb{I}}^{i-\frac{1}{3}} \left(\left(\mathbf{B} - \frac{1}{3} (\hat{\mathbf{C}}^i : \mathbf{B}) \hat{\mathbf{C}}^{i-1} \right) \otimes \hat{\mathbf{C}}^{i-1} + \hat{\mathbf{C}}^{i-1} \otimes \left(\mathbf{B} - \frac{1}{3} (\hat{\mathbf{C}}^i : \mathbf{B}) \hat{\mathbf{C}}^{i-1} \right) \right. \\ & \left. - (\hat{\mathbf{C}}^i : \mathbf{B}) \left(\hat{\mathbf{C}}^{i-1} \hat{\mathbf{C}}^{i-1} - \frac{1}{3} \hat{\mathbf{C}}^{i-1} \otimes \hat{\mathbf{C}}^{i-1} \right) \right) \end{aligned} \quad (3.119)$$

Subsequently, we find for the derivatives of π_Δ :

$$\begin{aligned} \frac{\partial^2 \pi_\Delta}{\partial \Delta \xi^2} = & \frac{\partial \mathbf{C}_r^i}{\partial \Delta \xi} : \left(\mathbb{D}_i \left(\frac{\partial(\psi_e + \psi_h + \psi_p)}{\partial \mathbf{C}^i} \right) + \mathbb{P}_i^\top : \frac{\partial^2(\psi_e + \psi_h + \psi_p)}{\partial \mathbf{C}^{i2}} : \mathbb{P}_i + \frac{\partial^2 \psi_r}{\partial \hat{\mathbf{C}}^{i2}} \right) : \frac{\partial \mathbf{C}_r^i}{\partial \Delta \xi} \\ & + 2 \frac{\partial \mathbf{C}_r^i}{\partial \Delta \xi} : \mathbb{P}_i^\top : \frac{\partial^2(\psi_e + \psi_h + \psi_p)}{\partial \Delta \xi \partial \mathbf{C}^i} + \frac{\partial(\psi_e + \psi_h + \psi_p + \psi_r)}{\partial \hat{\mathbf{C}}^i} : \frac{\partial^2 \mathbf{C}_r^i}{\partial \Delta \xi^2} + \sigma^{\text{reo}} \frac{\partial^2 \Delta \alpha}{\partial \xi^2} \\ & + \frac{\partial^2(\psi_e + \psi_h + \psi_p)}{\partial \Delta \xi^2} + H(\text{sg}(\Delta \xi)) \end{aligned} \quad (3.120)$$

$$\begin{aligned} \frac{\partial^2 \pi_\Delta}{\partial \Delta \gamma^2} = & 2\mathbf{N}^s : \left(\mathbb{D}_i \left(\frac{\partial(\psi_e + \psi_h + \psi_p)}{\partial \mathbf{C}^i} \right) + \mathbb{P}_i^\top : \frac{\partial^2(\psi_e + \psi_h + \psi_p)}{\partial \mathbf{C}^{i2}} : \mathbb{P}_i + \frac{\partial^2 \psi_r}{\partial \hat{\mathbf{C}}^{i2}} \right) : 2\mathbf{N}^s \\ & + \sigma^{\text{reo}} \frac{\partial^2 \Delta \alpha}{\partial \Delta \gamma^2} \end{aligned} \quad (3.121)$$

$$\begin{aligned} \frac{\partial^2 \pi_\Delta}{\partial \tilde{\mathbf{N}}^{s2}} = & \mathbb{D}_{\mathbf{N}^s} \left(2\Delta \gamma \frac{\partial(\psi_e + \psi_h + \psi_r + \psi_p)}{\partial \hat{\mathbf{C}}^i} \right) + \sigma^{\text{reo}} \frac{\partial^2 \Delta \alpha}{\partial \tilde{\mathbf{N}}^{s2}} + \frac{\partial^2 \psi_r}{\partial \tilde{\mathbf{N}}^{s2}} \\ & + \mathbb{P}_{\mathbf{N}^s}^\top : \left(2\Delta \gamma \left(\mathbb{P}_i^\top : \frac{\partial^2(\psi_e + \psi_h + \psi_p)}{\partial \mathbf{C}^{i2}} : \mathbb{P}_i + \frac{\partial^2 \psi_r}{\partial \hat{\mathbf{C}}^{i2}} \right) 2\Delta \gamma \right) : \mathbb{P}_{\mathbf{N}^s} \end{aligned} \quad (3.122)$$

$$\begin{aligned} \frac{\partial^2 \pi_\Delta}{\partial \Delta \xi \partial \Delta \gamma} = & 2\mathbf{N}^s : \left(\mathbb{P}_i^\top : \frac{\partial^2(\psi_e + \psi_h + \psi_p)}{\partial \mathbf{C}^{i2}} : \mathbb{P}_i + \frac{\partial^2 \psi_r}{\partial \hat{\mathbf{C}}^{i2}} \right) : \frac{\partial \mathbf{C}_r^i}{\partial \Delta \xi} + \sigma^{\text{reo}} \frac{\partial^2 \Delta \alpha}{\partial \Delta \xi \partial \Delta \gamma} \\ & + 2\mathbf{N}^s : \mathbb{P}_i^\top : \frac{\partial^2(\psi_e + \psi_h + \psi_p)}{\partial \Delta \xi \partial \mathbf{C}^i} \end{aligned} \quad (3.123)$$

$$\begin{aligned} \frac{\partial^2 \pi_\Delta}{\partial \xi \partial \tilde{\mathbf{N}}^s} &= \mathbb{P}_{N^s}^\top : \left(2\Delta\gamma \left(\mathbb{P}_i^\top : \frac{\partial^2(\psi_e + \psi_h + \psi_p)}{\partial \mathbf{C}^{i2}} : \mathbb{P}_i + \frac{\partial^2 \psi_r}{\partial \hat{\mathbf{C}}^{i2}} \right) : \frac{\partial \mathbf{C}_r^i}{\partial \Delta \xi} \right) + \sigma^{\text{reo}} \frac{\partial^2 \Delta \alpha}{\partial \Delta \xi \partial \tilde{\mathbf{N}}^s} \\ &\quad + 2\Delta\gamma \mathbb{P}_{N^s}^\top : \mathbb{P}_i^\top : \frac{\partial^2(\psi_e + \psi_h + \psi_p)}{\partial \xi \partial \mathbf{C}^i} \end{aligned} \quad (3.124)$$

$$\begin{aligned} \frac{\partial^2 \pi_\Delta}{\partial \Delta \gamma \partial \tilde{\mathbf{N}}^s} &= \sigma^{\text{reo}} \frac{\partial^2 \Delta \alpha}{\partial \Delta \gamma \partial \tilde{\mathbf{N}}^s} + \mathbb{P}_{N^s}^\top : \left(2\Delta\gamma \left(\mathbb{P}_i^\top : \frac{\partial^2(\psi_e + \psi_h + \psi_p)}{\partial \mathbf{C}^{i2}} : \mathbb{P}_i + \frac{\partial^2 \psi_r}{\partial \hat{\mathbf{C}}^{i2}} \right) : 2\mathbf{N}^s \right. \\ &\quad \left. + 2 \left(\mathbb{P}_i^\top : \frac{\partial(\psi_e + \psi_h + \psi_p)}{\partial \mathbf{C}^i} + \frac{\partial \psi_r}{\partial \hat{\mathbf{C}}^i} \right) \right) \end{aligned} \quad (3.125)$$

Here, for the occurring derivatives we find:

$$\frac{\partial^2 \psi_e}{\partial \Delta \xi^2} = \frac{1}{4} \frac{\partial^2 \lambda}{\partial \xi^2} (\det(\mathbf{b}^e) - 1 - 2 \ln J^e) + \frac{1}{2} \frac{\partial^2 \mu}{\partial \xi^2} (\text{tr}(\mathbf{b}^e) - 3 - 2 \ln J^e) \quad (3.126)$$

$$\frac{\partial^2 \lambda}{\partial \xi^2} = 2 \left(\frac{1-\xi}{\lambda_A} + \frac{\xi}{\lambda_M} \right)^{-3} \left(\frac{-1}{\lambda_A} + \frac{1}{\lambda_M} \right)^2 \quad (3.127)$$

$$\frac{\partial^2 \mu}{\partial \xi^2} = 2 \left(\frac{1-\xi}{\mu_A} + \frac{\xi}{\mu_M} \right)^{-3} \left(\frac{-1}{\mu_A} + \frac{1}{\mu_M} \right)^2 \quad (3.128)$$

$$\frac{\partial^2 \psi_e}{\partial \mathbf{C}^{i2}} = \mathbb{B}_i^e : \frac{\partial^2 \psi_e}{\partial \mathbf{b}^{e2}} : \mathbb{B}_i^{e\top} - \mathbf{C}^{i-1} \overset{s}{\square} \frac{\partial \psi_e}{\partial \mathbf{C}^i} - \frac{\partial \psi_e}{\partial \mathbf{C}^i} \overset{s}{\square} \mathbf{C}^{i-1} \quad (3.129)$$

$$\mathbb{B}_i^e = (\mathbf{C}^{i-1} \mathbf{F}^\top) \overset{s}{\square} (\mathbf{F} \mathbf{C}^{i-1}) \quad (3.130)$$

$$\frac{\partial^2 \psi_e}{\partial \mathbf{b}^{e2}} = \frac{\lambda(\xi)}{4} J^{e2} \mathbf{b}^{e-1} \otimes \mathbf{b}^{e-1} - \left(\frac{\lambda(\xi)}{4} (J^{e2} - 1) - \frac{\mu}{2} \right) \mathbf{b}^{e-1} \overset{s}{\square} \mathbf{b}^{e-1} \quad (3.131)$$

$$\frac{\partial^2 \psi_e}{\partial \xi \partial \mathbf{C}^i} = -\mathbb{B}_i^e : \frac{\partial^2 \psi_e}{\partial \xi \partial \mathbf{b}^e} \quad (3.132)$$

$$\frac{\partial^2 \psi_e}{\partial \xi \partial \mathbf{b}^e} = \frac{1}{4} \frac{\partial \lambda}{\partial \xi} \left(\left(J^e - \frac{1}{J^e} \right) J^e \mathbf{b}^{e-1} \right) + \frac{1}{2} \frac{\partial \mu}{\partial \xi} (\mathbf{I} - \mathbf{b}^{e-1}) \quad (3.133)$$

$$\frac{\partial^2 \mathbf{C}_r^i}{\partial \Delta \xi^2} = \frac{H_-(\xi)}{\Delta \xi_n^2} \mathbf{C}_r^i \ln^2(\mathbf{C}_n^i) \quad (3.134)$$

$$\frac{\partial^2 \psi_r}{\partial \hat{\mathbf{C}}^{i2}} = \left(\frac{\partial^2 \psi_r}{\partial \hat{\mathbb{I}}^{i2}} \hat{\mathbb{I}}^i + \frac{\partial \psi_r}{\partial \hat{\mathbb{I}}^i} \right) \hat{\mathbb{I}}^i \hat{\mathbf{C}}^{i-1} \otimes \hat{\mathbf{C}}^{i-1} - \frac{\partial \psi_r}{\partial \hat{\mathbb{I}}^i} \hat{\mathbb{I}}^i \hat{\mathbf{C}}^{i-1} \overset{s}{\square} \hat{\mathbf{C}}^{i-1} \quad (3.135)$$

$$\frac{\partial^2 \psi_r}{\partial \tilde{\mathbf{N}}^{s2}} = A \left(\mathbf{C}_n^{i-1} \overset{s}{\square} \mathbf{C}_n^{i-1} \left(1 - \frac{1}{\|\tilde{\mathbf{N}}^s\|_{U_n^{i-1}}} \right) + \frac{1}{\|\tilde{\mathbf{N}}^s\|_{U_n^{i-1}}} \mathbf{C}_n^{i-1} \mathbf{N}^s \mathbf{C}_n^{i-1} \otimes \mathbf{C}_n^{i-1} \mathbf{N}^s \mathbf{C}_n^{i-1} \right) \quad (3.136)$$

$$\frac{\partial^2 \psi_h}{\partial \xi^2} = 2 \frac{\partial^2 \psi_h}{\partial \langle \mathbf{E}^t \rangle \partial \xi} \frac{\partial \langle \mathbf{E}^t \rangle}{\partial \xi} + \frac{\partial^2 \psi_h}{\partial \langle \mathbf{E}^t \rangle^2} \left(\frac{\partial \langle \mathbf{E}^t \rangle}{\partial \xi} \right)^2 + \frac{\partial \psi_h}{\partial \langle \mathbf{E}^t \rangle} \frac{\partial^2 \langle \mathbf{E}^t \rangle}{\partial \xi^2} \quad (3.137)$$

$$\frac{\partial^2 \psi_h}{\partial \langle \mathbf{E}^t \rangle \partial \xi} = \begin{cases} k E^{\text{int}} \frac{2\langle \mathbf{E}^t \rangle + 2\langle \mathbf{E}^t \rangle^5}{(1 - \langle \mathbf{E}^t \rangle^4)^2} & \langle \mathbf{E}^t \rangle \leq c \\ k E^{\text{int}} a_1 & \langle \mathbf{E}^t \rangle > c \end{cases} \quad (3.138)$$

$$\frac{\partial^2 \psi_h}{\partial \langle \mathbf{E}^t \rangle^2} = \begin{cases} k E^{\text{int}} \xi \frac{2 + 24\langle \mathbf{E}^t \rangle^4 + 6\langle \mathbf{E}^t \rangle^8}{(1 - \langle \mathbf{E}^t \rangle^4)^3} & \langle \mathbf{E}^t \rangle \leq c \\ 0 & \langle \mathbf{E}^t \rangle > c \end{cases} \quad (3.139)$$

$$\frac{\partial^2 \langle \mathbf{E}^t \rangle}{\partial \xi^2} = \frac{2\langle \mathbf{E}^i \rangle}{\xi^3} \quad (3.140)$$

$$\frac{\partial^2 \psi_h}{\partial \mathbf{C}^i \partial \xi} = \frac{1}{2\xi} \left(\frac{\partial^2 \psi_h}{\partial \langle \mathbf{E}^t \rangle \partial \xi} - \frac{\partial^2 \psi_h}{\partial \langle \mathbf{E}^t \rangle^2} \frac{\langle \mathbf{E}^i \rangle}{\xi^2} - \frac{1}{\xi} \frac{\partial \psi_h}{\partial \langle \mathbf{E}^t \rangle} \right) \frac{\partial \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^i} \quad (3.141)$$

$$\frac{\partial^2 \psi_h}{\partial \mathbf{C}^{i2}} = \frac{1}{4} \left(\frac{\partial^2 \psi_h}{\partial \langle \mathbf{E}^t \rangle^2} \frac{1}{\xi^2} \frac{\partial \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^i} \otimes \frac{\partial \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^i} + \frac{\partial \psi_h}{\partial \langle \mathbf{E}^t \rangle} \frac{1}{\xi} \frac{\partial^2 \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^{i2}} \right) \quad (3.142)$$

$$\frac{\partial^2 \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^{i2}} = \sqrt{\frac{2}{3}} \frac{1}{k \|\mathbf{E}^i\|} \left(\mathbb{I}^s - \frac{\mathbf{E}^i}{\|\mathbf{E}^i\|} \otimes \frac{\mathbf{E}^i}{\|\mathbf{E}^i\|} \right) \quad (3.143)$$

$$\frac{\partial^2 \psi_p}{\partial \xi^2} = \begin{cases} H_p \left(\frac{\langle \mathbf{E}^t \rangle - c}{\sqrt{1+c^2}} \right)^2 & \xi \geq \xi_0 - c \langle \mathbf{E}^i \rangle \\ H_p (1 + \langle \mathbf{E}^t \rangle^2) & \text{else} \end{cases} \quad (3.144)$$

$$\frac{\partial^2 \psi_p}{\partial \mathbf{C}^{i2}} = \frac{1}{4} \left(\frac{\partial^2 \psi_p}{\partial \langle \mathbf{E}^t \rangle^2} \frac{1}{\xi^2} \frac{\partial \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^i} \otimes \frac{\partial \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^i} + \frac{\partial \psi_p}{\partial \langle \mathbf{E}^t \rangle} \frac{1}{\xi} \frac{\partial^2 \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^{i2}} \right) \quad (3.145)$$

$$\frac{\partial^2 \psi_p}{\partial \mathbf{C}^i \partial \xi} = \frac{1}{2\xi} \left(\frac{\partial^2 \psi_p}{\partial \langle \mathbf{E}^t \rangle \partial \xi} - \frac{\partial^2 \psi_p}{\partial \langle \mathbf{E}^t \rangle^2} \frac{\langle \mathbf{E}^i \rangle}{\xi^2} - \frac{1}{\xi} \frac{\partial \psi_p}{\partial \langle \mathbf{E}^t \rangle} \right) \frac{\partial \langle \mathbf{E}^i \rangle}{\partial \mathbf{E}^i} \quad (3.146)$$

$$\frac{\partial^2 \psi_p}{\partial \langle \mathbf{E}^t \rangle^2} = \begin{cases} H_p \left(\frac{\xi}{\sqrt{1+c^2}} \right)^2 & \xi \geq \xi_0 - c \langle \mathbf{E}^i \rangle \\ H_p \xi^2 & \text{else} \end{cases} \quad (3.147)$$

$$\frac{\partial^2 \psi_p}{\partial \xi \partial \langle \mathbf{E}^t \rangle} = \begin{cases} H_p \left(\frac{\langle \mathbf{E}^t \rangle - c}{\sqrt{1+c^2}} \frac{\xi}{\sqrt{1+c^2}} + l_F \frac{1}{\sqrt{1+c^2}} \right) & \xi \geq \xi_0 - c \langle \mathbf{E}^i \rangle \\ 2H_p \xi \langle \mathbf{E}^t \rangle & \text{else} \end{cases} \quad (3.148)$$

