Efficient and Globally Convergent Minimization Algorithms for Small- and Finite-Strain Plasticity Problems

M. Sc. Patrick Jaap

Dissertation

zur Erlangung des akademischen Grades

Doctor rerum naturalium (Dr. rer. nat.)

vorgelegt dem Bereich Mathematik und Naturwissenschaften
der Technischen Universität Dresden am 23. Februar 2023

verteidigt am 21. August 2023

1. Gutachter:
Prof. Dr. Oliver Sander

2. Gutachter:
Prof. Dr. Christian Meyer

Für meinen Opa Ulrich,
der meine Freude an Zahlen geweckt hat

Versicherung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Datum, Unterschrift

Benutzte Hilfmittel

Für die Ausarbeitung der vorliegenden Arbeit wurden technische Hilfsmittel verwendet. Diese sind im Detail im Folgenden aufgelistet:

- Wörterbücher: `https://www.dict.cc/`, `https://www.deepl.com/translator`
- Korrekturlesen: `https://www.deepl.com/write`
- LaTeX-Editor: TeXstudio (`https://www.texstudio.org/`)
- `C++`-Programmierung: KDevelop (`https://www.kdevelop.org/`)

Contents

I. An Efficient and Globally Convergent Minimization Algorithm for Small-Strain Plasticity Minimization Problems 11

1. Introduction 13

2. Strong Formulation of Small-Strain Plasticity 15
2.1. Kinematics 15
2.2. Forces and Stress 18
2.3. Elastic and Plastic Strain 19
2.4. Yield Stress 21
2.4.1. Von Mises Yield Criterion 22
2.4.2. Tresca Yield Criterion 22
2.5. Plastic Strain Evolution 23
2.6. Hardening 25
2.6.1. Kinematic Hardening 25
2.6.2. Isotropic Hardening 26
2.7. Generalized Formulation 28

3. Variational Formulations of Small-Strain Plasticity 31
3.1. Excursus: Dual Formulation of Plasticity 31
3.2. Support Functions and Dissipation 33
3.3. Variational Formulation of Primal Plasticity 37

4. Discretization 41
4.1. Time Discretization 41
4.2. Eliminating Internal Variables 44
4.3. Space Discretization 45
4.4. Excursus: Primal Plasticity with Gradient Regularization 48

5. Minimization Algorithms for Small-Strain Primal Plasticity 51
5.1. Predictor–Corrector Methods 51
5.2. The TNNMG Algorithm 52
5.2.1. TNNMG Stages 53
5.2.2. Convergence of TNNMG 55
5.3. Application to Primal Small-Strain Plasticity Problems 56
5.3.1. Minimizing in Displacement Subspaces 56
5.3.2. Minimizing in Plastic Strain Subspaces 56

5.3.3. Determining Maximal Differentiable Subspaces 63
5.4. Efficient Implementation of Norm Derivatives 64

6. Numerical Results 67
6.1. Von Mises Dissipation with Kinematic Hardening 67
6.1.1. Two-Dimensional Test 68
6.1.2. Three-Dimensional Test 71
6.2. Three-Dimensional Tresca Dissipation 73
6.3. Isotropic Hardening 76
6.4. Conclusion 77

II. An Efficient and Globally Convergent Minimization Algorithm for Finite-Strain Plasticity Problems 79

7. Introduction to Finite-Strain Mechanics 81
7.1. Kinematics 81
7.1.1. Deformation 81
7.1.2. Strain Tensors 81
7.1.3. Decomposition of the Deformation Gradient 82
7.2. Ranges 83
7.2.1. Matrix Manifolds 83
7.2.2. Plastic Spin 84
7.3. Stress Tensors 85
7.4. Material Laws 86
7.4.1. Hyperelastic Materials 86
7.4.2. Polyconvex Materials 88

8. Energetic Formulation of Rate-Independent Finite-Strain Plasticity 91
8.1. Plastic Energy and Back Stress 91
8.2. Kinematic Hardening 92
8.3. Plastic Flow 94
8.4. Global Formulation 95
8.5. Von Mises Dissipation 99
8.6. Existence Theory 100

9. Discrete Problem 103
9.1. Time Discretization 103
9.2. Reformulation of the Minimization Problem 103
9.2.1. Without Plastic Spin 104
9.2.2. With Plastic Spin 107
9.3. Space Discretization 110
9.3.1. Discrete Function Spaces 110
9.3.2. Discrete Plastic Strain Gradients 113