$$\begin{aligned} \frac{\partial^2 \Delta \alpha}{\partial \Delta \xi^2} &= \sqrt{\frac{2}{3}} \left(\frac{\partial \mathbf{C}_r^i}{\partial \xi} : \left((\mathbb{P}_i^\top - \mathbb{I}^s) : \frac{1}{2} \frac{\partial \tilde{\mathbf{M}}_{C^i}}{\partial \mathbf{C}^i} : (\mathbb{P}_i - \mathbb{I}^s) + \mathbb{D}_i \left(\frac{1}{2} \tilde{\mathbf{M}}_{C^i} \right) \right) : \frac{\partial \mathbf{C}_r^i}{\partial \xi} \right. \\ &\quad \left. + \frac{1}{2} \tilde{\mathbf{M}}_{C^i} : (\mathbb{P}_i - \mathbb{I}^s) : \frac{\partial^2 \mathbf{C}_r^i}{\partial \xi^2} \right) \end{aligned} \quad (3.149)$$

$$\frac{\partial^2 \Delta \alpha}{\partial \Delta \gamma^2} = \sqrt{\frac{2}{3}} 2\mathbf{N}^s : \left(\mathbb{D}_i \left(\frac{1}{2} \tilde{\mathbf{M}}_{C^i} \right) + \frac{1}{2} \mathbb{P}_i^\top : \frac{\partial \tilde{\mathbf{M}}_{C^i}}{\partial \mathbf{C}^i} : \mathbb{P}_i \right) : 2\mathbf{N}^s \quad (3.150)$$

$$\begin{aligned} \frac{\partial^2 \Delta \alpha}{\partial \tilde{\mathbf{N}}^s{}^2} &= \sqrt{\frac{2}{3}} \left(2\Delta \gamma \mathbb{D}_{N^s} \left(\mathbb{P}_i^\top : \frac{1}{2} \tilde{\mathbf{M}}_{C^i} \right) \right. \\ &\quad \left. + (2\Delta \gamma)^2 \mathbb{P}_{N^s}^\top : \left(\mathbb{D}_i \left(\frac{1}{2} \tilde{\mathbf{M}}_{C^i} \right) + \mathbb{P}_i^\top : \frac{1}{2} \frac{\partial \tilde{\mathbf{M}}_{C^i}}{\partial \mathbf{C}^i} : \mathbb{P}_i \right) : \mathbb{P}_{N^s} \right) \end{aligned} \quad (3.151)$$

$$\frac{\partial^2 \Delta \alpha}{\partial \Delta \xi \partial \Delta \gamma} = \sqrt{\frac{2}{3}} 2\mathbf{N}^s : \left(\mathbb{D}_i \left(\frac{1}{2} \tilde{\mathbf{M}}_{C^i} \right) : \frac{\partial \mathbf{C}_r^i}{\partial \xi} + \mathbb{P}_i^\top : \frac{1}{2} \frac{\partial \tilde{\mathbf{M}}_{C^i}}{\partial \mathbf{C}^i} : (\mathbb{P}_i - \mathbb{I}^s) : \frac{\partial \mathbf{C}_r^i}{\partial \xi} \right) \quad (3.152)$$

$$\frac{\partial^2 \Delta \alpha}{\partial \Delta \xi \partial \tilde{\mathbf{N}}^s} = \sqrt{\frac{2}{3}} 2\Delta \gamma \mathbb{P}_{N^s}^\top : \left(\mathbb{D}_i \left(\frac{1}{2} \tilde{\mathbf{M}}_{C^i} \right) : \frac{\partial \mathbf{C}_r^i}{\partial \xi} + \mathbb{P}_i^\top : \frac{1}{2} \frac{\partial \tilde{\mathbf{M}}_{C^i}}{\partial \mathbf{C}^i} : (\mathbb{P}_i - \mathbb{I}^s) : \frac{\partial \mathbf{C}_r^i}{\partial \xi} \right) \quad (3.153)$$

$$\frac{\partial^2 \Delta \alpha}{\partial \Delta \gamma \partial \tilde{\mathbf{N}}^s} = \sqrt{\frac{2}{3}} \left(2\mathbb{P}_{N^s}^\top : \mathbb{P}_i^\top : \frac{1}{2} \tilde{\mathbf{M}}_{C^i} + 2\Delta \gamma \mathbb{P}_{N^s}^\top : \left(\mathbb{D}_i \left(\frac{1}{2} \tilde{\mathbf{M}}_{C^i} \right) + \mathbb{P}_i^\top : \frac{1}{2} \frac{\partial \tilde{\mathbf{M}}_{C^i}}{\partial \mathbf{C}^i} : \mathbb{P}_i \right) : 2\mathbf{N}^s \right) \quad (3.154)$$

$$\frac{\partial \tilde{\mathbf{M}}_{C^i}}{\partial \mathbf{C}^i} = \frac{1}{\|\tilde{\mathbf{N}}^s\|_{U_n^{i-1}}} \left(\mathbf{C}_n^{i-1} \square \mathbf{C}_n^{i-1} - \tilde{\mathbf{M}}_{C^i} \otimes \tilde{\mathbf{M}}_{C^i} \right) \quad (3.155)$$

3.10 Appendix E: Consistent tangent

When using the finite element method to solve boundary value problems, we make use of the consistent tangent operator. To derive it with respect to \mathbf{d}_d from the potential, we start with the virtual work of the internal forces

$$\int_{V_0} \boldsymbol{\tau} : \mathbf{d}_\delta \, dV. \quad (3.156)$$

Here, \mathbf{d}_δ is defined as

$$\mathbf{d}_\delta = \text{sym} \left(\underbrace{\text{Grad}(\delta \mathbf{u}) \mathbf{F}^{-1}}_{\mathbf{l}_\delta} \right). \quad (3.157)$$

Now, we need to calculate the differential of virtual work of the internal forces:

$$\int_{V_0} d(\boldsymbol{\tau} : \mathbf{d}_\delta) \, dV, \quad (3.158)$$

where through using the symmetry of $\boldsymbol{\tau}$ we can simplify to

$$d(\boldsymbol{\tau} : \mathbf{d}_\delta) = d(\boldsymbol{\tau} : \mathbf{l}_\delta) = d(\mathbf{F}\mathbf{S}) : \delta \mathbf{F}. \quad (3.159)$$

Now, the differential reads

$$\begin{aligned} d(\mathbf{F}\mathbf{S}) : \delta \mathbf{F} &= (d\mathbf{F}\mathbf{F}^{-1}\mathbf{F}\mathbf{S}\mathbf{F}^\top (\delta \mathbf{F}\mathbf{F}^{-1})^\top) : \mathbf{I} + (\mathbf{F} d\mathbf{S}\mathbf{F}^\top) : \mathbf{l}_\delta \\ &= (\mathbf{l}_d \boldsymbol{\tau} \mathbf{l}_\delta^\top) : \mathbf{I} + (\mathbf{c}^a : \mathbf{d}_d) : \mathbf{d}_\delta, \end{aligned} \quad (3.160)$$

where the first term, i.e., $\mathbf{l}_d \boldsymbol{\tau} \mathbf{l}_\delta^\top$ is related to the so-called geometric tangent and the second term arises from the incremental material stiffness. Further, \mathbf{l}_d and \mathbf{d}_d are defined in analogy to \mathbf{l}_δ and \mathbf{d}_δ . Here, we neglected, in a first step, the influence of perturbations $d\theta$ of the temperature (for the full linearization, see below). Now, we can split \mathbf{c}^a into the elastic and

inelastic parts:

$$\begin{aligned}\mathbb{c}^a : \mathbf{d}_d &= \left(\mathbf{F} \stackrel{s}{\square} \mathbf{F}^\top : \frac{\partial \mathbf{S}}{\partial \mathbf{E}} : \mathbf{F}^\top \stackrel{s}{\square} \mathbf{F} \right) : \mathbf{d}_d + \frac{\partial(\mathbf{F} \mathbf{S} \mathbf{F}^\top)}{\partial \mathcal{A}} : \frac{\partial \mathcal{A}}{\partial \mathbf{E}} : d\mathbf{E} \\ &= \mathbb{c}^e : \mathbf{d}_d + \frac{\partial \boldsymbol{\tau}}{\partial \mathcal{A}} : \frac{\partial \mathcal{A}}{\partial \mathbf{E}} : \mathbf{F}^\top \mathbf{d}_d \mathbf{F},\end{aligned}\quad (3.161)$$

where (with a small abuse of notation) \mathcal{A} denotes the vector of active variables and may contain ξ , $\Delta\gamma$ and $\tilde{\mathbf{N}}^s$. Additionally, the elastic tangent \mathbb{c}^e is given by

$$\mathbb{c}^e = (\lambda(\xi) \det(\mathbf{b}^e)) \mathbf{I} \otimes \mathbf{I} + (2\mu(\xi) - \lambda(\xi)(\det(\mathbf{b}^e) - 1)) \mathbb{I}^s. \quad (3.162)$$

Next, we need to obtain $\frac{\partial \mathcal{A}}{\partial \mathbf{E}}$. Therefore, we take a look at the local Newton scheme, where

$$\frac{\partial \pi_\Delta}{\partial \mathcal{A}} = 0 \quad (3.163)$$

holds in the converged state. Now, calculating the differential of Eq. (3.163), we obtain

$$\begin{aligned}0 &= d\left(\frac{\partial \pi_\Delta}{\partial \mathcal{A}}\right) \\ &= \frac{\partial^2 \pi_\Delta}{\partial \mathcal{A}^2} d\mathcal{A} + \frac{\partial^2 \pi_\Delta}{\partial \mathcal{A} \partial \mathbf{E}} : d\mathbf{E} = 0 \\ &= \frac{\partial^2 \pi_\Delta}{\partial \mathcal{A}^2} d\mathcal{A} + \left(\frac{\partial \mathbf{S}}{\partial \mathcal{A}}\right)^\top : d\mathbf{E} = 0.\end{aligned}\quad (3.164)$$

Rearranging the result from Eq. (3.164) we obtain

$$\begin{aligned}d\mathcal{A} &= - \left(\frac{\partial^2 \pi_\Delta}{\partial \mathcal{A}^2}\right)^{-1} \frac{\partial^2 \pi_\Delta}{\partial \mathcal{A} \partial \mathbf{E}} : d\mathbf{E} \\ &= - \frac{\partial \mathcal{A}}{\partial \mathbf{E}} : d\mathbf{E} \\ &= - \left(\frac{\partial^2 \pi_\Delta}{\partial \mathcal{A}^2}\right)^{-1} \left(\frac{\partial \mathbf{S}}{\partial \mathcal{A}}\right)^\top : \mathbf{F}^\top \stackrel{s}{\square} \mathbf{F} : \mathbf{d}_d \\ &= - \left(\frac{\partial^2 \pi_\Delta}{\partial \mathcal{A}^2}\right)^{-1} \left(\frac{\partial \boldsymbol{\tau}}{\partial \mathcal{A}}\right)^\top : \mathbf{d}_d,\end{aligned}\quad (3.165)$$

which we can insert into Eq. (3.161) to finally obtain the algorithmic tangent:

$$\mathbb{c}^a = \mathbb{c}^e - \frac{\partial \boldsymbol{\tau}}{\partial \mathcal{A}} \left(\frac{\partial^2 \pi_\Delta}{\partial \mathcal{A}^2}\right)^{-1} \left(\frac{\partial \boldsymbol{\tau}}{\partial \mathcal{A}}\right)^\top. \quad (3.166)$$

In here, the term $\partial\tau/\partial\mathcal{A}$ is computed as

$$\frac{\partial\tau}{\partial\mathcal{A}} = \begin{pmatrix} \frac{\partial\tau}{\partial\xi} \\ \frac{\partial\tau}{\partial\Delta\gamma} \\ \frac{\partial\tau}{\partial\tilde{\mathbf{N}}^s} \end{pmatrix}, \quad (3.167)$$

with the additional derivatives

$$\frac{\partial\tau}{\partial\xi} = \frac{\partial\tau}{\partial\hat{\mathbf{C}}^i} : \frac{\partial\mathbf{C}_r^i}{\partial\xi} + 2\mathbf{b}^e \frac{\partial^2\psi_e}{\partial\mathbf{b}^e\partial\xi} \quad (3.168)$$

$$\frac{\partial\tau}{\partial\Delta\gamma} = 2 \frac{\partial\tau}{\partial\hat{\mathbf{C}}^i} : \mathbf{N}^s, \quad (3.169)$$

$$\frac{\partial\tau}{\partial\tilde{\mathbf{N}}^s} = 2\Delta\gamma \frac{\partial\tau}{\partial\hat{\mathbf{C}}^i} : \mathbb{P}_{\mathbf{N}^s} \quad (3.170)$$

$$\begin{aligned} \frac{\partial\tau}{\partial\hat{\mathbf{C}}^i} = & -2 \left(\mathbf{I} \stackrel{s}{\square} \frac{\partial\psi_e}{\partial\mathbf{b}^e} + \frac{\partial\psi_e}{\partial\mathbf{b}^e} \stackrel{s}{\square} \mathbf{I} + \left(\mathbf{I} \stackrel{s}{\square} \mathbf{b}^e + \mathbf{b}^e \stackrel{s}{\square} \mathbf{I} \right) : \frac{\partial^2\psi_e}{\partial\mathbf{b}^e{}^2} \right) \\ & : \mathbf{F} \stackrel{s}{\square} \mathbf{F}^\top : \mathbf{C}^{i-1} \stackrel{s}{\square} \mathbf{C}^{i-1} : \mathbb{P}_i. \end{aligned} \quad (3.171)$$

Additionally, the algorithmic consistent tangent with respect to θ as well as the mixed contributions can be computed through linearizing the weak form in Eq. (3.74):

$$\begin{aligned} & \int_{V_0} d \left(\frac{\partial\pi_\Delta}{\partial\theta} \delta\theta - \frac{\kappa\Delta\theta}{\theta_n} \text{Grad}(\theta) \cdot \text{Grad}(\delta\theta) \right) dV \\ = & \int_{V_0} \delta\theta \left(\left(-\frac{\partial^2\pi_\Delta}{\partial\mathcal{A}\partial\theta} \left(\frac{\partial^2\pi_\Delta}{\partial\mathcal{A}^2} \right)^{-1} \frac{\partial^2\pi_\Delta}{\partial\mathcal{A}\partial\theta} - \frac{c}{\theta} \right) d\theta + \frac{\partial\tau}{\partial\theta} : \mathbf{d}_d \right) - \text{Grad}(\delta\theta) \left(\frac{\kappa\Delta t}{\theta_n} \text{Grad}(d\theta) \right) dV. \end{aligned} \quad (3.172)$$

Here, the derivative $\partial\tau/\partial\theta$ is given by

$$\frac{\partial\tau}{\partial\theta} = \frac{\partial\tau}{\partial\mathcal{A}} \left(\frac{\partial^2\pi_\Delta}{\partial\mathcal{A}^2} \right)^{-1} \frac{\partial^2\pi_\Delta}{\partial\mathcal{A}\partial\theta}, \quad (3.173)$$

with









$$\frac{\partial^2\pi_\Delta}{\partial\mathcal{A}\partial\theta} = \left(\Delta s^{\text{AM}} + \frac{\sigma^{\text{reo}}}{\theta_n} \frac{\partial\Delta\alpha}{\partial\xi} + \frac{Q}{\theta_n} + \frac{H\Delta\xi}{\theta_n}; \frac{\sigma^{\text{reo}}}{\theta_n} \frac{\partial\Delta\alpha}{\partial\Delta\gamma}; \frac{\sigma^{\text{reo}}}{\theta_n} \left(\frac{\partial\Delta\alpha}{\partial\tilde{\mathbf{N}}^s} \right)^\top \right)^\top. \quad (3.174)$$

Now, using standard formulations for the element ansatz functions \mathbf{N} and it's gradients \mathbf{B} , one can assemble the element stiffness matrix (see, e.g., Wriggers [146]).

CHAPTER 4

Article 3: TiNiHf/SiO₂/Si shape memory film composites for bidirectional micro actuation

This article was published as:

Curtis, S. M. , Sielenkämper, M. , Arivanandhan, G. , Dengiz, D., Li, Z. , Jetter, J. , Hanke, L., Bumke, L. , Quandt, E. , Wulfinghoff, S. and Kohl, M.  [2022], 'TiNiHf/SiO₂/Si shape memory film composites for bidirectional micro actuation', *International Journal of Smart and Nano Materials* **13**(2), 293-314.

Own contributions to the following article:

- planning (simulation part)
- implementation of algorithm (large fraction)
- numerical simulations and parameter fitting (large fraction)
- writing of the manuscript (large fraction regarding simulations and their results, small fraction of the introduction and conclusions)
- interpretation of the results (large fraction for simulations, small fraction for all other results)
- editing of the manuscript (small fraction)

Abstract

The martensitic phase transformation in Ti_{40.4}Ni₄₈Hf_{11.6} shape memory alloys is leveraged for bi-directional actuation with TiNiHf/SiO₂/Si composites. The shape memory properties of magnetron sputtered Ti_{40.4}Ni₄₈Hf_{11.6} films annealed at 635 °C – 5 min are influenced by film thickness and the underlying substrate. Decreasing TiNiHf film thickness from 21 μm to 110 nm results in the reduction of all characteristic transformation temperatures until a critical thickness is reached. Particularly, Ti_{40.4}Ni₄₈Hf_{11.6} thin films as low as 220 nm show transformations above room temperature when deposited on SiO₂ buffer layer, which is of great interest in nano-actuation. In comparison, 220 nm films on Si substrates are austenitic at room temperature, and thus not suitable for actuation. Thermal fatigue tests on TiNiHf/SiO₂/Si bimorphs demonstrate better functional fatigue characteristics than freestanding films, with an average reduction of 15 °C after 125 cycles, with temperature stabilization subsequently. Experimental bi-directional actuation results are promising in the development of bistable actuators within a PMMA/TiNiHf/Si trimorph composite, whereby the additional PMMA layer undergoes a glass transition at 105 °C. With the aid of constitutive modeling, a route is elaborated on how bistable actuation can be achieved at micro- to nanoscales by showing favorable thickness combinations of PMMA/TiNiHf/Si composite.