9.3.3. Discrete Minimization Functional 115

10. The Inexact Proximal Newton Method 117
10.1. General Setting 117
10.2. Proximal Newton Steps 118
10.2.1. Second-Order Model Problems 118
10.2.2. Sufficient Decrease Criterion 119
10.3. Choice of Regularization Weights 119
10.3.1. A Simple Strategy 120
10.3.2. Adaptive Strategies 120
10.4. Inexactness Criteria 121
10.4.1. Local Criterion 121
10.4.2. Global Criterion 121
10.5. Convergence Properties 122
10.6. Application to Finite-Strain Plasticity 124
10.6.1. Solving Second-Order Subproblems with TNNMG 125
10.6.2. Local Problems 125
10.6.3. Automatic Differentiation 129
10.6.4. Evaluating Inexactness Criteria 129

11. Numerical Results 131
11.1. Test Parameters 132
11.1.1. Elastic Energies 132
11.1.2. Kinematic Hardening 132
11.1.3. Other Parameters 133
11.1.4. Solver Settings 133
11.1.5. Test Machine 134
11.2. Tests on a Simple Geometry 134
11.2.1. Regularization Strategies 134
11.2.2. Gradient Regularization 137
11.2.3. Inexact Proximal Newton 140
11.3. Different Material Models on a Complex Geometry 143
11.4. Comparison to other Methods 147
11.4.1. Proximal Gradient Method 148
11.4.2. Subgradient Method 149
11.5. Conclusion 150

A. Appendix 151
A.1. The Dissipation Function for the Tresca Yield Criterion 151
A.2. Geometric Multigrid Method 152
A.2.1. Gauss–Seidel Steps 153
A.2.2. Prolongation and Restriction Operators 153
A.2.3. Multigrid Steps 154

A.3. The Matrix Group SL(d) 155
  A.3.1. Geometry 155
  A.3.2. Positive Definite, Symmetric Subset SL(d)$^+_{\mathrm{sym}}$ 156
  A.3.3. Projection onto SL(d) 157
A.4. A Property of Symmetric Matrices 157
A.5. Robust and Efficient Computation of Matrix Exponential 158
A.6. Convergence of the Newton Method with Armijo Damping 158

B. C++ Source Code **161**
B.1. Docker File 161
B.2. Small-Strain Tests 162
B.3. Finite-Strain Tests 163

Notation

In this thesis we try to stick to the following notation policy.

- The domains of interest, as subsets of $\mathbb{R}^d$, are denoted by Ω with possible super- or subscripts.
- Boundary portions of Ω are denoted by Γ with possible super- or subscripts.
- x is usually an element of Ω.
- t is always a “time” parameter.
- Scalar values use Greek letters α, γ, λ.
- Vectors in $\mathbb{R}^d$ use lowercase Latin letters u, v.
- Matrices are written in uppercase Latin letters A, P, Q.
- Standard continuum mechanics quantities, such as σ and ϵ, do not follow the rules above.
- Local quantities use standard fonts, while global functions (on Ω) are boldfaced:

  $$p = \mathbf{p}(x) \quad \text{for} \quad x \in \Omega.$$

- Dual pairings are given is sharp brackets $\langle \cdot, \cdot \rangle$.
- The scalar product of two vectors is written by a single dot $a \cdot b$ for $a, b \in \mathbb{R}^d$.
- Matrix scalar products are written using two dots $A : B$ for $A, B \in \mathbb{R}^{n \times m}$.
- Multiplications of tensors $\mathbf{C}$ and matrices M are written as $\mathbf{C}M$.
- The identification operator $\equiv$ is used whenever two functions are equal in the pointwise meaning, i.e., for $\mathbf{f}, \mathbf{g} : D \to R$

  $$\mathbf{f} \equiv \mathbf{g} \quad \Leftrightarrow \quad \mathbf{f}(x) = \mathbf{g}(x)$$

  for all (or almost all, depending on the function space) $x \in D$.

Why do we Differentiate between Small- and Finite-Strain Plasticity?

This thesis covers two fundamental models of continuum mechanics: small-strain plasticity and finite-strain plasticity. Although the two models have a similar name, the mathematical theory involved is very different. However, we are able to show that the finite-strain plasticity model is an extension of the small-strain model.

In the first part, we cover the topic of small-strain plasticity. After an introduction to the continuum mechanics model, a numerical approximation scheme leads to convex but nonsmooth minimization problems. The assumption of small strains allows us to justify a number of linearizations of the continuum mechanics model. Hence, the minimization problems have a simple form, but their solution is non-trivial due to the nonsmoothness. Therefore, an efficient solver is presented to handle these problems.