Keywords: bimorph microactuator, NiTiHf, bistable actuator, shape memory alloy, thin films.

4.1 Introduction

Next-generation of silicon (Si)-based nanophotonic and nanomechanical devices, such as optical waveguide switches and routers in advanced communication technology, demand for ultra-small (micro and nano) actuators allowing for large displacements compared to their footprint size [20]. In particular, bi-directional and bistable actuators that could be directly integrated onto Si chips are attractive for such applications [88]. Actuators based on sputtered thin-film NiTiX (X = Zr, Hf, Pd, Cu, Pt, Au) shape memory alloys (SMAs) have the largest power-to-weight ratio among lightweight technologies. SMA actuators are known to exhibit limited speed. However, heat transfer scales favorably with downscaling allowing for an increase of actuation frequency. At the nanoscale, frequencies may reach the kilohertz regime depending on the heat transfer rate and fraction of transforming SMA material [62]. This investigation focuses on a route for design and fabrication of quasi-stationary bi-directional and bistable switching devices with large stroke and force that could be scaled down to nanometer dimensions, which may not be achieved by other technologies. Sputtered TiNiHf SMAs are attractive for this investigation because they are low in cost, have large transformation temperatures, and offer high work density among various SMAs [47, 49, 56]. SMA/Si bimorph nano-actuators have been fabricated using standard E-beam lithography,

wet etching, and micro-machining with conventional plasma etching techniques (e.g. RIE, IBE) [12, 59, 100, 140]. These bimorph actuators take advantage of the combined bimorph and shape memory effect, which arises when SMAs are sputtered on a substrate with a different coefficient of thermal expansion (CTE) [59, 100, 140]. In the case of this work, this is between the substrate, silicon ($\alpha_{\text{Si}} = 2.63 \times 10^{-6} \frac{1}{\text{K}}$ [138]), and the two different crystallographic phases of TiNiHf (cubic and monoclinic), which are dependent on atomic composition. For example, for a TiNiHf₂₀ alloy, austenite's CTE is $\alpha_{\text{A}} = 49 \times 10^{-6} \frac{1}{\text{K}}$ and the isotropic polycrystalline CTE value for martensite phase is calculated to be $\alpha_{\text{M}} = 9.5 \times 10^{-6} \frac{1}{\text{K}}$ [113, 114]. Another advantage of TiNiHf/Si bimorph actuators is the prospect of developing bistable actuation by adding a third layer of polymer with a glass transition temperature (T_g) that falls in between the martensite and austenite phase transformation temperatures of the SMA [123, 126, 142]. Alternatively, bistable actuation can be achieved by adding an SMA with a narrow hysteresis (e.g. TiNiCu) as the third layer [132, 140, 141].

Controlling actuation precisely with Ni-lean TiNiHf is challenging compared to other SMAs like Ni-rich TiNiHf, NiTi, and TiNiCu, as they suffer from lower transformation strains [49], have unstable transformation temperatures with thermal cycling [48], and have a low strength against dislocation plasticity [50]. Furthermore, the large thermal hysteresis reflects poor crystallographic compatibility between the martensite and austenite phases [23, 129], which is known to lead to unwanted effects such as structural/functional fatigue, and a change in volume during the phase transformation, known as the volume effect [114]. The larger thermal hysteresis is detrimental to the lifetime and energy efficiency of TiNiHf devices [8]. These problems may be overcome by training TiNiHf SMAs [75, 82], aging Ni-rich NiTiHf compositions [50, 70, 71, 129], or by designing TiNiHf-based layered bi-/multimorph composites [69]. The functional and structural fatigue properties in TiNiHf can be controlled by composition and microstructure (grain size, precipitate size, precipitate homogeneity), which is dependent on annealing conditions [48, 56]. Even though Ni-lean TiNiHf alloys are claimed to have poor functional stability, previously Bechtold et al. [8] showed that 20 μm -25 μm thick Ni_{49.2}Ti_{31.4}Hf_{19.3} films can undergo a phase transformation for an average of 1.5×10^6 actuation cycles when tested to a maximum critical stress of 300 MPa. However, when the maximum critical stress was increased to 450 MPa, the number of cycles until failure was reduced to an average of 65 K. Improved fatigue in these sputtered Ni-lean TiNiHf alloy samples could be due to small grain size ($\sim 100 \text{ nm}$) and finely dispersed small precipitates at the grain boundaries [8]. Another advantage of Ni-lean TiNiHf alloys is that they can be heat treated at moderate temperatures for a short amount of time to generate fine, homogeneous distributions of coherent zone structures and Ti₂Ni-type precipitates, which can strengthen the lattice against slip [33, 56].

Film thickness effects on the transformation temperatures of TiNiHf/Si bimorphs must be taken into account when designing TiNiHf actuators. This is specifically of interest for bistable nano-actuators because the SMA transformation temperatures need to be coupled to the transition temperatures of the third layer (e.g. polymer, additional SMA). Decreasing

SMA film thickness from micro- to nanoscale is known to decrease the martensitic transformation temperatures for several SMA systems [44, 55, 134]. Sputtered TiNiHf thin films are also reported to have lower transformation temperatures compared to bulk material of similar compositions, attributed possibly to be due to finer grain structures [33]. Several studies explore the properties of bulk TiNiHf [72, 73, 74, 150], freestanding TiNiHf films [8], TiNiHf films on Mo substrates [140, 144], and TiNiHf films on silicon substrates [33, 82, 107]; however, there are only a few studies that show the properties of sputtered TiNiHf films on silicon oxide (SiO₂) buffer layers on Si substrates [56].

In this work, thickness effects and functional fatigue characteristics are investigated in freestanding TiNiHf films (21 μm - 5 μm) and films 5 μm down to 110 nm on Si substrates with and without SiO₂ buffer layers. The influence of film thickness on the transition temperatures of TiNiHf/Si and TiNiHf/SiO₂/Si bimorphs is compared, and limits for downscaling to a scale usable for nanodevices are discussed. The possibilities of downscaling and tailoring the properties make TiNiHf a promising material in the development of bistable nano-actuators with an additional PMMA layer to the bimorph. A thermomechanically coupled finite element model is also implemented to guide the design of favorable thickness combinations of the PMMA/TiNiHf/Si composite layers to achieve such bistable nano-actuators.

4.2 Methods and materials

4.2.1 Preparation of freestanding TiNiHf films

Structured freestanding amorphous Ti_{40.4}Ni₄₈Hf_{11.6} films were fabricated into dogbone geometries through a combination of UV-lithography, DC magnetron sputtering, and a wet chemical etching process, as described in detail by Lima de Miranda et al. [66]. A Von Ardenne CS730S (Von Ardenne, Germany, base pressure $< 3 \times 10^7$ mbar) cluster magnetron sputtering device was used to sputter amorphous Ti_{40.4}Ni₄₈Hf_{11.6} films onto a pre-structured substrate using a multilayer sputter deposition approach. A 4in Ti₄₂Ni₄₃Hf₁₅ target (Ing-puls, Germany) was sputtered for 35 s (deposition layer thickness of ~ 57 nm, pressure of 2.3×10^{-3} mbar, argon flow of 25 sccm, and power of 150 W). Next, an 8in pure Ti target was sputtered for 10 s (deposition layer thickness of ~ 10 nm, pressure of 2.3×10^{-3} mbar, argon flow of 25 sccm, power of 100 W). By repeating this multilayer deposition sequence, amorphous freestanding Ti/TiNiHf films were sputtered with varying thickness ($5 \pm 0.5 \mu\text{m}$, $10 \pm 1 \mu\text{m}$, and $21 \pm 2 \mu\text{m}$). Due to the sputter yield, films are typically 10 – 15% thinner at the edge of a 4in wafer compared to the center. Rapid thermal annealing (Createc Fischer RTA-6 SY09, Germany) was used to crystallize the amorphous films and homogenize the microstructure. Different RTA temperatures between 635 °C and 750 °C and times between 5 min and 60 min were tested before selecting on the final heat treatment for all samples of 635 °C - 5 min.

The nominal film composition was determined using a Helios NanoLab 600 scanning elec-

tron microscopy (SEM) (FEI, Germany) equipped with an energy-dispersive X-ray spectroscopy (EDX) silicon drift detector (Oxford Instruments, UK). Qualitative analysis used Ti_{49.60}Ni_{50.40} binary standard; however, the error on all reported compositional data is around ± 0.5 at.%. EDX measurements were taken for sputtered amorphous TiNiHf film on a 100 mm silicon substrate. The average film composition was determined to be Ti_{40.4}Ni₄₈Hf_{11.6} for all samples. Slight compositional variations are expected to be sources of error in the following experiments.

4.2.2 Preparation of TiNiHf/Si and TiNiHf/SiO₂/Si bimorphs

TiNiHf/Si and TiNiHf/SiO₂/Si bimorph structures were prepared using the same multi-layer sputter approach described above onto chips with lateral dimensions of 20 mm \times 20 mm. Amorphous Ti_{40.4}Ni₄₈Hf_{11.6} films of different thicknesses ($5 \pm 0.5 \mu\text{m}$, $2 \pm 0.2 \mu\text{m}$, $0.88 \pm 0.08 \mu\text{m}$, $0.44 \pm 0.04 \mu\text{m}$, $0.22 \pm 0.02 \mu\text{m}$, and $0.11 \pm 0.01 \mu\text{m}$) were deposited and annealed on $525 \pm 20 \mu\text{m}$ (100) silicon substrates (Siegert Wafer, Germany) and $1.5 \mu\text{m}$ SiO₂/ $525 \pm 25 \mu\text{m}$ (100) silicon substrates (MicroChemicals, Germany). TiNiHf films on SiO₂/Si substrates are $\sim 10\%$ thicker than the films on Si substrates because the SiO₂/Si substrates were sputtered in the direct center of the device, while the pure Si substrates were sputtered adjacent to the SiO₂/Si sample. TiNiHf films with two different thicknesses ($2 \pm 0.1 \mu\text{m}$ and $0.88 \pm 0.04 \mu\text{m}$) were also sputtered onto 100 nm SiO₂/ $300 \pm 3 \mu\text{m}$ (100) Si substrates (Si-mat silicon, Germany), pre-structured into cantilevers (3.5 mm \times 20 mm) for bi-directional actuation measurements. All film composites were annealed via RTA at 635 °C - 5 min.

4.2.3 Tensile testing

The mechanical properties of TiNiHf films were determined using a high-temperature tensile test on freestanding TiNiHf films. The dogbone geometry has been chosen with a width and length of 500 μm and 4 mm, respectively. The tensile tests were conducted using a displacement-controlled micro-tensile setup equipped with a load cell (KM26z-0.2 kN, ME measuring systems) for force measurements and a digital camera (Pike 505, Allied Vision Technology) to record images for measurement of corresponding strain values using non-contact digital image correlation method. The cross-correlation of images with speckle patterns is evaluated using MATLAB code.

4.2.4 Differential scanning calorimetry

Thermal analysis and thermal cycling on freestanding SMA films (thickness of 5 μm , 10 μm , 21 μm) were conducted on a DSC 204 F1 Phoenix (Netzsch, Germany) with a heating and cooling rate of $10 \frac{^\circ\text{C}}{\text{min}}$. The transformation temperatures and latent heat of the SMA are determined using the software Proteus 7.1.0 by the tangent method. Thermal hysteresis is

calculated according to $\Delta T = (A_s + A_f - M_s - M_f)/2$, where A_s/A_f correspond to austenite start and finish temperatures and M_s/M_f correspond to martensite start and finish temperatures.

4.2.5 Electrical resistance measurements

Four-point resistance measurements were carried out inside a cryostat to determine the temperature-dependent electrical resistance of TiNiHf films constrained by Si and SiO₂/Si substrates. Quasi-stationary conditions guaranteed that influence of temperature change is negligible during measurement. For films of micrometer thickness, probes contacting the film surface were used to connect to the setup, and for films of nanometer thickness, films were connected by wire bonding. Phase transformations were determined by change in resistance with temperature, and transformation temperatures were calculated using the tangent method.

4.2.6 X-Ray Diffraction (XRD)

The martensitic crystal structures and transformation temperatures of TiNiHf films on Si and SiO₂/Si substrates were confirmed using temperature-dependent XRD with a SmartLab 9 kW diffractometer (Rigaku, Japan), CuK_α radiation ($\lambda = 1.5406 \text{ \AA}$), and 2-D Hypix3000 detector operated in 1D mode. $\theta - 2\theta$ scans in the range of $10 - 100^\circ \text{C}$ were conducted with a step size of 0.02° and a scan speed of $20 \frac{^\circ}{\text{min}}$. A heating stage (AntonPaar DHS 1100, Germany) and cooling stage (AntonPaar DCS 350, Germany) was used in the temperature range between -100°C and 130°C .

4.2.7 Cantilever deflection measurements

Functional fatigue behavior in bi-directional actuators was characterized by measuring cantilever deflection as a function of temperature. The cantilever deflection measurement setup contains a laser, a mirror for directing the laser onto the surface, a position-sensitive detector (PSD) to detect deflection, a Peltier element for heating and cooling, and a thermocouple. The actuator's stroke, D , of a cantilever can be calculated using Equation (4.1):

$$D = \frac{\Delta x l}{4a} \quad (4.1)$$

where Δx corresponds to laser dislocation on the PSD surface, l is length of the cantilever beam, a is the distance between cantilever tip and PSD ($a = 117 \text{ mm}$), and deflection is measured as volts on the PSD (V) which is converted to displacement (mm) by using a factor $dx/dU = 0.794 \frac{\text{mm}}{\text{V}}$. Each cantilever has a size of $3.5 \text{ mm} \times 20 \text{ mm}$, with a freestanding cantilever length of 14 mm . The experiments were performed under vacuum ($10 \times 10^{-4} \text{ mbar}$)

for each sample with a $10 \frac{^\circ\text{C}}{\text{min}}$ heating/cooling rate within a temperature range of 30°C - 170°C .

4.2.8 Constitutive modeling

The model used to describe the SMA is in line with the model by Sedláč et al. [109], extended to the finite strain case with inhomogeneous temperature fields. The model details regarding numerical procedures are given in [119]. In contrast to other phenomenological finite strain SMA models, it makes use of a projection method to satisfy the inelastic volume preservation constraint [43, 118]. Further, we assume a multiplicative split of the deformation gradient in the form of $\mathbf{F} = \mathbf{F}^e \mathbf{F}^i \mathbf{F}^\theta$, where \mathbf{F}^e is the elastic, \mathbf{F}^i the inelastic, and \mathbf{F}^θ the thermal part of the deformation gradient \mathbf{F} . Moreover, the volume fraction of martensite $\xi \in [0, 1]$ and the inelastic right Cauchy-Green tensor $\mathbf{C}^i = \mathbf{F}^{iT} \mathbf{F}^i$ are used as the internal variables of the model at hand. Additionally, the model is implemented into the framework of generalized standard materials [36], which ensures thermodynamic consistency when carefully choosing the potentials. This framework is then extended to non-constant temperatures θ using a variational formulation [149].

Thus, the total potential is given by $\pi = \dot{\psi} + \phi$, where $\dot{\psi}$ is the rate of the Helmholtz free energy density and ϕ is the dissipation potential. This potential is integrated into the system and minimized to solve the energy balance as well as the linear momentum balance in the bodies considered. We assume ψ to be the sum of an elastic, a chemical, and a hardening-type energy. The elastic energy is given by

$$\psi_e = \frac{\lambda(\xi)}{4} (J^{e2} - 1 - 2 \ln J^e) + \frac{\mu(\xi)}{2} (\text{tr}(\mathbf{b}^e) - 3 - 2 \ln J^e) \quad (4.2)$$

where $\lambda(\xi)$ and $\mu(\xi)$ are the Lamé parameters determined by a Reuss-like rule of mixture and J^e is the determinant of \mathbf{F}^e . Further, the Kirchhoff stress $\boldsymbol{\tau}$ is given by $\boldsymbol{\tau} = 2\mathbf{b}^e \partial \psi_e / \partial \mathbf{b}^e$, where $\mathbf{b}^e = \mathbf{F}^e \mathbf{F}^{eT}$ is the elastic left Cauchy-Green tensor. For the chemical energy, we assume a standard relationship [65, 94]:

$$\psi_c = u_0^A - \theta s_0^A + \xi (\theta - \theta_0) \Delta s^{\text{AM}} + c \left(\theta - \theta_0 - \theta \ln \frac{\theta}{\theta_0} \right) \quad (4.3)$$

with the volume-specific internal energy of the austenite phase u_0^A , the specific entropy of the austenite phase s_0^A , the specific heat capacity c , the difference in specific entropy of the austenite and martensite phase Δs^{AM} , the equilibrium temperature θ_0 , and the absolute temperature θ . Since the inelastic strains vanish with $\xi \rightarrow 0$, the inelastic strain $\mathbf{E}^i = \frac{1}{2}(\mathbf{C}^i - \mathbf{I})$ is assumed to be given by $\mathbf{E}^i = \xi \mathbf{E}^t$, where \mathbf{E}^t is a measure for the effective transformation strain in twinned regions. Additionally, the hardening-type energy is given

by (compare to Sedláček et al. [109])

$$\psi_h = kE^{\text{int}}\xi \frac{\langle \mathbf{E}^t \rangle^2}{1 - \langle \mathbf{E}^t \rangle^4} + H \left(\text{tr}(\mathbf{C}^i) - 3 \right), \quad \langle \mathbf{E}^t \rangle = \sqrt{\frac{2}{3}} \frac{\|\mathbf{E}^t\|}{k} \quad (4.4)$$

where k is the maximum transformation strain and E^{int} and H are hardening parameters. The thermal strains and jump in volume between martensite and austenite cells are modeled by $\mathbf{F}^\theta = \det(\mathbf{F}^\theta)^{\frac{1}{3}} \mathbf{I} = (1 + \varepsilon^\theta(\xi, \theta)) \mathbf{I}$ with the thermal strain

$$\varepsilon^\theta = \xi \alpha_M (\theta - \theta_{\text{refM}}) + (1 - \xi) \alpha_A (\theta - \theta_{\text{refA}}) \quad (4.5)$$

where $\alpha_{A/M}$ are the coefficients of thermal expansion and $\theta_{\text{refA/M}}$ are the reference temperatures of austenite and martensite. To model the inelastic behavior, we define the dissipation potential, depending on the direction of the transformation, to be

$$\begin{aligned} & \phi(\dot{\xi}, \mathbf{D}^i, \xi) \\ &= \begin{cases} \dot{\xi} \Delta s^{\text{AM}}((\theta_0 - M_s) + \xi(M_s - M_f)) + \sqrt{\frac{2}{3}} \sigma^{\text{reo}} \|\mathbf{D}^i\| & \dot{\xi} \geq 0 \wedge \text{tr}(\mathbf{D}^i) = 0 \\ \dot{\xi} \Delta s^{\text{AM}}((\theta_0 - A_f) + \xi(A_f - A_s)) + \sqrt{\frac{2}{3}} \sigma^{\text{reo}} \left(\left\| \frac{\dot{\xi}}{\xi} \boldsymbol{\varepsilon}^i \right\| + \|\mathbf{D}^i - \frac{\dot{\xi}}{\xi} \boldsymbol{\varepsilon}^i\| \right) & \dot{\xi} < 0 \wedge \text{tr}(\mathbf{D}^i) = 0 \\ \infty & \text{else} \end{cases} \end{aligned} \quad (4.6)$$

where $\mathbf{D}^i = \text{sym}(\dot{\mathbf{F}}^i \mathbf{F}^{i-1})$ is the symmetric part of the inelastic 'velocity gradient', $\boldsymbol{\varepsilon}^i = \frac{1}{2} \ln \mathbf{b}^i$ with $\mathbf{b}^i = \mathbf{F}^i \mathbf{F}^{i\top}$ and σ^{reo} the stress at which reorientation occurs. Furthermore, the heat conduction is assumed to follow Fourier's law with thermal conductivity κ . After further numerical treatment of the volume preservation constraint and differentiability problems of the dissipation potential, we solve the problem using an active set algorithm, which is embedded into the finite element analysis software FEAP [127]. The material parameters employed are given in Table S1 (in the Supporting Information).

4.3 Results

4.3.1 Functional properties of freestanding TiNiHf films

Thermal fatigue evaluation

The characteristic martensitic phase transformation temperatures, austenite start (A_s), austenite finish (A_f), martensite start (M_s), martensite finish (M_f), and latent heat of transformation of forward (ΔH^{AM}) and reverse transformations (ΔH^{MA}) are determined by DSC. Figure 4.1a) shows the first DSC cycle of 21 μm , 10 μm , and 5 μm thick freestanding films depicting a reduction in all characteristic transformation temperatures, with decreasing film thickness. The thermal transformation temperatures, latent heat, and thermal hysteresis

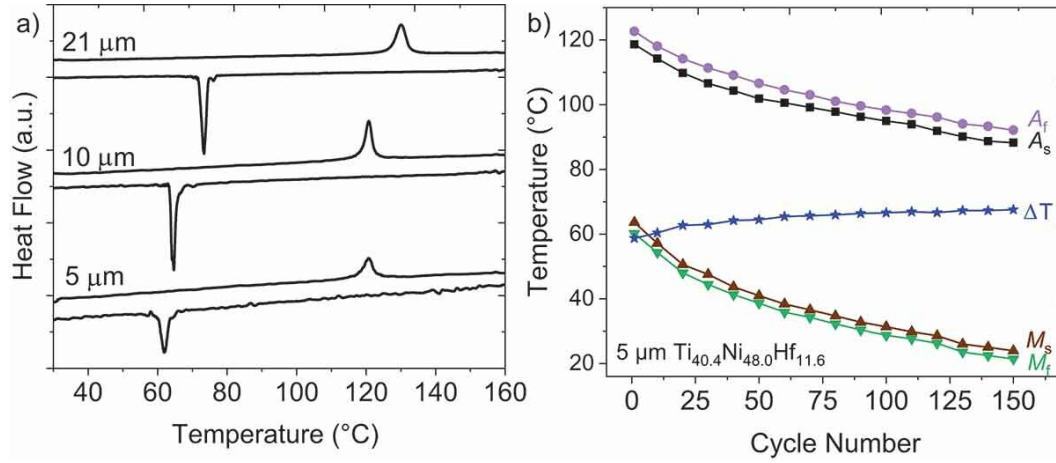


Fig. 4.1: a) First DSC cycle showing Ni_{47.7}Ti_{40.7}Hf_{11.6} freestanding films annealed at 635 °C - 5 min with decreasing film thickness from 21 μm to 5 μm. b) Change in characteristic transformation temperatures in 5 μm freestanding sample in first 150 thermal cycles.

Tab. 4.1: Thermal transformation properties of fabricated freestanding TiNiHf films, determined by DSC.

Cycle #	Thickness/μm	$A_s/^\circ\text{C}$	$A_f/^\circ\text{C}$	$\Delta H^{\text{MA}}/\frac{\text{J}}{\text{g}}$	$M_s/^\circ\text{C}$	$M_f/^\circ\text{C}$	$\Delta H^{\text{AM}}/\frac{\text{J}}{\text{g}}$	$\Delta T/^\circ\text{C}$
1	21	127.3	132.6	17.5	71.9	74.5	21.9	56.8
1	10	119.0	122.2	14.7	65.3	63.5	19.5	56.2
1	5	118.7	122.4	9.2	63.6	60.2	15.8	58.7
50	5	101.9	106.6	15.0	41.0	38.6	17.0	64.4
90	5	96.3	99.6	12.1	32.8	30.4	16.9	66.4
120	5	91.9	96.1	10.3	28.6	26.2	15.9	66.6
140	5	88.7	93.3	11.4	25.0	22.4	14.2	66.9

(ΔT) are given for all samples in Table 4.1. Only a minor influence on thermal hysteresis is noticeable with a reduction in TiNiHf film thickness from 21 μm to 5 μm. However, the transition temperatures, latent heat of transformation, and thermal hysteresis width of TiNiHf films can also be slightly increased by annealing the films at a higher temperature as shown in our previous work [4].

Bi-directional and bistable SMA-based micro- and nano-actuators ideally should last for thousands of actuation cycles with little functional fatigue. The wide thermal hysteresis of $\sim \Delta T = 57^\circ\text{C}$ indicates poor crystallographic compatibility between the martensite and austenite phase for the fabricated TiNiHf alloy [129]. Figure 4.1b) shows functional fatigue of 5 μm thick film after 150 thermal cycles between 0 °C and 140 °C. A decrease in all characteristic transformation temperatures can be seen where the largest change occurs within the first ~ 20 cycles. The A_s , A_f , M_s , and M_f temperatures continue to steadily decrease with every cycle without stabilization up to 150 cycles. A decrease in transition temperatures with thermal cycling is expected in TiNi-based alloys due to the introduction of dislocations

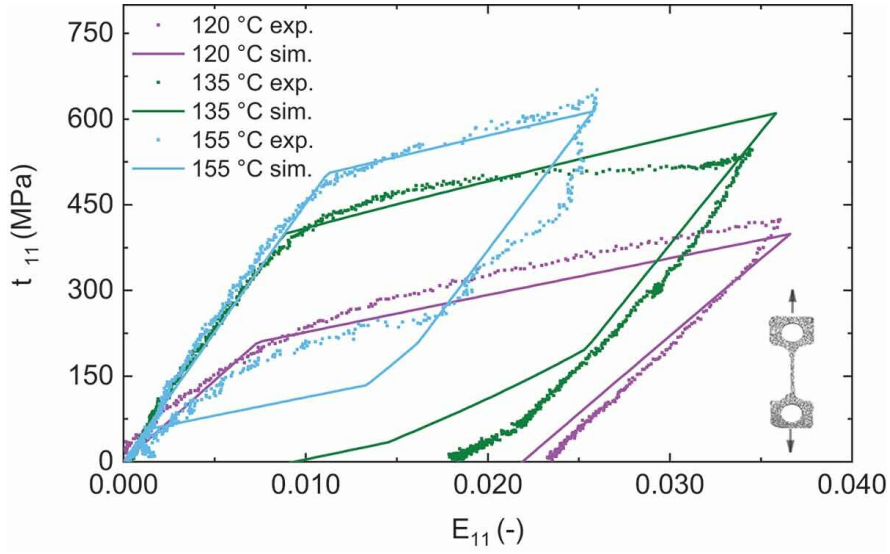


Fig. 4.2: Experimental and simulated stress-strain curve of the investigated 21 μm thick TiNiHf freestanding films of dogbone geometry size of 500 $\mu\text{m} \times 4 \text{ mm}$ (see inset). The tests are performed at three different ambient temperatures of 120 $^{\circ}\text{C}$, 135 $^{\circ}\text{C}$, and 155 $^{\circ}\text{C}$ with a strain rate of $1 \times 10^{-3} \frac{1}{\text{s}}$.

that compensate for the crystallographic compatibility between the martensite and austenite phases [129]. This fatigue behavior is observed in fabricated freestanding TiNiHf alloys indicated by a slight increase in ΔT with every cycle. Decreasing film thickness and thermal cycling lead to a significant reduction in the latent heat of both the forward and reverse transformation (ΔH^{AM} , ΔH^{MA}), which might be beneficial for SMA actuation [125].

Tensile testing

Figure 4.2 shows a comparison between experimental and simulated stress-strain characteristics of 21 μm thick freestanding TiNiHf films annealed at 635 $^{\circ}\text{C}$ - 5 min, structured into a dogbone geometry, at various temperatures in the phase transformation regime. The samples are investigated at three different ambient temperatures of 120 $^{\circ}\text{C}$, 135 $^{\circ}\text{C}$, and 155 $^{\circ}\text{C}$ at a constant strain rate of $1 \times 10^{-3} \frac{1}{\text{s}}$. Simulation results are fit to the experimental results to obtain mechanical parameters such as elastic modulus, maximum transformation strain, critical loading stress, and critical unloading stress values for the stress-induced martensitic transformation.

At 120 $^{\circ}\text{C}$, one-way shape memory behavior is obtained upon loading/unloading the samples up to 3.5 % strain, indicating the material is in the martensite phase [130]. Reorientation of the martensite variants (detwinning) is suggested to introduce the slip of dislocations. The lack of a stress plateau in Ni-lean TiNiHf alloys is characterized by their low critical stress for slipping [129], strong work hardening behavior, and continuous yielding [75]. The open hysteresis loop with a residual strain of 2.4 % indicates the retained stress-induced marten-

site that is known to form when loading an SMA at a temperature $M_s < T < A_f$ [130]. At 135 °C, the critical stress (σ_{crit}) for phase transformation to martensite is about 430 MPa and a maximum tested strain of 3.4 %. The open hysteresis loop with a residual strain of 1.8 % indicates the sample is not fully transformed to the austenite phase by 135 °C. Upon loading and unloading at 155 °C, the critical stress for martensitic transformation shifts to 490 MPa, and the reverse transformation starts at 260 MPa. As a result, a superelastic behavior is obtained with recoverable strain up to 2.6 %. The simulated model fit for the 135 °C tensile test shows a certain discrepancy with experimental data. Upon unloading, the simulation model overestimates the recoverable strain from 1.8 % to 0.95 %. However, the superelastic and martensitic forward transformation behavior for all three test conditions is well described by the simulation model. The small discrepancy could be due to microstructural and grain size effects which currently are not captured by the model. This includes the hardening effect that is responsible for the steep slope instead of a stress plateau for TiNiHf films.

4.3.2 Functional properties of TiNiHf/Si and TiNiHf/SiO₂/Si bimorphs

Electrical resistance

The film thickness dependencies of the phase transformation properties of TiNiHf films on Si and SiO₂/Si substrates are investigated to understand their impact on bi-directional actuation performance. The transformation hysteresis curves are compared in Figure 4.3 for all investigated thicknesses from 5 µm down to 0.11 µm on both substrates. Decreasing film thickness leads to a significant increase in resistance with a similar thermal hysteresis width until a critical film thickness is reached for very thin films.

Figure 4.4 summarizes the corresponding thickness dependence of transformation temperatures and thermal hysteresis width ΔT of bimorph films. Data is not shown for the 0.11 µm film on Si substrate as electrical resistance measurements revealed only a partial phase transformation ($M_s = -157.4$ °C and $A_f = -23.4$ °C) when cooling the sample from room temperature to -160 °C. According to Figure 4.4, ΔT witnesses a dramatic increase when the film thickness is below 0.44 µm for films on Si substrate and below 0.22 µm for films on SiO₂/Si substrate, which appear to be critical thicknesses on the corresponding substrates. When the film thickness is below this critical value, the thermal hysteresis shifts to a lower temperature range and the hysteresis width nearly doubles. In contrast, when the film thickness is larger than the critical thickness, the thermal hysteresis width and transformation temperatures show only minor variations.

Structural properties

TiNiHf is known to undergo a solid-to-solid phase transformation from a low-symmetry, low-temperature, martensite phase (monoclinic, B19') to a high-symmetry, high-temperature, austenite phase (cubic, B2). The crystal structures of all TiNiHf films (between 110 nm and

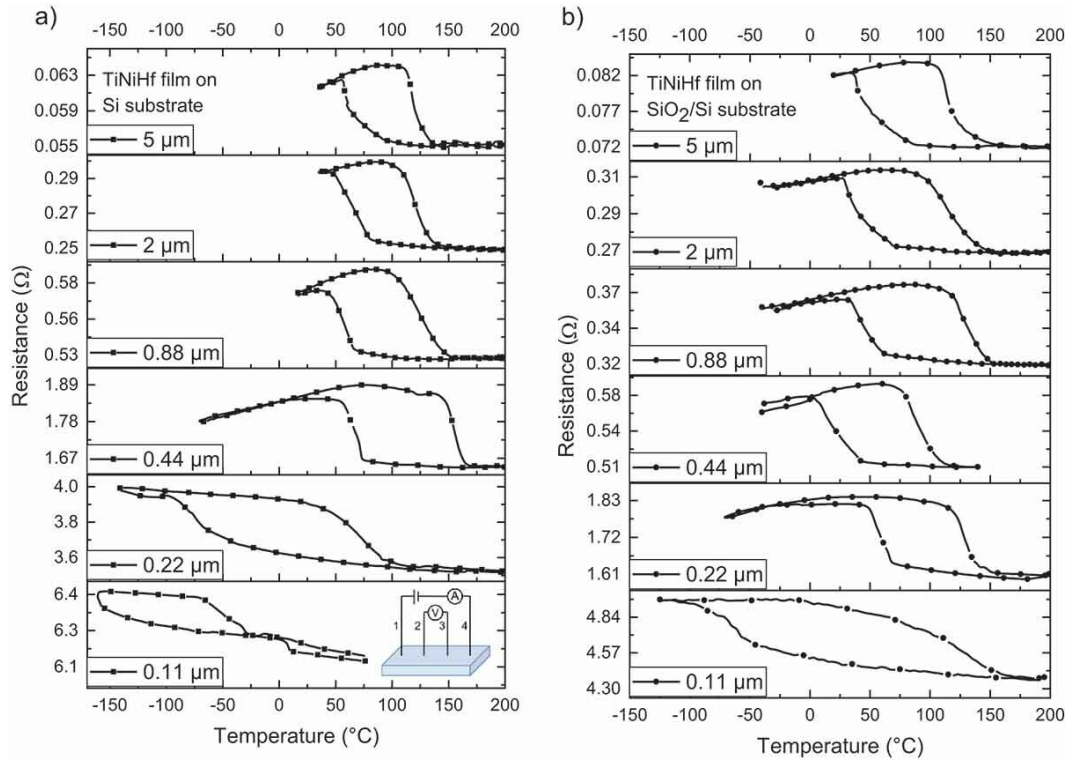


Fig. 4.3: Comparison of resistance-temperature curves of TiNiHf films of different thicknesses on a) Si and b) SiO₂/Si substrates of 5 mm × 10 mm (see inset for the schematic). Hysteresis width gets significantly larger when films are below critical thickness, which are 0.44 μm for films on Si substrate and 0.22 μm for films on SiO₂/Si substrate.

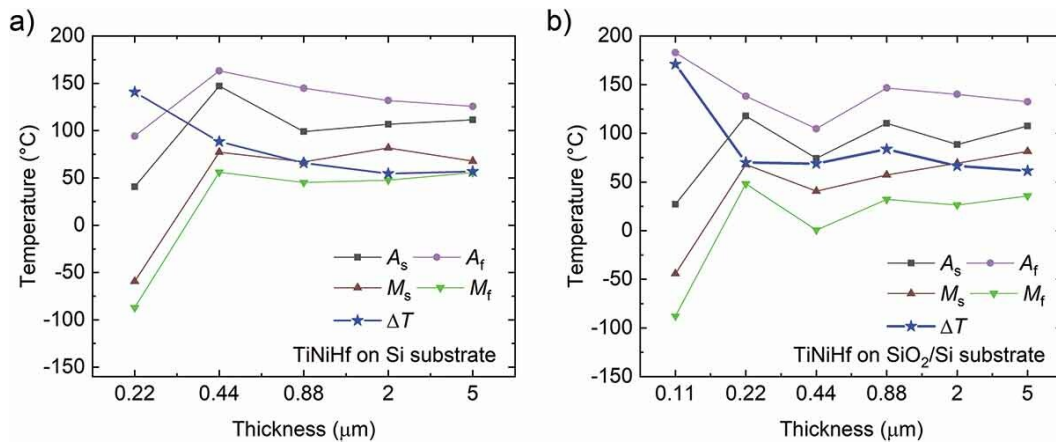


Fig. 4.4: Transformation temperatures and calculated thermal hysteresis width ΔT of TiNiHf films with different thicknesses on a) Si substrate and (b) SiO₂/Si substrate. Data is not shown for 0.1 μm-thin film on Si substrate, as transformation hysteresis is not fully seen in the temperature range -160°C to 200°C . All films were annealed at the same condition 635°C - 5 min.