The second part of this thesis deals with the theory of finite-strain plasticity. No linearizations are applied to the continuum mechanics model. Therefore, the mathematical theory is more involved and we have to consider nonlinear manifold spaces to describe the problem and the solutions. A numerical approximation scheme again leads to nonsmooth minimization problems. However, these problems are not convex. Hence, another method has been developed to handle these minimization problems numerically. An interesting property of the solver is that it constructs a series of small-strain plasticity problems. These small-strain plasticity problems can be solved efficiently by methods from the first part of this thesis. Therefore, we consider the second part of this thesis as an extension of the first part, both in terms of the continuum mechanics model and the solution algorithms.

Part I.

An Efficient and Globally Convergent Minimization Algorithm for Small-Strain Plasticity Minimization Problems

1. Introduction

Small strain plasticity problems are of great interest in engineering. Over the decades, much progress has been made in modeling the physical problem and in improving the solution algorithms. Small-strain plasticity is one of the most studied areas. The fundamental assumption of small strains allows a variety of convenient simplifications of the physical interactions of the quantities involved.

The content of this part of the thesis is based on the collaboration with Oliver Sander, resulting in the published article [30]. In this thesis, we extend that article in several places by introducing another nonsmooth yield function in Section 2.4 and new hardening laws in Section 2.6. The general structure of this topic, including the article [30], is taken in large parts from the monograph by Weimin Han and B. Daya Reddy [16]. This book gives a comprehensive overview of possible further ideas of small-strain plasticity. We consider the parts of this book to the extent that we can establish a solid foundation for the problem of small-strain plasticity without at the same time going beyond the scope of this thesis. Several aspects from the extended theory could easily be adopted in this thesis, while other parts would require extensive changes.

This part of the thesis is structured as follows. The physical strong form of the small-strain plasticity model is constructed from scratch in Chapter 2. In Chapter 3, equivalent variational formulations of the strong formulation are presented. An algebraic form of the variational problems is presented in Chapter 4, followed by the introduction of solution algorithms in Chapter 5. Finally, in Chapter 6, several numerical tests are performed to demonstrate the efficiency of the algorithms.

2. Strong Formulation of Small-Strain Plasticity

We begin with an overview of important results from continuum mechanics. Most of the results are taken from [16, Part I], including the majority of the notation. Intended variations in the notation are made in order to fit consistently to the second part of this thesis about finite-strain mechanics. It should be mentioned that other commonly used notations exist in literature.

2.1. Kinematics

The physical problem is defined over an open and bounded domain $\Omega \subset \mathbb{R}^d$. We will also refer to Ω as the initial state. In this thesis we consider only $d \in \{2, 3\}$ since these are the relevant cases for continuum mechanics. The physical observation we want to explain here is the change of the shape of the domain undergoing external loads. This change in shape is modeled by a time-dependent deformation field $\mathbf{y} : \Omega \times [0, T] \to \mathbb{R}^d$, such that for a time point $t \in [0, T]$, the value of $\mathbf{y}(x, t)$ describes the current position of $x \in \Omega$. Hence, the current domain at time t is given by

$$\Omega^{\mathbf{y},t} := \mathbf{y}(\Omega, t).$$

Besides the deformation field, we will use the displacement field $\mathbf{u} : \Omega \times [0, T] \to \mathbb{R}^d$, defined by

$$\mathbf{u}(x, t) := \mathbf{y}(x, t) - x.$$

A representation of deformation and displacement is visualized in Figure 2.1.

Although the deformation and the displacement field describe the shape change well, they are no suitable quantity for measuring local deformations. This becomes obvious if we consider a *rigid body motion* given by the deformation field

$$\mathbf{y}(x, t) = Q(t)x + b(t) \tag{2.1}$$

with an orthogonal matrix $Q(t)$ and a vector $b(t)$ for all t. This describes for instance the movement of a pen on the desk, if the pen is given by Ω. For those rigid body motions, no local deformation takes place.

Therefore, we will construct another quantity (the *strain*) on Ω which will be invariant under rigid body motions. To this end, we consider a point $x \in \Omega$, and two directions $a, b \in \mathbb{R}^d$, such that $x + a$ and $x + b$ are in Ω. The goal is to capture both changes in the length scales and in the angle between a and b. The scalar product $\langle a, b\rangle$ captures both of these.