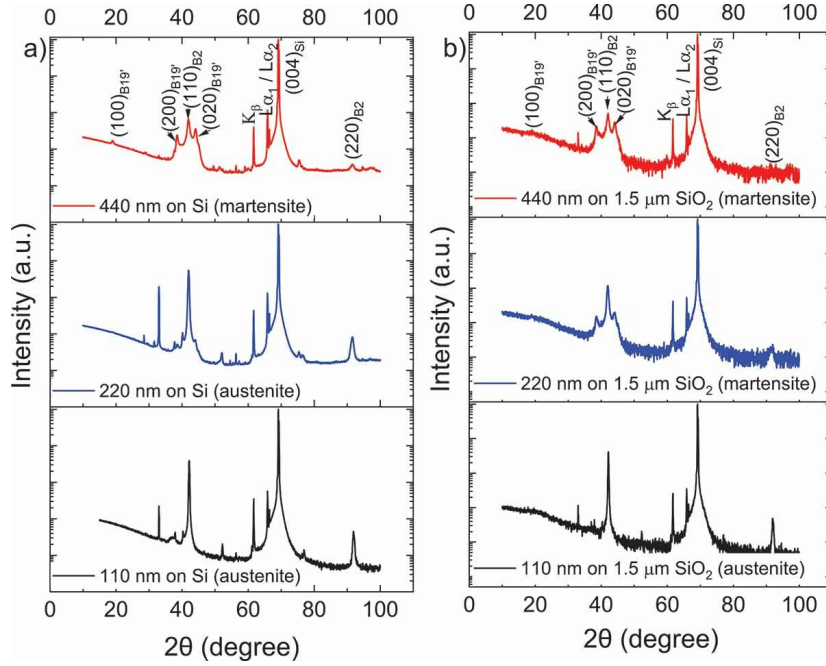


Fig. 4.5: XRD results at 30 °C for 110 nm, 220 nm, and 440 nm TiNiHf films on a) Si substrates and b) on 1.5 μm SiO₂/Si substrates. Samples were 20 mm \times 20 mm in size.

5 μm) on Si and SiO₂ substrates were evaluated at room temperature with XRD. Figure 4.5 shows XRD results at 30 °C for 110 nm, 220 nm, and 440 nm TiNiHf films on both Si and 1.5 μm SiO₂/Si substrates. The preferred orientations of martensite and (110) austenite peaks are labeled. Two dominant peaks arise for the martensite phase at 40° for the (002) B19' phase and 44° for the (020) B19' phase. The dominant peaks for the austenite phase are located at 42° for the (110) B2 phase and 91° for the (220) B2 phase. The absence (or a weak signal) of these two austenite peaks is the indication that the sample is in martensite phase at room temperature. The large signal at 69° is the peak related to the (100) Si substrates. There were a handful of unidentifiable secondary phase peaks for TiNiHf films on Si substrates that were not observed for TiNiHf films on SiO₂/Si substrates. The additional peaks could be due to a (Ti,Hf)₂Ni phase that occurs when using Ti-rich TiNi film compositions [56, 129]. For a more precise study on the precipitate phases, additional TEM measurements are necessary, which are out of scope for the current study.

To realize many applications, the main criterion for selecting the film and film composite for martensitic phase transformation has to be above room temperature transformation. Figure 4.5a) shows that 110 nm and 220 nm TiNiHf films on Si substrates are in austenite phase at room temperature. However, Figure 4.5b) for 220 nm films on SiO₂/Si buffer layers, XRD shows the SMA is composed of martensite B19' phase with some residual austenite at room temperature. XRD results indicate the minimum film thicknesses required to obtain a transformation above room temperature are \sim 200 nm on SiO₂/Si substrates and \sim 440 nm

Tab. 4.2: Transition temperatures of TiNiHf films on Si substrates and SiO₂/Si substrates for different thicknesses determined by XRD.

Substrate	Thickness/ μm	$A_s/^{\circ}\text{C}$	$A_f/^{\circ}\text{C}$	$M_s/^{\circ}\text{C}$	$M_f/^{\circ}\text{C}$	$\Delta T/^{\circ}\text{C}$
Si substrate	5	88.2	114.2	72.7	50.4	39.6
	0.88	95.3	110.1	69.3	56.9	39.6
	0.44	95.4	108.7	67.9	56.5	39.9
	0.22	No transformation between 30 $^{\circ}\text{C}$ and 130 $^{\circ}\text{C}$				
	0.11	No transformation between -100 $^{\circ}\text{C}$ and 130 $^{\circ}\text{C}$				
1.5 μm SiO ₂ /Si substrate	0.44	92.9	108.3	68.2	50.9	41.1
	0.11	No transformation between 30 $^{\circ}\text{C}$ and 130 $^{\circ}\text{C}$				

on Si substrates. This confirms both the dramatic film thickness effect below critical film thicknesses and the substrate dependence of this film thickness effect observed in electrical resistance measurements. Fig. S1 (Supporting Information) shows all TiNiHf film thicknesses above the critical thickness value on SiO₂/Si substrates transform from the martensite phase at 30 $^{\circ}\text{C}$, to the austenite phase upon heating to 110 $^{\circ}\text{C}$. 220 nm and 440 nm films on SiO₂/Si substrates were fully transformed to the austenite phase by 110 $^{\circ}\text{C}$. Films with a thickness greater than 880 nm on SiO₂/Si substrates still showed weak (200) and (020) martensite peaks indicating that these samples are still partially martensite at 110 $^{\circ}\text{C}$.

Temperature-dependent XRD measurements were used to determine the transformation temperatures of TiNiHf films with thicknesses of 110 nm, 220 nm, 440 nm, 880 nm, and 5 μm on Si substrates and a 440 nm thick TiNiHf film on 1.5 μm SiO₂/Si substrates. The samples were heated and cooled in 5 $^{\circ}\text{C}$ increments between 30 $^{\circ}\text{C}$ and 130 $^{\circ}\text{C}$ at a heating rate of 10 $\frac{^{\circ}\text{C}}{\text{s}}$. By monitoring the change in peak intensity of the B2 austenite peaks at 42 $^{\circ}$ and 91 $^{\circ}$ and using a Rietveld refinement on the (110) B2 cubic austenite phase, it was possible to determine the SMA transition temperatures [131]. The results of the transformation temperatures for all tested films are given in Table 4.2. 110 nm films on SiO₂/Si substrates were in the cubic austenite phase, with no transformation taking place in the temperature range between 30 $^{\circ}\text{C}$ and 130 $^{\circ}\text{C}$. 220 nm films on Si substrates were also found to be in the cubic austenite phase at 30 $^{\circ}\text{C}$ with no transformations taking place between 30 $^{\circ}\text{C}$ and 130 $^{\circ}\text{C}$ (Fig. S2, Supporting Information). 440 nm TiNiHf films on Si substrates were martensite at room temperature and displayed a high temperature phase transformation with $A_s = 95.4$ $^{\circ}\text{C}$ and $A_f = 108.7$ $^{\circ}\text{C}$ (Fig. S3, Supporting Information). There was not a significant change in transformation temperatures for 440 nm on Si versus 440 nm films on 1.5 μm SiO₂/Si substrates. Scanning electron microscopy cross-sectional images of different TiNiHf film thicknesses (5 μm , 2 μm , 880 nm, and 440 nm) on Si substrates are shown in Fig. S4 (Supporting Information).

Temperature-dependent XRD data was used to estimate the CTE for austenite phase of the Ti_{40.4}Ni₄₈Hf_{11.6} alloy by monitoring the change in the cubic lattice parameter while cooling the SMA over a temperature range of 120 $^{\circ}\text{C}$ -85 $^{\circ}\text{C}$ [154]. The CTE of the austenite phase (α_A) was calculated to be $30 \times 10^{-6} \frac{1}{\text{K}} \pm 5 \times 10^{-6} \frac{1}{\text{K}}$ from XRD measurements. This

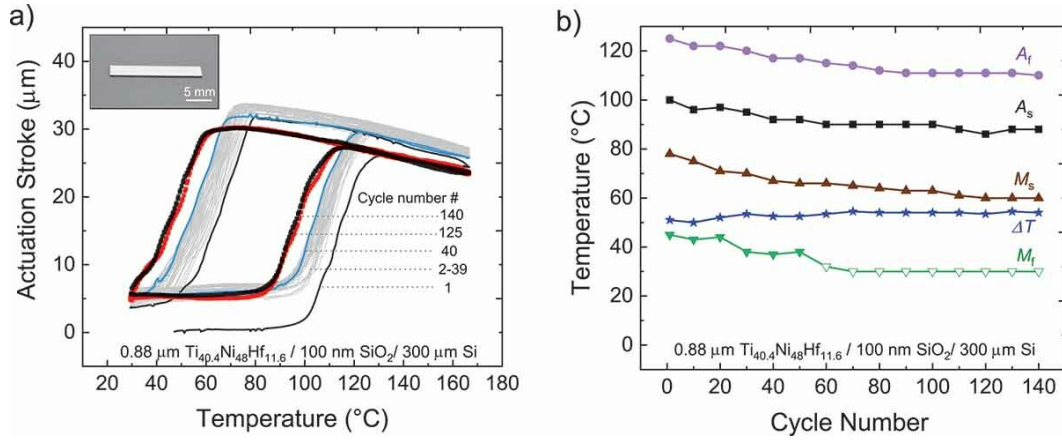


Fig. 4.6: a) Deflection versus temperature measurements of bi-directional TiNiHf/SiO₂/Si actuators with TiNiHf film thicknesses of 0.88 μm on 100 nm buffer layer of SiO₂ on 300 μm Si substrates (see inset for the sample picture). b) Influence of functional fatigue on transition temperatures for the 0.88 μm TiNiHf/SiO₂/Si composite. Cantilevers were 3.5 mm × 20 mm in size, with a freestanding cantilever length of 14 mm during measurement.

value falls between the austenite CTEs reported for binary NiTi ($\alpha_A = 13 \times 10^{-6} \frac{1}{K}$) [99] and for a Ti_{29.7}Ni_{50.3}Hf₂₀ alloy ($\alpha_A = 49.6 \times 10^{-6} \frac{1}{K} \pm 3.4 \times 10^{-6} \frac{1}{K}$) [113]. For computational simulations, the calculated value from XRD data is used ($\alpha_A = 30 \times 10^{-6} \frac{1}{K}$), and for simplicity, the CTE value for martensite is estimated to be half the value of austenite ($\alpha_M = 15 \times 10^{-6} \frac{1}{K}$).

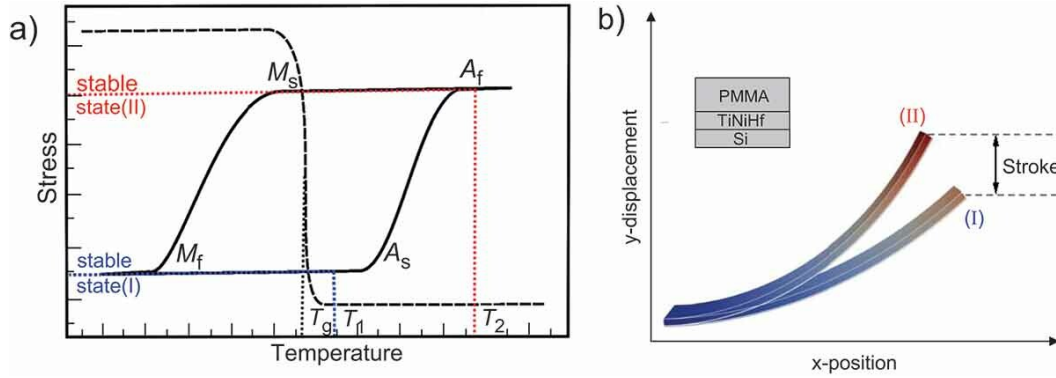
4.3.3 Bi-directional actuation based on TiNiHf/SiO₂/Si bimorph film composites

Actuation in SMA/Si bimorph cantilevers arises from the combined shape memory effect and CTE difference between bimorph layers [49]. The different stress states in the bimorph cantilever that lead to desirable deflection arise when heating the composite to temperatures above A_f and cooling the composite below M_f . The maximum actuation stroke is determined by the y -deflection obtained between the martensite and austenite phases using CDM.

Figure 4.6a) shows 145 cycles of deflection for 0.88 μm TiNiHf/100 nm SiO₂/300 μm Si cantilever composite. The transition temperatures, thermal hysteresis, and maximum stroke upon thermal cycling are given in Table 4.3. A decrease in transformation temperatures and a slight decrease in the actuation stroke are observed with increasing cycles. Figure 4.6b) shows the change in characteristic transformation temperatures as a function of thermal cycles for 0.88 μm TiNiHf/100 nm SiO₂/300 μm Si bi-directional actuator. The thermal-mechanical hysteresis stabilized after ~ 120 cycles, with minor shift in transformation temperatures between cycles 120 and 145. After ~ 60 cycles, the M_f temperature dropped below minimum test temperature of 30 °C; therefore, the material was not fully transformed. The M_f temperature is marked with a green open symbol in Figure 4.6b) to indicate $M_f < 30$ °C. Very

Tab. 4.3: Influence of functional fatigue on bi-directional stroke and characteristic transformation temperatures of different TiNiHf thicknesses on 100 nm SiO₂/Si substrates.

Cycle #	Thickness/ μm	$A_s/^\circ\text{C}$	$A_f/^\circ\text{C}$	$M_s/^\circ\text{C}$	$M_f/^\circ\text{C}$	$\Delta T/^\circ\text{C}$	Stroke/ μm
1	2	110	134	85	59	50	39
40	2	101	127	74	45	54.5	35
1	0.88	100	125	78	45	51	32
40	0.88	92	117	67	37	52.5	31.5
140	0.88	88	110	60	<30	54	31.5

**Fig. 4.7:** a) Schematic of interlaced polymer and SMA hysteresis, enabling bistability. b) Superimposed stable states (I) and (II) for comparison. The cross section of the trimorph layers (PMMA/TiNiHf/Si) is sketched.

small reduction in maximum actuation stroke from 32 μm to 31.5 μm within the first 40 cycles (i.e. an indication of functional fatigue) is observed which then stabilized in the subsequent cycles.

4.3.4 Bistable actuation based on PMMA/TiNiHf/Si trimorph film composites

Bistable actuation principle

The unique properties of TiNiHf/Si bimorph films of large hysteresis and an A_f temperature larger than 105 $^\circ\text{C}$ open the opportunity for the design of bistable actuation using a polymer layer on TiNiHf/Si composite. As depicted in the schematic in Figure 4.7a), SMA/Si bimorph structures can become bistable if combined with a polymer with matching glass transition temperature (T_g). That means the glass transition temperature of the polymer is required to fall within the thermal hysteresis of the SMA ($M_s < T_g < A_s$) [143]. Assume the composite is in stable state I, as shown in Figure 4.7b). In order to switch to stable state II, the composite is heated above austenite finish temperature ($T_g < A_f < T_2$), allowing the polymer to soften and adapt to the deflected shape of the composite. Upon cooling below T_g , the polymer will become hard and, thereby, fixes the deflected shape that corresponds to stable state II. In order to recover stable state I, the composite is heated to intermediate temperature T_1

($T_g < T_1 < A_s$), and the polymer becomes soft again and adapts to the initial shape in martensite phase, which will be fixed upon cooling to room temperature. The fixation effect will be effective if the thickness of the polymer is large enough.

Here, we use PMMA as the polymer in a PMMA/TiNiHf/Si trimorph composite. PMMA is a standard resist material in various lithography methods such as X-ray lithography, deep UV-lithography, and E-beam lithography techniques [4, 29, 46, 60]. The glass transition temperature of PMMA is 105 °C, which fulfills the criteria for bistability.

Polymer and Si model

Since the PMMA in mind does not exhibit shape memory effects, the polymer is modeled as a passive material. This is done by using a thermally coupled viscoelastic Maxwell model for finite strains [102], where the viscosity is 1×10^7 MPa s at low temperatures below $T_g = 105$ °C and 1 MPa s at high temperatures. Thus, the polymer is soft above T_g , while still being stiff at colder temperatures. For simplicity, the CTE ($70 \times 10^{-6} \frac{1}{K}$) [31, 42, 158], Young's modulus (500 MPa), and Poisson's ratio (0.4) of the PMMA polymer are held constant in simulations. The Si layer is modeled using a thermally coupled Neo-Hookean elastic model with Young's modulus of 130 GPa and Poisson's ratio being equal to 0.278 [42]. The CTE of Si is $2.6 \times 10^{-6} \frac{1}{K}$ [138]. The SiO₂ buffer layer is neglected in simulations as it is not expected to influence the mechanical properties of the composite.