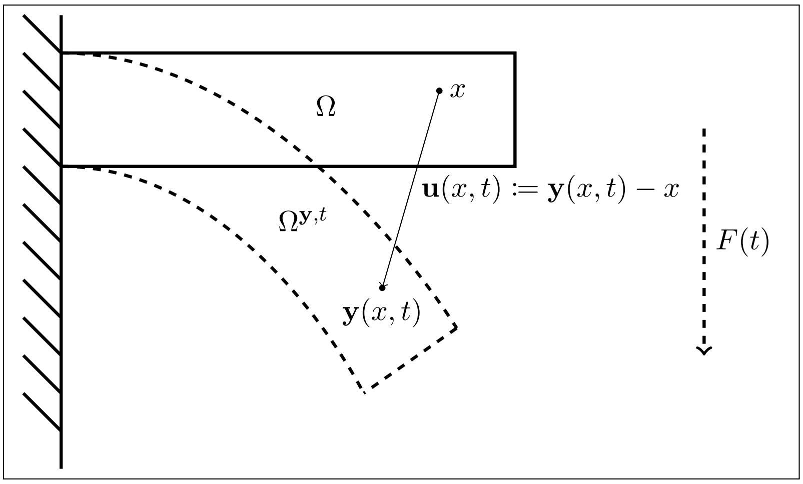

Ω
x
$\Omega^{\mathbf{y},t}$
$\mathbf{u}(x,t) := \mathbf{y}(x,t) - x$
$F(t)$
$\mathbf{y}(x,t)$

Figure 2.1.: Deformation of the two-dimensional initial state (solid lines) to the deformed state (dashed lines) at time t.

After the deformation at time t, the deformed directions are given by $\mathbf{y}(x+a,t)-\mathbf{y}(x,t)$ and $\mathbf{y}(x+b,t)-\mathbf{y}(x,t)$, respectively, in the deformed body. The situation is displayed in Figure 2.2.

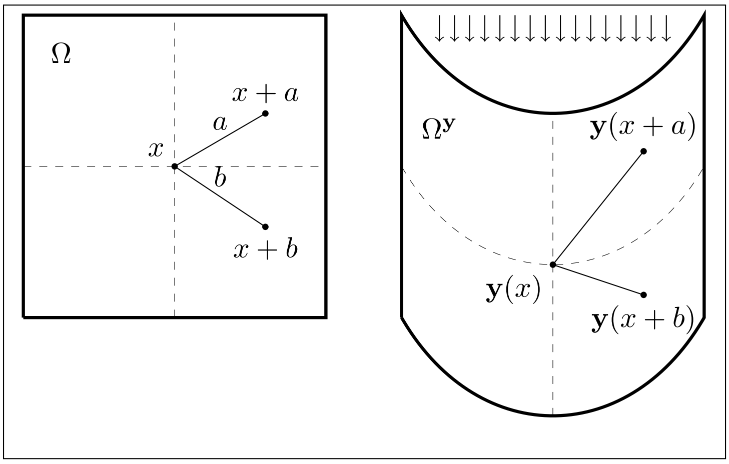

Ω
$x+a$
a
x
b
$x+b$
$\Omega^{\mathbf{y}}$
$\mathbf{y}(x+a)$
$\mathbf{y}(x)$
$\mathbf{y}(x+b)$

Figure 2.2.: Location of x and the directions a and b in the initial state (left), and in the deformed state (right). Dependency on t is omitted for readability.

The infinitesimal change of the scalar product of a and b is given by

$$S(x;a,b) := \lim_{h\to 0} \frac{\langle \mathbf{y}(x+ha,t)-\mathbf{y}(x,t), \mathbf{y}(x+hb,t)-\mathbf{y}(x,t)\rangle - \langle ha, hb\rangle}{h^2}. \tag{2.2}$$

If we assume that $\mathbf{y}$ is sufficiently smooth, we can write

$$\mathbf{y}(x+ha,t) = \mathbf{y}(x,t) + h\nabla\mathbf{y}(x,y)\,a + \mathcal{O}(h^2)$$

where ∇ is the gradient operator with respect to $x \in \Omega$. Plugging this into the scalar products leads to

$$\begin{aligned}\langle \mathbf{y}(x+ha,t)-\mathbf{y}(x,t), \mathbf{y}(x+hb,t)-\mathbf{y}(x,t)\rangle &= \langle h\nabla\mathbf{y}(x,y)\cdot a, h\nabla\mathbf{y}(x,y)\cdot b\rangle + \mathcal{O}(h^3)\\ &= h^2 a^T \left(\nabla\mathbf{y}(x,t)^T\nabla\mathbf{y}(x,t)\right) b + \mathcal{O}(h^3).\end{aligned}$$

Inserting this into (2.2) results in

$$S(x;a,b) = a^T\left(\nabla\mathbf{y}(x,t)^T\nabla\mathbf{y}(x,t) - I\right) b.$$

Now we use $\nabla\mathbf{u} = \nabla\mathbf{y} - \mathbf{I}$ to obtain

$$S(x;a,b) = a^T\left(\nabla\mathbf{u}(x,t) + \nabla\mathbf{u}(x,t)^T + \nabla\mathbf{u}(x,y)^T\nabla\mathbf{u}(x,t)\right) b.$$

Hence, the local change of the deformation can be expressed as a bilinear form independent of the directions a, b.