Actuator geometry and boundary conditions

The modeled actuator has a length of 1 mm and a width of 100 μ m. The Si and SMA films have a thickness of 2 μ m and 1 μ m, respectively. The geometry is discretized by 20 elements over its length and 6 over its width, whereas each layer of material is discretized by 4 elements over the thickness. At the left side, it is clamped, i.e., the displacements on the left side are constrained to be zero. Additionally, at the top and bottom of the thin film, a Robin boundary condition with a convective heat transfer coefficient of $70 \frac{W}{m^2 K}$ [51] is used to model heat convection with the surrounding air, which has a temperature of 20 °C. Furthermore, Joule heating is realized through a heat source term in the TiNiHf. The actuator model is shown in Figure 4.7b).

Optimization of polymer thickness

The actuator's maximum bistable stroke is strongly dependent on the polymer layer thickness. If the layer is too thin, it is not able to hold the austenite-related shape at room temperature. If the polymer layer is too thick, it hampers the stroke and limits the actuation speed as well as power consumption. Additionally, larger body forces and less freedom for the thermal hysteresis of the SMA limit the design space of the actuator. Therefore, we optimize the

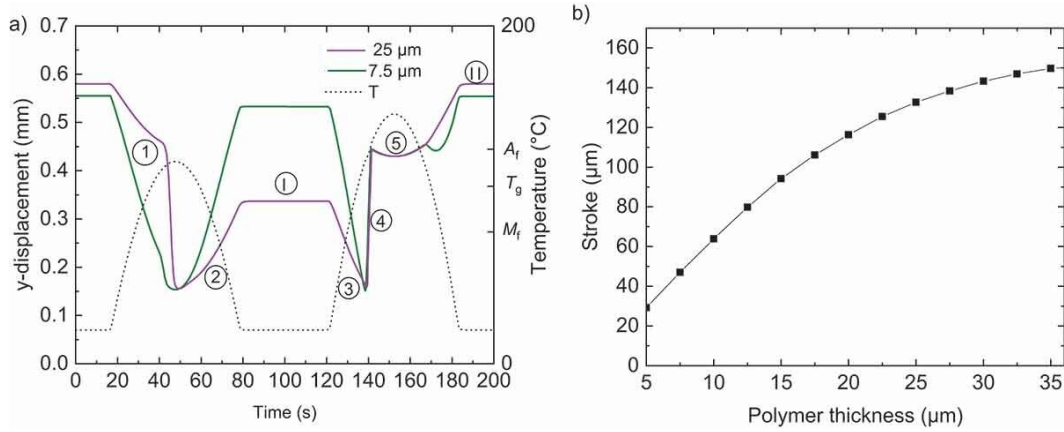


Fig. 4.8: a) Demonstration of bistable actuation of PMMA/TiNiHf/Si through simulation assuming a T_g of 105 °C. b) FEM simulation of maximum bistable stroke versus polymer thickness for a cantilever consisting of a PMMA/TiNiHf/Si trimorph composite with TiNiHf and Si layer thicknesses of 1 μm and 2 μm, respectively. The cantilever length L and width w are 1 mm mm and 100 μm, respectively.

Tab. 4.4: Simulated stroke of downsized trimorph actuator geometries.

Length L	Width w	Layer thicknesses t			Bistable stroke Δy	Relative bistable stroke $\Delta y/L$
		Si	SMA	Polymer		
1 mm	100 μm	2 μm	1 μm	20 μm	116 μm	12 %
100 μm	10 μm	2 μm	1 μm	20 μm	4.4 μm	4.4 %
20 μm	2 μm	0.4 μm	0.2 μm	4 μm	0.8 μm	4 %

polymer geometry by simulating several thicknesses. For clarity, only two polymer thicknesses of this sweep are shown in Figure 4.8a), where the achieved vertical displacement and the temperature in the middle of the SMA at the end of the beam are plotted over time. In the beginning, we start with the actuator at room temperature, which is held by the polymer in stable state (I). Subsequently, the actuator undergoes a low heat cycle (1-2) in Figure 4.8a), in which the polymer softens and releases the stress, which results in the flattened, stable state (I). After that, the actuator is heated above A_f (3-4). Next, it is cooled down to room temperature again (5), at which it returns to stable state (II) again.

When reaching M_f , depending on the polymer thickness, it can hold the shape (e.g. for 25 μm) or release most of the actuation (e.g. for 7.5 μm, shown in Figure 4.8a)). Additionally, since the area moment of inertia depends in a cubic manner on the polymer layer thickness, any thinner layers result in less achievable stroke. The influence of the polymer layer onto the achievable bistable stroke is depicted in Figure 4.8b), where the difference between the vertical displacement of stable states (I) and (II) is compared. Furthermore, favorable thickness combinations for smaller actuator geometries and stack sizes are listed in Table 4.4. The maximum relative bistable strokes ($\Delta y/L$) show a scaling-dependent decrease when reducing the lateral dimensions $L \times w$ from 1000 μm × 100 μm down to 20 μm × 2 μm. Yet, even for the

smallest bistable device, a considerable bistable stroke of 4 % is expected, which is of special interest for actuator applications at ultra-small scales. Because the model at hand does not inherit any size effects, further model refinement might be required considering additional effects such as grain size, oxidation, compositional or other influences affecting the actuation behavior. Furthermore, the T_g of PMMA is dependent on parameters such as film thickness and polymer molecular weight. Additionally, both the T_g and CTE are dependent on the interfacial energy between the polymer film and substrate and the exposure dose used during structuring [29, 46, 60]. These processing and scaling effects are not captured by the model and must be taken into account when designing bistable actuators with PMMA/TiNiHf/Si trimorph composites.

4.4 Discussion

The shape memory properties of TiNiHf/SiO₂/Si bimorph composites were characterized experimentally for use as micro and nano bi-directional actuators. Controlling functional fatigue characteristics upon thermal cycling in Ni-lean TiNiHf is pivotal in the development of bi-directional actuators. DSC and CDM results showed that the transition temperatures of fabricated TiNiHf freestanding films and TiNiHf/SiO₂/Si bimorph composites decrease with thermal cycling. The change in transition temperatures with cycling is reduced for TiNiHf films bound to an SiO₂/Si substrate compared to the functional fatigue DSC results reported for freestanding films from Table 4.1. An average reduction in transition temperatures of 12 °C - 18 °C is obtained for 0.88 μm TiNiHf/SiO₂/Si composites after 140 cycles compared to an average reduction in transition temperatures of 29 °C - 38 °C for 5 μm freestanding TiNiHf films after 140 cycles. Additionally, XRD and CDM show thermal hysteresis width is also reduced for TiNiHf films constrained by SiO₂ and Si substrates compared to freestanding films of similar thicknesses. This might imply an improvement in the compatibility of the austenite and martensite phases for films constrained by a substrate. The Si and SiO₂/Si substrates cause an inhomogeneous stress profile at the film-substrate interface, which may affect phase transformation temperatures as it poses an additional energy barrier for the formation of martensite upon cooling [59].

Based on the analysis of our experimental results, we can infer a pronounced substrate influence as well as a dramatic film thickness effect below a substrate-dependent critical film thickness on the phase transformation properties of TiNiHf films. Freestanding TiNiHf films show a thickness-dependent decrease of all characteristic transformation temperatures of ~ 10 °C when reducing thickness from 21 μm to 5 μm. Similarly, decreasing the SMA's thickness was found to decrease the martensitic transformation temperatures for SMA wires (e.g. NiMnGa [51]), freestanding films (e.g. TiNiCu [55], NiTi [44, 134]), and sputtered bimorph systems (e.g. NiMnGa/Si [59], NiTi/Si [44], NiTi/Pt [45, 63], Ti₂NiCu/Pt [55]). This thickness effect can also be seen for TiNiHf with decreasing film thickness from 5 μm down to 110 nm constrained on Si and SiO₂/Si substrates. There is a critical thickness of

TiNiHf films on both Si and SiO₂/Si substrates, below which the thermal hysteresis width increases significantly. Temperature-dependent electrical resistance measurements and XRD measurements are in agreement that these values are between 110 nm and 220 nm for films on SiO₂/Si substrates and between 220 nm and 440 nm for films on Si substrates.

This shows evidence that the interface between the substrate and TiNiHf film plays an important role in the phase transformation and is more pronounced below a critical thickness. The existence of the observed critical thickness values could also be ascribed to a change in the composition of the TiNiHf film due to diffusion at the surface and interface of the substrate during the annealing process. One explanation of the change is a TiO_x oxidation layer on the surface of the SMA film [44, 45], which is claimed to affect films with a thickness below 1 μ m [45, 55]. The oxidation layer will create a Ti-lean zone beneath, which is much thicker than the oxidation layer, changing the composition of the film. The effect of the SiO₂ buffer layer on the change of the critical thickness to lower values might be attributed to the effect of diffusion as well. Jarrige et al. [45] propose that for binary NiTi systems, there is an additional TiO_x oxidation layer between the film and SiO₂/Si substrate, and this layer can hinder the diffusion of Ni and Si atoms between the film and the substrate, as was observed for films annealed on Si substrates. Furthermore, the stress at the TiNiHf/Si interface might be reduced due to stress release by the intermediate SiO_x layer, which affects phase transformation at small SMA layer thicknesses. These considerations help to understand that the critical thickness of TiNiHf films on SiO₂/Si substrate could be smaller than that on Si substrate. However, further investigations on the layer sequence (addition of buffer layers and diffusion barriers like Si₃N₄ or Ta) of bimorph sample sections are needed for better understanding.

The transition temperatures deduced from different experimental methods of DSC, XRD, electrical resistance, and bi-directional deflection reveal systematic deviations. The different length scales and different loading conditions may cause major differences. While XRD probes undergo phase transformation at the local scale, electrical resistance reveals average values of the whole test specimen. Bi-directional bending reflects the non-uniform stress profiles during beam bending, which is absent in DSC. Another difference is ascribed to the different time scales of temperature cycling in the phase transformation regime. In particular, electrical resistance measurements have been performed by providing sufficient waiting time between data points to enable quasi-stationary equilibrium conditions, while DSC experiments have been conducted using a fixed heating and cooling rate. However, these details are not the focus of this investigation. Conclusions are drawn based on the dependencies observed by each experimental method independently, while direct comparisons are avoided.

The experimental results shown in this work envisage a possible route for developing bistable actuators using additional polymer layer to TiNiHf/Si bimorphs. The biggest challenge in experimentally realizing such bistable devices is tailoring transformation temperatures of TiNiHf to match the T_g of PMMA. Simulations incorporated the experimental results on TiNiHf phase transformation properties accounting for substrate influence and film thickness effects that arise when downscaling to nanofilms. Simulation results indicated that trimorph

PMMA/TiNiHf/Si film composites exhibit large bistable actuation with scaling-dependent bistable stroke per length ($\Delta y/\Delta L$). The simulation results should be considered as an outlook and guideline for future experimental work.

4.5 Conclusions

For the investigated chemical composition of Ti_{40.4}Ni₄₈Hf_{11.6} and optimized annealing conditions of 635 °C - 5 min, it is found that decreasing the layer thickness of TiNiHf in bimorph composites decreases the phase transformation temperatures. A large reduction of M_s/A_s by more than 100 °C is observed when TiNiHf film thickness drops below a critical thickness. The onset of the film thickness effect depends on the substrate. For Si substrates, critical thickness occurs between 440 nm and 220 nm, while for 1.5 μm SiO₂ buffer/Si substrate it occurs between 220 nm and 110 nm. Functional fatigue was improved for TiNiHf films in bi-directional TiNiHf/SiO₂/Si composites as compared to freestanding TiNiHf films. The fabricated bi-directional actuator demonstrated a stable actuation stroke after 40 thermal cycles and thermal stability after 125 cycles. Freestanding films, on the other hand, achieved were not stabilized in the first 150 thermal cycles. Tailoring the substrate and film properties of TiNiHf films above room temperature leads to the possibility to use these films for advanced nano- or micro-actuation. To further investigate the utility of these films for the actuation, simulations were carried out by adding an additional PMMA layer to show that with this material system, bistable actuation can be achieved at nanoscales. Once functional fatigue is controlled in sputtered TiNiHf/SiO₂/Si film composites, these results show they are promising materials for nanoscale actuation to enable novel applications in nanomechanics and photonics.

Acknowledgments

S. Curtis gratefully acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. DGE 1840340. We acknowledge Kaiju Lu at the Institute for Applied Materials (KIT-IAM) for access to temperature-dependent tensile test experiments. The authors would like to thank Afrin Shara for her assistance with fabrication and thermal characterization.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research has received funding from the Deutsche Forschungsgemeinschaft (DFG) within the priority program "SPP2206-Cooperative Multistage Multistable Microactuator Systems".

CHAPTER 5

Conclusions and outlook

5.1 Conclusions

This dissertation dealt with modeling the incompressibility constraint for finite deformations in shape memory alloys. In contrast to the gold standard of making use of the exponential map, a new projection ansatz was used in this work. It allows to include the entire model formulation in the generalized standard materials framework, from which all model equations follow.

Article 1 introduced this new projection approach, which due to its simpler nature, is implemented for finite deformation von Mises elastoplasticity instead of a shape memory alloy model. The projection approach is embedded in the generalized standard materials framework, which ensures thermodynamic consistency. Besides this, it also enables better control over the model by giving the possibility to modify the time-discretized potential to, e.g., obtain a numerically favorable formulation. Additionally, a novel way of improving the logarithm function for tensors was employed to greatly improve the algorithm's convergence behavior for large time steps. Furthermore, we extended the model to include Armstrong-Frederick type kinematic hardening. Finally, it was shown that in contrast to the exponential map, our approach intrinsically preserves the system's symmetry.

Article 2 adapted the projection ansatz for the small strain shape memory alloy model by Sedláč et al. [109], which was extended to the finite strain case. The model was formulated in the generalized standard materials framework, which itself was extended to include the temperature as an additional field of unknowns. The from the potentials resulting model and evolution equations are, due to their highly nonlinear nature, numerically challenging.

To overcome the numerical difficulties, regularizations and additional energy terms are introduced and analyzed. A novel way using a logarithmic strain measure to ensure vanishing inelastic strains upon reverse transformation is employed. Furthermore, the volume change upon phase transformation that occurs in some shape memory alloys is included in the model. The model results are shown in several numerical examples.

Article 3 showed a way to manufacture TiNiHf/SiO₂/Si shape memory film composites for bidirectional micro actuation. Several different annealing conditions for magnetron sputtered thin films with different layer thicknesses are analyzed for their actuation properties and functional fatigue. Here, a multitude of experiments, including tensile testing at different temperatures, differential scanning calorimetry, electrical resistance measurements, X-ray diffraction, cantilever deflection measurements as well as scanning electron microscopy was used to obtain critical material parameters and information about the downscaling phenomena happening in the shape memory alloy we utilized. Additionally, the in the experiments obtained material parameters were used to fit the model parameters of the model developed in Article 2, which was extended to better approximate the apparent hardening behavior of the NiTiHf thin films. The model was then used to demonstrate a concept for shape memory alloy nanoactuators which utilize PMMA as an additional layer to gain bistability by curing the polymer in different stroke states.

5.2 Outlook

Modeling of shape memory alloys is still a very active area of research. Some possibilities to improve and extend the shown works are summarized in the following:

5.2.1 One to one numerical comparison of the exponential map and projection approach

The proposed projection scheme to circumvent the exponential mapping was analyzed with respect to numerical accuracy and convergence behavior in Article 1. Here, it showed a slightly worse convergence behavior with respect to the number of time steps required to obtain an accurate solution. However, another interesting aspect of replacing the exponential mapping was to circumvent its rather high numerical cost arising from the singular value decomposition. Hurtado et al. [43] showed a decreased computational cost using a similar projection ansatz, when comparing it to the exponential map. Still, for our approach, a computational performance comparison with the exponential map is yet to be done.

5.2.2 Inclusion of functional fatigue into the shape memory alloy model

Recently, many authors have included fatigue phenomena such as transformation induced plasticity into their models. This allows to predict cyclic loading behavior over many cycles as well as to analyze and optimize the training process used in many shape memory alloy

applications.

For the application to shape memory alloy nanoactuators, adding the capability for modeling functional fatigue behavior is a key ingredient in predicting the lifetime of them as well as optimizing them for long-term use.

5.2.3 Improve shape memory model for nanoscale simulations via a gradient extension

In Articles 2 and 3, the shape memory alloy model was applied to nanoactuator simulations. Here, it became apparent that for composites thin films in the nanometer range, thermo-mechanical size effects play a large role in the materials behavior. Article 3 suggests that these size effects mainly arise from the boundary layers, which might undergo diffusion. A possibility to include these effects would be to include a gradient extension for the martensite volume fraction ξ into the model.

5.2.4 Better validation by conducting more shape memory alloy experiments

The proposed shape memory alloy model possesses many mechanical, thermo-mechanical and thermal parameters. Some of these parameters were taken from the literature, while most of them were obtained by experiments in Article 3. However, to minimize the influence of possible experimental influences, there is additional need to validate the model in additional experiments.

Bibliography

- [1] Achenbach, M. and Müller, I. (1985). Simulation of material behaviour of alloys with shape memory. *Archiwum Mechaniki Stosowanej*, 37(6):573–585.
- [2] Aravas, N. (1994). Finite-strain anisotropic plasticity and the plastic spin. *Modelling and Simulation in Materials Science and Engineering*, 2(3A):483–504.
- [3] Arghavani, J., Auricchio, F., and Naghdabadi, R. (2011). A finite strain kinematic hardening constitutive model based on Hencky strain: General framework, solution algorithm and application to shape memory alloys. *International Journal of Plasticity*, 27(6):940–961.
- [4] Arivanandhan, G., Li, Z., Curtis, S., Velvaluri, P., Quandt, E., and Kohl, M. (2020). Temperature homogenization of co-integrated shape memory-silicon bimorph actuators. *Proceedings*, 64(1).
- [5] Armstrong, P. J., Frederick, C., et al. (1966). *A mathematical representation of the multiaxial Bauschinger effect*, volume 731. Berkeley Nuclear Laboratories Berkeley, CA.
- [6] Auricchio, F., Bonetti, E., Scalet, G., and Ubertini, F. (2014). Theoretical and numerical modeling of shape memory alloys accounting for multiple phase transformations and martensite reorientation. *International Journal of Plasticity*, 59:30–54.
- [7] Auricchio, F., Reali, A., and Stefanelli, U. (2007). A three-dimensional model describing stress-induced solid phase transformation with permanent inelasticity. *International Journal of Plasticity*, 23(2):207–226.
- [8] Bechtold, C., Chluba, C., Zamponi, C., Quandt, E., and de Miranda, R. L. (2019). Fabrication and characterization of freestanding NiTi based thin film materials for shape memory micro-actuator applications. *Shape Memory and Superelasticity*, 5(4):327–335.
- [9] Bellini, C., Berto, F., Di Cocco, V., and Iacoviello, F. (2021). A cyclic integrated microstructural-mechanical model for a shape memory alloy. *International Journal of Fatigue*, 153:106473.

-
- [10] Bernardini, D. and Pence, T. J. (2002). Models for one-variant shape memory materials based on dissipation functions. *International Journal of Non-Linear Mechanics*, 37(8):1299–1317.
- [11] Brepols, T., Wulfinghoff, S., and Reese, S. (2017). Gradient-extended two-surface damage-plasticity: Micromorphic formulation and numerical aspects. *International Journal of Plasticity*, 97:64–106.
- [12] Choudhary, N. and Kaur, D. (2016). Shape memory alloy thin films and heterostructures for MEMS applications: A review. *Sensors and Actuators A: Physical*, 242:162–181.
- [13] Chowdhury, P. and Sehitoglu, H. (2017). Deformation physics of shape memory alloys - Fundamentals at atomistic frontier. *Progress in Materials Science*, 88:49–88.
- [14] Cisse, C., Zaki, W., and Ben Zineb, T. (2016). A review of constitutive models and modeling techniques for shape memory alloys. *International Journal of Plasticity*, 76:244–284.
- [15] Cuitino, A. and Ortiz, M. (1992). A material-independent method for extending stress update algorithms from small-strain plasticity to finite plasticity with multiplicative kinematics. *Engineering computations*, 9(4):437–451.
- [16] Dettmer, W. and Reese, S. (2004). On the theoretical and numerical modelling of Armstrong-Frederick kinematic hardening in the finite strain regime. *Computer Methods in Applied Mechanics and Engineering*, 193(1):87–116.
- [17] Du, H., Chau, F. S., and Zhou, G. (2016). Mechanically-tunable photonic devices with on-chip integrated MEMS/NEMS actuators. *Micromachines*, 7(4):69.
- [18] Eckart, C. (1948). The thermodynamics of irreversible processes. IV. The theory of elasticity and anelasticity. *Physical Review*, 73(4):373.
- [19] Eichenfield, M., Camacho, R., Chan, J., Vahala, K. J., and Painter, O. (2009). A picogram- and nanometre-scale photonic-crystal optomechanical cavity. *nature*, 459(7246):550–555.
- [20] Errando-Herranz, C., Takabayashi, A. Y., Edinger, P., Sattari, H., Gylfason, K. B., and Quack, N. (2020). MEMS for photonic integrated circuits. *IEEE Journal of Selected Topics in Quantum Electronics*, 26(2):1–16.
- [21] Es-Souni, M., Es-Souni, M., and Fischer-Brandies, H. (2005). Assessing the biocompatibility of NiTi shape memory alloys used for medical applications. *Analytical and Bioanalytical Chemistry*, 381(3):557–567.

- [22] Eterovic, A. L. and Bathe, K.-J. (1990). A hyperelastic-based large strain elasto-plastic constitutive formulation with combined isotropic-kinematic hardening using the logarithmic stress and strain measures. *International Journal for Numerical Methods in Engineering*, 30(6):1099–1114.
- [23] Evrigen, A., Karaman, I., Santamarta, R., Pons, J., Hayrettin, C., and Noebe, R. (2016). Relationship between crystallographic compatibility and thermal hysteresis in Ni-rich Ni-TiHf and NiTiZr high temperature shape memory alloys. *Acta Materialia*, 121:374–383.
- [24] Fischer, A. (1992). A special Newton-type optimization method. *Optimization*, 24(3-4):269–284.
- [25] Fish, J. and Shek, K. (2000). Finite deformation plasticity based on the additive split of the rate of deformation and hyperelasticity. *Computer Methods in Applied Mechanics and Engineering*, 190(1):75–93.
- [26] Flory, P. J. and Volkenstein, M. (1969). Statistical mechanics of chain molecules. *Biopolymers*, 8(5):699–700.
- [27] Frost, M., Benešová, B., and Sedlák, P. (2016). A microscopically motivated constitutive model for shape memory alloys: Formulation, analysis and computations. *Mathematics and Mechanics of Solids*, 21(3):358–382.
- [28] Frost, M., Benešová, B., Seiner, H., Kružík, M., Šittner, P., and Sedlák, P. (2021). Thermomechanical model for NiTi-based shape memory alloys covering macroscopic localization of martensitic transformation. *International Journal of Solids and Structures*, 221:117–129. Special Issue dedicated to papers from the International Union of Theoretical and Applied Mechanics 2019 Symposium on Phase Transformations in Shape Memory Materials: Modeling and Applications.
- [29] Fryer, D. S., Peters, R. D., Kim, E. J., Tomaszewski, J. E., de Pablo, J. J., Nealey, P. F., White, C. C., and Wu, W.-l. (2001). Dependence of the glass transition temperature of polymer films on interfacial energy and thickness. *Macromolecules*, 34(16):5627–5634.
- [30] Gall, K., Sehitoglu, H., Chumlyakov, Y. I., and Kireeva, I. V. (1999). Tension-compression asymmetry of the stress-strain response in aged single crystal and polycrystalline NiTi. *Acta Materialia*, 47(4):1203–1217.
- [31] Goods, S. H. (2003). Thermal expansion and hydration behavior of PMMA moulding materials for LIGA applications.
- [32] Govindjee, S., Mielke, A., and Hall, G. J. (2003). The free energy of mixing for n-variant martensitic phase transformations using quasi-convex analysis. *Journal of the Mechanics and Physics of Solids*, 51(4):I–XXVI.

-
- [33] Grummon, D. S. (2003). Thin-film shape-memory materials for high-temperature applications. *JOM*, 55(12):24–32.
- [34] Hackl, K. (1997). Generalized standard media and variational principles in classical and finite strain elastoplasticity. *Journal of the Mechanics and Physics of Solids*, 45(5):667–688.
- [35] Hairer, E. (2001). Geometric integration of ordinary differential equations on manifolds. *BIT Numerical Mathematics*, 41(5):996–1007.
- [36] Halphen, B. and Nguyen, Q. S. (1975). Sur les matériaux standard généralisés. *Journal de Mécanique*, 14(1):39–63.
- [37] Han, S., Seok, T. J., Quack, N., Yoo, B.-W., and Wu, M. C. (2015). Large-scale silicon photonic switches with movable directional couplers. *Optica*, 2(4):370–375.
- [38] Han, W. and Reddy, B. D. (1999). *Plasticity: Mathematical Theory and Numerical Analysis*, volume 9. Springer Science & Business Media.
- [39] Hartl, D. J., Chatzigeorgiou, G., and Lagoudas, D. C. (2010). Three-dimensional modeling and numerical analysis of rate-dependent irrecoverable deformation in shape memory alloys. *International Journal of Plasticity*, 26(10):1485–1507.
- [40] Helm, D. (2006). Stress computation in finite thermoviscoplasticity. *International Journal of Plasticity*, 22(9):1699–1727.
- [41] Hoffmann, M., Bezzaoui, H., and Voges, E. (1994). Micromechanical cantilever resonators with integrated optical interrogation. *Sensors and Actuators A: Physical*, 44(1):71–75.
- [42] Hopcroft, M. A., Nix, W. D., and Kenny, T. W. (2010). What is the Young’s modulus of silicon? *Journal of Microelectromechanical Systems*, 19(2):229–238.
- [43] Hurtado, D., Stainier, L., and Ortiz, M. (2014). The special-linear update: An application of differential manifold theory to the update of isochoric plasticity flow rules. *International Journal for Numerical Methods in Engineering*, 97(4):298–312.
- [44] Ishida, A. and Sato, M. (2003). Thickness effect on shape memory behavior of Ti-50.0at.%Ni thin film. *Acta Materialia*, 51(18):5571–5578.
- [45] Jarrige, I., Holliger, P., and Jonnard, P. (2004). Diffusion processes in NiTi/Si, NiTi/SiO₂ and NiTi/Si₃N₄ systems under annealing. *Thin Solid Films*, 458(1):314–321.
- [46] Kahle, O., Wielsch, U., Metzner, H., Bauer, J., Uhlig, C., and Zawatzki, C. (1998). Glass transition temperature and thermal expansion behaviour of polymer films investigated by variable temperature spectroscopic ellipsometry. *Thin Solid Films*, 313-314:803–807.

- [47] Karaca, H., Saghaian, S., Ded, G., Tobe, H., Basaran, B., Maier, H., Noebe, R., and Chumlyakov, Y. (2013). Effects of nanoprecipitation on the shape memory and material properties of an Ni-rich NiTiHf high temperature shape memory alloy. *Acta Materialia*, 61(19):7422–7431.
- [48] Karaca, H. E., Acar, E., Tobe, H., and Saghaian, S. M. (2014). NiTiHf-based shape memory alloys. *Materials Science and Technology*, 30(13):1530–1544.
- [49] Karakoc, O., Hayrettin, C., Bass, M., Wang, S., Canadinc, D., Mabe, J., Lagoudas, D., and Karaman, I. (2017). Effects of upper cycle temperature on the actuation fatigue response of NiTiHf high temperature shape memory alloys. *Acta Materialia*, 138:185–197.
- [50] Karakoc, O., Hayrettin, C., Evirgen, A., Santamarta, R., Canadinc, D., Wheeler, R., Wang, S., Lagoudas, D., and Karaman, I. (2019). Role of microstructure on the actuation fatigue performance of Ni-Rich NiTiHf high temperature shape memory alloys. *Acta Materialia*, 175:107–120.
- [51] Knick, C. R., Smith, G. L., Morris, C. J., and Bruck, H. A. (2019). Rapid and low power laser actuation of sputter-deposited NiTi shape memory alloy (SMA) MEMS thermal bimorph actuators. *Sensors and Actuators A: Physical*, 291:48–57.
- [52] Kockar, B., Karaman, I., Kim, J., and Chumlyakov, Y. (2006). A method to enhance cyclic reversibility of NiTiHf high temperature shape memory alloys. *Scripta Materialia*, 54(12):2203–2208.
- [53] Kohl, M. (2004). *Shape memory microactuators*. Springer Science & Business Media.
- [54] Kohl, M., Krevet, B., and Just, E. (2002). SMA microgripper system. *Sensors and Actuators A: Physical*, 97-98:646–652. Selected papers from Eurosenors XV.
- [55] König, D., Ehmann, M., Thienhaus, S., and Ludwig, A. (2010). Micro- to nanostructured devices for the characterization of scaling effects in shape-memory thin films. *Journal of Microelectromechanical Systems*, 19(5):1264–1269.
- [56] König, D., Zarnetta, R., Savan, A., Brunken, H., and Ludwig, A. (2011). Phase transformation, structural and functional fatigue properties of Ti-Ni-Hf shape memory thin films. *Acta Materialia*, 59(8):3267–3275.
- [57] Kröner, E. (1959). Allgemeine Kontinuumstheorie der Versetzungen und Eigenspannungen. *Archive for Rational Mechanics and Analysis*, 4(1):273–334.
- [58] Lagoudas, D. C., Entchev, P. B., Popov, P., Patoor, E., Brinson, L. C., and Gao, X. (2006). Shape memory alloys, part II: Modeling of polycrystals. *Mechanics of Materials*, 38(5):430–462. Shape Memory Alloys.

- [59] Lambrecht, F., Lay, C., Aseguinolaza, I. R., Chernenko, V., and Kohl, M. (2016). NiMnGa/Si shape memory bimorph nanoactuation. *Shape Memory and Superelasticity*, 2(4):347–359.
- [60] Lan, T. and Torkelson, J. M. (2014). Methacrylate-based polymer films useful in lithographic applications exhibit different glass transition temperature-confinement effects at high and low molecular weight. *Polymer*, 55(5):1249–1258.
- [61] Lee, E. H. (1969). Elastic-plastic deformation at finite strains. *Journal of Applied Mechanics*, 36(1):1–6.
- [62] Lee, H.-T., Kim, M.-S., Lee, G.-Y., Kim, C.-S., and Ahn, S.-H. (2018). Shape memory alloy (SMA)-based microscale actuators with 60% deformation rate and 1.6 kHz actuation speed. *Small*, 14(23):1801023.
- [63] Lega, P., Nedospasov, I., Orlov, A., Koledov, V., and Tabachkova, N. (2019). On the fundamental limits of the size of the shape memory nanoactuators posed by martensitic transition in Ti_2NiCu shape memory alloy on nano-scale. In *2019 IEEE International Conference on Manipulation, Manufacturing and Measurement on the Nanoscale (3M-NANO)*, pages 90–93.
- [64] Lester, B. T., Baxevanis, T., Chemisky, Y., and Lagoudas, D. C. (2015). Review and perspectives: Shape memory alloy composite systems. *Acta Mechanica*, 226(12):3907–3960.
- [65] LExcellent, C., Boubakar, M., Bouvet, C., and Calloch, S. (2006). About modelling the shape memory alloy behaviour based on the phase transformation surface identification under proportional loading and anisothermal conditions. *International Journal of Solids and Structures*, 43(3):613–626.
- [66] Lima de Miranda, R., Zamponi, C., and Quandt, E. (2013). Micropatterned freestanding superelastic TiNi films. *Advanced engineering materials*, 15(1-2):66–69.
- [67] Lion, A. (2000). Constitutive modelling in finite thermoviscoplasticity: a physical approach based on nonlinear rheological models. *International Journal of Plasticity*, 16(5):469–494.
- [68] Lührs, G., Hartmann, S., and Haupt, P. (1997). On the numerical treatment of finite deformations in elastoviscoplasticity. *Computer Methods in Applied Mechanics and Engineering*, 144(1):1–21.
- [69] Mandepudi, S. K. and Ackler, H. D. (2010). Processing and characterization of composite shape memory alloy (SMA) thin film structures for microactuators. In *Behavior and Mechanics of Multifunctional Materials and Composites 2010*, volume 7644, pages 154 – 165. International Society for Optics and Photonics, SPIE.

- [70] Meng, X., Cai, W., Chen, F., and Zhao, L. (2006). Effect of aging on martensitic transformation and microstructure in Ni-rich TiNiHf shape memory alloy. *Scripta Materialia*, 54(9):1599–1604.
- [71] Meng, X., Cai, W., Fu, Y., Li, Q., Zhang, J., and Zhao, L. (2008). Shape-memory behaviors in an aged Ni-rich TiNiHf high temperature shape-memory alloy. *Intermetallics*, 16(5):698–705.
- [72] Meng, X., Cai, W., Fu, Y., Zhang, J., and Zhao, L. (2010). Martensite structure in Ti-Ni-Hf-Cu quaternary alloy ribbons containing (Ti,Hf)₂Ni precipitates. *Acta Materialia*, 58(10):3751–3763.
- [73] Meng, X., Cai, W., Wang, L., Zheng, Y., Zhao, L., and Zhou, L. (2001). Microstructure of stress-induced martensite in a Ti-Ni-Hf high temperature shape memory alloy. *Scripta Materialia*, 45(10):1177–1182.
- [74] Meng, X., Cai, W., Zheng, Y., Tong, Y., Zhao, L., and Zhou, L. (2002). Stress-induced martensitic transformation behavior of a Ti-Ni-Hf high temperature shape memory alloy. *Materials Letters*, 55(1):111–115.
- [75] Meng, X., Zheng, Y., Cai, W., and Zhao, L. (2004). Two-way shape memory effect of a TiNiHf high temperature shape memory alloy. *Journal of Alloys and Compounds*, 372(1):180–186.
- [76] Miehe, C. (1996a). Exponential map algorithm for stress updates in anisotropic multiplicative elastoplasticity for single crystals. *International Journal for Numerical Methods in Engineering*, 39(19):3367–3390.
- [77] Miehe, C. (1996b). Multisurface thermoplasticity for single crystals at large strains in terms of eulerian vector updates. *International Journal of Solids and Structures*, 33(20):3103–3130.
- [78] Miehe, C. (2011). A multi-field incremental variational framework for gradient-extended standard dissipative solids. *Journal of the Mechanics and Physics of Solids*, 59(4):898–923.
- [79] Miehe, C., Stein, E., and Wagner, W. (1994). Associative multiplicative elasto-plasticity: Formulation and aspects of the numerical implementation including stability analysis. *Computers & Structures*, 52(5):969–978.
- [80] Mirzaeifar, R., DesRoches, R., Yavari, A., and Gall, K. (2013). A micromechanical analysis of the coupled thermomechanical superelastic response of textured and untextured polycrystalline NiTi shape memory alloys. *Acta Materialia*, 61(12):4542–4558.
- [81] Moler, C. and Van Loan, C. (2003). Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, 45(1):3–49.