Definition 2.1.1 (Strain tensor). *For a differentiable displacement field* $\mathbf{u} : \Omega \times [0,T] \to \mathbb{R}^d$ *the strain tensor is given by*

$$\mathbf{E} = \mathbf{E}(\mathbf{u}) \coloneqq \frac{1}{2}\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T + \nabla\mathbf{u}^T\nabla\mathbf{u}\right).$$

Remarks 2.1.2.

(i) $\mathbf{E}$ *is called the St. Venant–Green strain tensor.*

(ii) Often we find the equivalent definition

$$\mathbf{E} = \mathbf{E}(\mathbf{y}) \coloneqq \frac{1}{2}\left(\nabla\mathbf{y}^T\nabla\mathbf{y} - \mathbf{I}\right)$$

if the deformation field $\mathbf{y}$ *is the quantity of interest.*

(iii) For a rigid body motion (2.1)*, we have* $\mathbf{E} = 0$*. Our introductory motivation for the strain of rigid body motions is therefore fulfilled.*

The connection of strain and displacement is obviously not linear, which makes both analysis and numerical handling of the problem complicated. This is studied in detail in the second part of this thesis. In the theory of small strains the following name-giving assumption is made: $\|\nabla\mathbf{u}\|$ is small in a given norm. This implies that $\|\mathbf{E}\|$ is also small in the same norm. We will never specify what *small* means exactly in practice since it is a model choice made beforehand. This choice is justified afterwards when numerical results match real observations. For a small $\nabla\mathbf{u}$, the quadratic term $\nabla\mathbf{u}^T\nabla\mathbf{u}$ can be neglected in practice.

Definition 2.1.3 (Small-Strain Tensor). *For a given differentiable displacement field* $\mathbf{u} : \Omega \times [0,T] \to \mathbb{R}^d$ *the small-strain tensor is*

$$\boldsymbol{\epsilon}(\mathbf{u}) \coloneqq \frac{1}{2}\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right).$$

2.2. Forces and Stress

Now we introduced quantities that describe the current form of the material undergoing the deformation process. Next, external load is studied. In continuum mechanics, load is described by forces. Let $\partial\Omega$ denote the boundary of Ω. The total force F that acts on Ω is given by the integral of the acceleration times the material mass density $\boldsymbol{\rho}$

$$F = \int_\Omega \boldsymbol{\rho}\ddot{\mathbf{u}}\,dx \in \mathbb{R}^d.$$

The total force F is split into volumetric and surface parts. Formally, there are force densities $\mathbf{b} : \Omega \times [0,T] \to \mathbb{R}^d$ and $\mathbf{s} : \partial\Omega \times [0,T] \to \mathbb{R}^d$ such that

$$F(t) = \int_\Omega \mathbf{b}(x,t)\,dx + \int_{\partial\Omega} \mathbf{s}(x,t)\,dS. \tag{2.3}$$

A crucial result from continuum mechanics states that there is a time-dependent, symmetric matrix field $\boldsymbol{\sigma}\colon \Omega \times [0,T] \to \mathbb{R}^{d\times d}$ which represents the force density s, cf. [16, Section 2.2]

Proposition 2.2.1 (Cauchy stress at small strains)**.** *Assuming that* $\|\nabla\mathbf{u}\|$ *is small there is a differentiable and symmetric matrix field* $\boldsymbol{\sigma} : \Omega \times [0,T] \to \mathbb{R}^{d\times d}$ *such that*

$$\mathbf{s}(x,t) = \boldsymbol{\sigma}(x,t) \cdot \mathbf{n}(x)$$

for the outer normal $\mathbf{n}(x)$ *of* $\partial\Omega$ *at* x*. Furthermore, we can express the total force* $F(t)$ *by*

$$F(t) = \int_\Omega \operatorname{div}(\boldsymbol{\sigma}(x,t)) + \mathbf{b}(x,t)\,dx.$$

The div*-operator is the row-wise divergence with respect to* x*.*

Remark 2.2.2. *Later on in the second part of this thesis in Section 7.3 a more detailed insight into the origin of the stress tensor field* $\boldsymbol{\sigma}$ *will be given. There it will also be stated why the assumption that* $\|\nabla\mathbf{u}\|$ *is small plays a crucial role in the proposition above.*

The results above can be formulated for each subset $\Omega' \subseteq \Omega$. Hence we conclude the strong pointwise equation

$$\operatorname{div}(\boldsymbol{\sigma}) + \mathbf{b} = \boldsymbol{\rho}\ddot{\mathbf{u}} \qquad \text{in } \Omega \times [0,T].$$

This is called the *balance of linear momentum.* Under the further assumption that the process of deformation is slow and independent of the rate of the external variables, the equation simplifies more. But first, let us formalize what rate-independence is precisely, cf. [22, Definition 1.2.1].