- [82] Motemani, Y., McCluskey, P. J., Zhao, C., Tan, M. J., and Vlassak, J. J. (2011). Analysis of Ti-Ni-Hf shape memory alloys by combinatorial nanocalorimetry. *Acta Materialia*, 59(20):7602–7614.
- [83] Moumni, Z., Zaki, W., and Nguyen, Q. S. (2008). Theoretical and numerical modeling of solid-solid phase change: Application to the description of the thermomechanical behavior of shape memory alloys. *International Journal of Plasticity*, 24(4):614–645.
- [84] Müller, C. and Bruhns, O. (2006). A thermodynamic finite-strain model for pseudoelastic shape memory alloys. *International Journal of Plasticity*, 22(9):1658–1682.
- [85] Nagtegaal, J., Parks, D., and Rice, J. (1974). On numerically accurate finite element solutions in the fully plastic range. *Computer Methods in Applied Mechanics and Engineering*, 4(2):153–177.
- [86] Nemat-Nasser, S. (1982). On finite deformation elasto-plasticity. *International Journal of Solids and Structures*, 18(10):857–872.
- [87] Noebe, R., Biles, T., and Padula, S. (2006). NiTi-based high-temperature shape-memory alloys: properties, prospects, and potential applications. *MATERIALS ENGINEERING-NEW YORK*-, 32.
- [88] Ollier, E. (2002). Optical MEMS devices based on moving waveguides. *IEEE Journal of Selected Topics in Quantum Electronics*, 8(1):155–162.
- [89] Ortiz, M., Radovitzky, R. A., and Repetto, E. A. (2001). The computation of the exponential and logarithmic mappings and their first and second linearizations. *International Journal for Numerical Methods in Engineering*, 52(12):1431–1441.
- [90] Ortiz, M. and Stainier, L. (1999). The variational formulation of viscoplastic constitutive updates. *Computer Methods in Applied Mechanics and Engineering*, 171(3):419–444.
- [91] Oshida, Y. and Tominaga, T. (2020). *Nickel-Titanium Materials: Biomedical Applications*. Walter de Gruyter GmbH & Co KG.
- [92] Otsuka, K. and Ren, X. (2005). Physical metallurgy of Ti-Ni-based shape memory alloys. *Progress in Materials Science*, 50(5):511–678.
- [93] Otsuka, K. and Wayman, C. M. (1999). *Shape memory materials*. Cambridge university press.
- [94] Panico, M. and Brinson, L. (2007). A three-dimensional phenomenological model for martensite reorientation in shape memory alloys. *Journal of the Mechanics and Physics of Solids*, 55(11):2491–2511.

- [95] Patoor, E., Lagoudas, D. C., Entchev, P. B., Brinson, L. C., and Gao, X. (2006). Shape memory alloys, part I: General properties and modeling of single crystals. *Mechanics of Materials*, 38(5):391–429. Shape Memory Alloys.
- [96] Perić, D., Owen, D., and Honnor, M. (1992). A model for finite strain elasto-plasticity based on logarithmic strains: Computational issues. *Computer Methods in Applied Mechanics and Engineering*, 94(1):35–61.
- [97] Potapov, P., Shelyakov, A., Gulyaev, A., Svistunov, E., Matveeva, N., and Hodgson, D. (1997). Effect of Hf on the structure of Ni-Ti martensitic alloys. *Materials Letters*, 32(4):247–250.
- [98] Qidwai, M. and Lagoudas, D. (2000). On thermomechanics and transformation surfaces of polycrystalline NiTi shape memory alloy material. *International Journal of Plasticity*, 16(10):1309–1343.
- [99] Qiu, S., Krishnan, V. B., Padula, S. A., Noebe, R. D., Brown, D. W., Clausen, B., and Vaidyanathan, R. (2009). Measurement of the lattice plane strain and phase fraction evolution during heating and cooling in shape memory NiTi. *Applied Physics Letters*, 95(14):141906.
- [100] Rastjoo, S., Fechner, R., Bumke, L., Kötz, M., Quandt, E., and Kohl, M. (2020). Development and co-integration of a SMA/Si bimorph nanoactuator for Si photonic circuits. *Microelectronic Engineering*, 225:111257.
- [101] Reese, S. and Christ, D. (2008). Finite deformation pseudo-elasticity of shape memory alloys - constitutive modelling and finite element implementation. *International Journal of Plasticity*, 24(3):455–482.
- [102] Reese, S. and Govindjee, S. (1997). Theoretical and numerical aspects in the thermo-viscoelastic material behaviour of rubber-like polymers. *Mechanics of Time-Dependent Materials*, 1(4):357–396.
- [103] Rezaee-Hajidehi, M., Tuma, K., and Stupkiewicz, S. (2020). Gradient-enhanced thermomechanical 3D model for simulation of transformation patterns in pseudoelastic shape memory alloys. *International Journal of Plasticity*, 128:102589.
- [104] Rockafellar, R. T. (2015). *Convex Analysis*. Princeton University Press.
- [105] Rondelli, G. (1996). Corrosion resistance tests on NiTi shape memory alloy. *Biomaterials*, 17(20):2003–2008.
- [106] Saghaian, S., Karaca, H., Tobe, H., Turabi, A., Saedi, S., Saghaian, S., Chumlyakov, Y., and Noebe, R. (2017). High strength NiTiHf shape memory alloys with tailorable properties. *Acta Materialia*, 134:211–220.

- [107] Sanjabi, S., Cao, Y., and Barber, Z. (2005). Multi-target sputter deposition of $\text{Ni}_{50}\text{Ti}_{50-x}\text{Hf}_x$ shape memory thin films for high temperature microactuator application. *Sensors and Actuators A: Physical*, 121(2):543–548.
- [108] Scalet, G., Karakalas, A., Xu, L., and Lagoudas, D. (2021). Finite strain constitutive modelling of shape memory alloys considering partial phase transformation with transformation-induced plasticity. *Shape Memory and Superelasticity*, 7(2):206–221.
- [109] Sedlák, P., Frost, M., Benešová, B., Ben Zineb, T., and Šittner, P. (2012). Thermo-mechanical model for NiTi-based shape memory alloys including R-phase and material anisotropy under multi-axial loadings. *International Journal of Plasticity*, 39:132–151.
- [110] Seelecke, S. and Müller, I. (2004). Shape memory alloy actuators in smart structures: Modeling and simulation. *Applied Mechanics Reviews*, 57(1):23–46.
- [111] Sehitoglu, H., Wu, Y., Patriarca, L., Li, G., Ojha, A., Zhang, S., Chumlyakov, Y., and Nishida, M. (2017). Superelasticity and shape memory behavior of NiTiHf alloys. *Shape Memory and Superelasticity*, 3(2):168–187.
- [112] Shen, J., Zeng, Z., Nematollahi, M., Schell, N., Maawad, E., Vasin, R., Safaei, K., Poorganji, B., Elahinia, M., and Oliveira, J. (2021). In-situ synchrotron X-ray diffraction analysis of the elastic behaviour of martensite and H-phase in a NiTiHf high temperature shape memory alloy fabricated by laser powder bed fusion. *Additive Manufacturing Letters*, 1:100003.
- [113] Shuitcev, A., Vasin, R., Balagurov, A., Li, L., Bobrikov, I., and Tong, Y. (2020a). Thermal expansion of martensite in $\text{Ti}_{29.7}\text{Ni}_{50.3}\text{Hf}_{20}$ shape memory alloy. *Intermetallics*, 125:106889.
- [114] Shuitcev, A., Vasin, R., Fan, X., Balagurov, A., Bobrikov, I., Li, L., Golovin, I., and Tong, Y. (2020b). Volume effect upon martensitic transformation in $\text{Ti}_{29.7}\text{Ni}_{50.3}\text{Hf}_{20}$ high temperature shape memory alloy. *Scripta Materialia*, 178:67–70.
- [115] Shutov, A. (2016). Efficient implicit integration for finite-strain viscoplasticity with a nested multiplicative split. *Computer Methods in Applied Mechanics and Engineering*, 306:151–174.
- [116] Shutov, A. and Kreißig, R. (2008). Finite strain viscoplasticity with nonlinear kinematic hardening: Phenomenological modeling and time integration. *Computer Methods in Applied Mechanics and Engineering*, 197(21):2015–2029.
- [117] Shutov, A. and Kreißig, R. (2010). Geometric integrators for multiplicative viscoplasticity: Analysis of error accumulation. *Computer Methods in Applied Mechanics and Engineering*, 199(9):700–711.

- [118] Sielenkämper, M., Dittmann, J., and Wulfinghoff, S. (2022). Numerical strategies for variational updates in large strain inelasticity with incompressibility constraint. *International Journal for Numerical Methods in Engineering*, 123(1):245–267.
- [119] Sielenkämper, M. and Wulfinghoff, S. (2022). A thermomechanical finite strain shape memory alloy model and its application to bistable actuators. *Acta Mechanica*, 233(8):3059–3094.
- [120] Simo, J. (1992). Algorithms for static and dynamic multiplicative plasticity that preserve the classical return mapping schemes of the infinitesimal theory. *Computer Methods in Applied Mechanics and Engineering*, 99(1):61–112.
- [121] Simo, J., Taylor, R., and Pister, K. (1985). Variational and projection methods for the volume constraint in finite deformation elasto-plasticity. *Computer Methods in Applied Mechanics and Engineering*, 51(1):177–208.
- [122] Šittner, P., Landa, M., Lukáš, P., and Novák, V. (2006). R-phase transformation phenomena in thermomechanically loaded NiTi polycrystals. *Mechanics of Materials*, 38(5):475–492. Shape Memory Alloys.
- [123] Sterzl, T., Winzek, B., Rumpf, H., and Quandt, E. (2002). Bistable shape memory composites for switches, grippers and adjustable capacitors. In *Proceedings of the 8th International Conference on New Actuators, Actuator*, pages 91–94.
- [124] Stupkiewicz, S., Rezaee-Hajidehi, M., and Petryk, H. (2021). Multiscale analysis of the effect of interfacial energy on non-monotonic stress-strain response in shape memory alloys. *International Journal of Solids and Structures*, 221:77–91. Special Issue dedicated to papers from the International Union of Theoretical and Applied Mechanics 2019 Symposium on Phase Transformations in Shape Memory Materials: Modeling and Applications.
- [125] Tabesh, M., Lester, B., Hartl, D., and Lagoudas, D. (2012). Influence of the latent heat of transformation and thermomechanical coupling on the performance of shape memory alloy actuators. volume 2: Mechanics and Behavior of Active Materials; Integrated System Design and Implementation; Bio-Inspired Materials and Systems; Energy Harvesting of *Smart Materials, Adaptive Structures and Intelligent Systems*, pages 237–248.
- [126] Taya, M., Liang, Y., Namli, O. C., Tamagawa, H., and Howie, T. (2013). Design of two-way reversible bending actuator based on a shape memory alloy/shape memory polymer composite. *Smart Materials and Structures*, 22(10):105003.
- [127] Taylor, R. L. (2017). FEAP - finite element analysis program. <http://projects.ce.berkeley.edu/feap/>.

- [128] Tong, J., Zhan, Z.-L., and Vermeulen, B. (2004). Modelling of cyclic plasticity and viscoplasticity of a nickel-based alloy using Chaboche constitutive equations. *International journal of Fatigue*, 26(8):829–837.
- [129] Tong, Y., Shuitcev, A., and Zheng, Y. (2020). Recent development of TiNi-based shape memory alloys with high cycle stability and high transformation temperature. *Advanced Engineering Materials*, 22(4):1900496.
- [130] Turner, T. L. (2001). *Thermomechanical response of shape memory alloy hybrid composites*. DIANE Publishing.
- [131] Uchil, J., Fernandes, F. B., and Mahesh, K. (2007). X-ray diffraction study of the phase transformations in NiTi shape memory alloy. *Materials Characterization*, 58(3):243–248.
- [132] Vitushinsky, R., Schmitz, S., and Ludwig, A. (2009). Bistable thin-film shape memory actuators for applications in tactile displays. *Journal of Microelectromechanical Systems*, 18(1):186–194.
- [133] Vladimirov, I. N., Pietryga, M. P., and Reese, S. (2008). On the modelling of non-linear kinematic hardening at finite strains with application to springback - comparison of time integration algorithms. *International Journal for Numerical Methods in Engineering*, 75(1):1–28.
- [134] Waitz, T., Kazykhanov, V., and Karnthaler, H. (2004). Martensitic phase transformations in nanocrystalline NiTi studied by TEM. *Acta Materialia*, 52(1):137–147.
- [135] Wang, B. and Zhu, S. (2018). Cyclic tension-compression behavior of superelastic shape memory alloy bars with buckling-restrained devices. *Construction and Building Materials*, 186:103–113.
- [136] Wang, J., Moumni, Z., and Zhang, W. (2017a). A thermomechanically coupled finite-strain constitutive model for cyclic pseudoelasticity of polycrystalline shape memory alloys. *International Journal of Plasticity*, 97:194–221.
- [137] Wang, J., Moumni, Z., Zhang, W., Xu, Y., and Zaki, W. (2017b). A 3D finite-strain-based constitutive model for shape memory alloys accounting for thermomechanical coupling and martensite reorientation. *Smart Materials and Structures*, 26(6):065006.
- [138] Watanabe, H., Yamada, N., and Okaji, M. (2004). Linear thermal expansion coefficient of silicon from 293 to 1000 K. *International Journal of Thermophysics*, 25(1):221–236.
- [139] Weber, G. and Anand, L. (1990). Finite deformation constitutive equations and a time integration procedure for isotropic, hyperelastic-viscoplastic solids. *Computer Methods in Applied Mechanics and Engineering*, 79(2):173–202.

- [140] Winzek, B., Schmitz, S., Rumpf, H., Sterzl, T., Hassdorf, R., Thienhaus, S., Feydt, J., Moske, M., and Quandt, E. (2004a). Recent developments in shape memory thin film technology. *Materials Science and Engineering: A*, 378(1):40–46. European Symposium on Martensitic Transformation and Shape-Memory.
- [141] Winzek, B., Schmitz, S., and Vitushinsky, R. (2004b). Shape memory actuators in mobile robots for planetary surface exploration. In *Tools and Technologies for Future Planetary Exploration*, volume 543, pages 115–120.
- [142] Winzek, B., Sterzl, T., and Quandt, E. (2001). Bistable thin film composites with TiHfNi-shape memory alloys. In *Transducers '01 Eurosensors XV*, pages 706–709, Berlin, Heidelberg. Springer Berlin Heidelberg.
- [143] Winzek, B., Sterzl, T., Rumpf, H., Botkin, N., and Quandt, E. (2002). Thin film shape memory composites. In *NatoConference on Martensitic Phase Transformation, Metz, Frankreich., Metz*.
- [144] Winzek, B., Sterzl, T., Rumpf, H., and Quandt, E. (2003). Composites of different shape memory alloys and polymers for complex actuator motions. In *Journal de Physique IV (Proceedings)*, volume 112, pages 1163–1168. EDP sciences.
- [145] Woodworth, L. A., Lohse, F., Kopelmann, K., Cherif, C., and Kaliske, M. (2022). Development of a constitutive model considering functional fatigue and pre-stretch in shape memory alloy wires. *International Journal of Solids and Structures*, 234-235:111242.
- [146] Wriggers, P. (2008). *Nonlinear finite element methods*. Springer Science & Business Media.
- [147] Wulfinghoff, S. and Böhlke, T. (2013). Equivalent plastic strain gradient crystal plasticity – enhanced power law subroutine. *GAMM-Mitteilungen*, 36(2):134–148.
- [148] Xu, L., Solomou, A., Baxevanis, T., and Lagoudas, D. (2021). Finite strain constitutive modeling for shape memory alloys considering transformation-induced plasticity and two-way shape memory effect. *International Journal of Solids and Structures*, 221:42–59. Special Issue dedicated to papers from the International Union of Theoretical and Applied Mechanics 2019 Symposium on Phase Transformations in Shape Memory Materials: Modeling and Applications.
- [149] Yang, Q., Stainier, L., and Ortiz, M. (2006). A variational formulation of the coupled thermo-mechanical boundary-value problem for general dissipative solids. *Journal of the Mechanics and Physics of Solids*, 54(2):401–424.
- [150] Yi, X., Meng, X., Cai, W., and Zhao, L. (2019). Multi-stage martensitic transformation behaviors and microstructural characteristics of Ti-Ni-Hf high temperature shape memory alloy powders. *Journal of Alloys and Compounds*, 781:644–656.

-
- [151] Yoshida, F. and Uemori, T. (2003). A model of large-strain cyclic plasticity and its application to springback simulation. *International Journal of Mechanical Sciences*, 45(10):1687–1702.
- [152] Yu, C., Kang, G., Sun, Q., and Fang, D. (2019). Modeling the martensite reorientation and resulting zero/negative thermal expansion of shape memory alloys. *Journal of the Mechanics and Physics of Solids*, 127:295–331.
- [153] Zaki, W., Moumni, Z., and Morin, C. (2011). Modeling tensile-compressive asymmetry for superelastic shape memory alloys. *Mechanics of Advanced Materials and Structures*, 18(7):559–564.
- [154] Zamkovskaya, A., Maksimova, E., Nauhatsky, I., and Shapoval, M. (2017). X-ray diffraction investigations of the thermal expansion of iron borate FeBO_3 crystals. *Journal of Physics: Conference Series*, 929:012030.
- [155] Zarinejad, M., Liu, Y., and White, T. J. (2008). The crystal chemistry of martensite in NiTiHf shape memory alloys. *Intermetallics*, 16(7):876–883.
- [156] Zhang, M. and Baxevanis, T. (2021). An extended three-dimensional finite strain constitutive model for shape memory alloys. *Journal of Applied Mechanics*, 88(11). 111010.
- [157] Zhang, Y., Kang, G., Miao, H., and Yu, C. (2022). Cyclic degeneration of elastocaloric effect for NiTi shape memory alloy: Experimental observation and constitutive model. *International Journal of Solids and Structures*, 248:111638.
- [158] Zhang, Y., Li, M., Wang, Y. D., Lin, J. P., Dahmen, K. A., Wang, Z. L., and Liaw, P. K. (2014). Superelasticity and serration behavior in small-sized NiMnGa alloys. *Advanced Engineering Materials*, 16(8):955–960.