Definition 2.2.3 (Rate-independent process)**.** *Let* $[0,T]$ *a time interval, and* F *the external load. Consider a process depending on* F *with a solution* $q = q(t)$ *in the time interval* $[0,T]$*. We call the system* (F,q) *rate-independent, if for every* $\alpha > 0$ *the time-scaled control* $t \mapsto F(\alpha t)$ *obtains a time-scaled solution* $t \mapsto q(\alpha t)$*.*

In other words, the process is independent of time-scaling. An immediate consequence of rate-independent solutions q is that speed $\dot{q}$ and acceleration $\ddot{q}$ cannot be part of a rate-independent system. In this thesis we consider only rate-independent processes and remove the acceleration term from our system.

Definition 2.2.4 (Quasi-static balance of momentum)**.** *For a rate-independent deformation process we consider the balance of moment linear momentum*

$$-\operatorname{div}(\boldsymbol{\sigma}) = \mathbf{b} \qquad \text{in } \Omega \times [0, T]. \tag{2.4}$$

2.3. Elastic and Plastic Strain

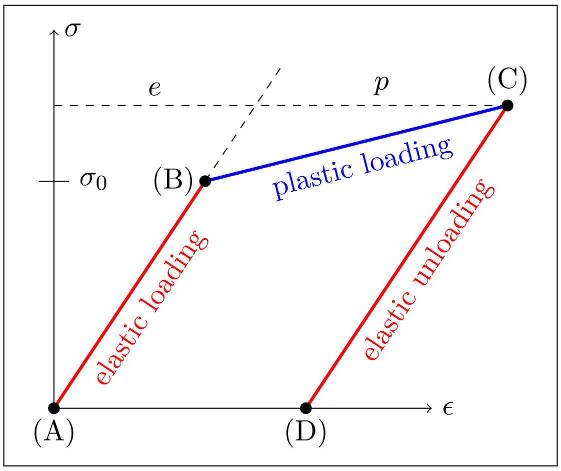


Figure 2.3.: Simplified stress–strain diagram for a loading–unloading test for $d = 1$ and plastic behavior.

In this section we want to connect the strain $\boldsymbol{\epsilon}$ to the stress $\boldsymbol{\sigma}$. A typical experiment to introduce this topic is given by a tensile test with a metal wire. We fix the wire on one end an apply an increasing pull load on the other end. After the maximal load is reached, the load is reduced back to zero. A simplified linear stress–strain curve is displayed in Figure 2.3.

What we can see are three different segments of the stress–strain curve. We start at an initial stress- and strain-free state (A). During the loading phase the increase of the stress reaches a critical stress value σ_0 at (B). Now the slope changes until we arrive at the maximal load, and hence the maximal stress at (C). During unloading, starting at (C), the stress decreases linearly until we arrive at (D). There is a remaining strain without any stress.

If the maximal stress is small enough that we never reach σ_0 at the first place, then the connection between σ and ϵ is linear. This is called elastic behavior and generates the elastic strain $\mathbf{e}$. After the critical stress σ_0 has been reached, an irreversible strain occurs. Irreversibility is understood in a way that further external load is necessary to return to the initial state. This kind of strain is called the plastic strain $\mathbf{p}$ and does not depend on the current load, but rather on the history of previous loads[1]. This phenomenon is visible in Figure 2.3. In the moment of decreasing the stress (C), there is already a portion p of the total strain ϵ created. This remains unchanged till the end (D).

In order to put this into a mathematical context, we recall the definition of the small-strain tensor

$$\boldsymbol{\epsilon} = \boldsymbol{\epsilon}(\mathbf{u}) := \frac{1}{2}\left(\nabla\mathbf{u} + (\nabla\mathbf{u})^T\right).$$

This needs to be decomposed into the elastic strain $\mathbf{e}$ and the plastic strain $\mathbf{p}$. How the decomposition is done in detail is not obvious at all. In our motivational experiment above, we use an additive split.

$$\boldsymbol{\epsilon} = \mathbf{e} + \mathbf{p}, \tag{2.5}$$

which will be the starting point for the small-strain theory. In [20] the product decomposition

$$\boldsymbol{\epsilon} = \mathbf{e}\mathbf{p}$$

was introduced. In fact, this will be used for the finite strain theory later in the second part of this thesis and discussed there in detail. The additive decomposition (2.5) allows us to work completely in a linear vector space setting, since we only rely on sums and differences in the strain space.

An important vector space is given by the space of symmetric matrices

$$\mathbb{S}^d := \{A \in \mathbb{R}^{d\times d},\ A^T = A\}.$$

We have

$$\boldsymbol{\epsilon}(x), \boldsymbol{\sigma}(x) \in \mathbb{S}^d \quad \forall x \in \Omega.$$

An important subspace of the above is given by the trace-tree, symmetric matrices

$$\mathbb{S}^d_0 := \{A \in \mathbb{S}^d,\ \operatorname{tr}(A) = 0\},$$

where the trace is the sum over the diagonal elements

$$\operatorname{tr}(A) := \sum_{i=1}^{d} A_{ii}.$$

We assume that the plastic strain $\mathbf{p}$ is volume-preserving [16, page 52]. This is realized by

$$\mathbf{p}(x) \in \mathbb{S}^d_0 \quad \forall x \in \Omega.$$

We assume that $\mathbf{p}(t)$ does not correspond to the current stress $\boldsymbol{\sigma}(t)$, since the current plastic strain depends only on the load history. Hence, the stress only depends on the elastic strain $\mathbf{e}$. In this part of this thesis we only consider a linear connection between the stress and the elastic strain. This is realized by a fourth order tensor $\mathbf{H}$ such that

---

[1] In [15] the current plastic strain is described as a model parameter for a fixed point in time and only the evolution $\dot{\mathbf{p}}$ is a quantity of interest.

$$\boldsymbol{\sigma} = \mathbf{H} : \mathbf{e}. \tag{2.6}$$

This is called Hooke's law. For the one dimensional case we see this law as the linear connection from (A) to (B) in Figure 2.3. In general, $\mathbf{H}$ contains d^4 components, but from demanding symmetry on both $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}$ some degrees of freedom are already fixed. If the material properties are independent of the coordinate system, we have symmetry of $\mathbf{H}$ in all components. In this case there are only two main components left: A bulk and a shear component. These are described by two scalar parameters λ and μ (called Lamé parameters), such that

$$\mathbf{H} : \mathbf{e} = \lambda \operatorname{tr}(\mathbf{e})\mathbf{I} + 2\mu\mathbf{e}. \tag{2.7}$$

2.4. Yield Stress

Hooke's law only holds for elastic processes. Therefore, we have to extend the elastic theory. We assume that the additional plastic strain $\mathbf{p}$ evolves only if a certain amount of stress is reached, namely the *yield stress*. The connection to the stress field $\boldsymbol{\sigma}$ is given by a function

$$\phi : \mathbb{S}^d \to \mathbb{R}$$

which can be interpreted as a measure of the stress values. We will refer to the function above as *yield function*. Once the yield function is established, we say a yield stress is reached if for $x \in \Omega$

$$\phi(\boldsymbol{\sigma}(x)) = 0.$$

For the construction of yield functions, we use that a yield stress is given at the value zero. We postulate that negative values indicate elastic behavior. We call the set of stresses that induce purely elastic material responses the *elastic region*.

Definition 2.4.1. *For a yield function $\phi : \mathbb{S}^d \to \mathbb{R}$ elastic region is*

$$\mathcal{E}(\phi) := \{\sigma \in \mathbb{S}^d : \ \phi(\sigma) \le 0\}.$$

Determining a suitable yield function ϕ requires engineering knowledge. We restrict ourself to convex yield functions that only depend on the singular values of $\boldsymbol{\sigma}$.

Two major classes of yield functions can be found in the literature. The first class is rather simple and creates elastic regions $\mathcal{E}$ with smooth boundary. On the other hand, the second class is more involved with a nonsmooth boundary of the elastic region. Of course, special numerical methods are necessary to deal with nonsmooth components.

Convexity of the yield function implies directly that the elastic region is also convex [8]. For both cases (smooth and nonsmooth) a classical candidate is considered in detail in the following.

2.4.1. Von Mises Yield Criterion

The von Mises yield criterion creates smooth elastic regions. It was developed such that hydrostatic stresses $\sigma = \gamma I$ should not lead to plastic deformation. This is achieved by considering only the deviatoric part $\operatorname{dev}(\sigma) := \sigma - \frac{1}{d}\operatorname{tr}(\sigma)I$ of the stress.
Then we combine this with the Frobenius norm to arrive at the yield function

$$\phi_{\mathrm{VM}}(\sigma) := \|\operatorname{dev}(\sigma)\|_F - \sigma_0. \tag{2.8}$$

The scalar value $\sigma_0 > 0$ is an experimentally measured quantity. If we only consider the singular values $\sigma_1, \dots, \sigma_d$ of σ, the elastic region $\mathcal{E}$ as an infinite cylinder centered at the origin, and orientated parallel to the $(1, \dots, 1)^T$ vector. The cylinder intersects at the axes at the coordinate $\sqrt{\frac{d}{d-1}}\sigma_0$, as seen in Figure 2.4 for the case $d = 3$.

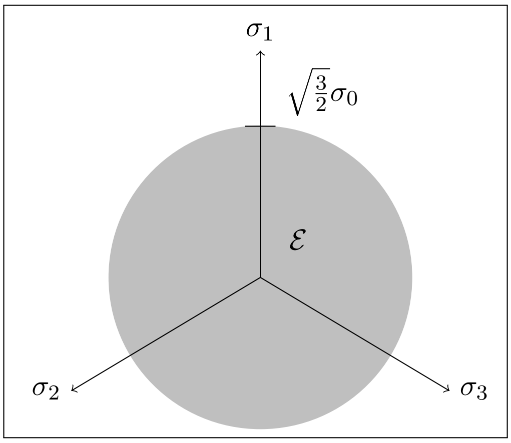


Figure 2.4.: Elastic region of the von Mises yield function in three-dimensional principal space. The view is along the $(1, 1, 1)^T$-vector.

Remark 2.4.2. *The von Mises yield function is sometimes defined with a scaled* σ_0 *as*

$$\phi_{VM}(\sigma) := \|\operatorname{dev}(\sigma)\|_F - \sqrt{\frac{d-1}{d}}\sigma_0.$$

This ensures that the intersection is at σ_0 *but we have to carry the factor* $\sqrt{\frac{d-1}{d}}$ *thought all computations. This is why we stick to original definition* (2.8).

2.4.2. Tresca Yield Criterion

A classical candidate for a nonsmooth elastic region is the Tresca yield criterion. Instead of measuring the total deviatoric stress, like the von Mises yield function, the Tresca yield criterion measures the shear stress. The shear stress is given by the maximal difference in the singular values $\sigma_1, \dots, \sigma_d$ of σ. Thus, we define

$$\phi_{\mathrm{Tr}}(\sigma) := \max_{i,j=1,\dots,d} |\sigma_i - \sigma_j| - \sigma_0. \tag{2.9}$$

Again, $\sigma_0 > 0$ is an experimentally measured value. Note that the Tresca yield function is not affected by hydrostatic stresses like the von Mises criterion. For $d = 3$, the geometry of the elastic region is displayed in Figure 2.5. We can see that the boundary is nonsmooth. The intersection with the principal axes is directly given at σ_0.

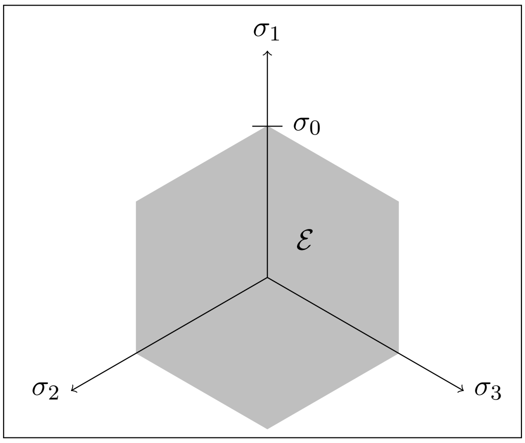


Figure 2.5.: Elastic region of the Tresca yield function in three-dimensional principal space. The view is along the $(1, 1, 1)^T$-vector.

In two dimensions the von Mises and Tresca yield function are equivalent if we use a scaled yield stress σ_0.

Lemma 2.4.3. *Let $d = 2$ and $\sigma \in \mathbb{S}^2$. Then*

$$\|\operatorname{dev}(\sigma)\|_F = \frac{1}{\sqrt{2}}|\sigma_1 - \sigma_2|.$$

Proof. Consider the eigenvalues σ_1, σ_2 of σ. The eigenvalues of the deviatoric part of σ are $\sigma_1 - \frac{1}{2}(\sigma_1 + \sigma_2)$ and $\sigma_2 - \frac{1}{2}(\sigma_1 + \sigma_2)$, respectively, which is $\pm\frac{1}{2}(\sigma_1 - \sigma_2)$. Hence the Frobenius norm is given by $\frac{1}{\sqrt{2}}|\sigma_1 - \sigma_2|$ and the proof is complete. ☐

2.5. Plastic Strain Evolution

In order to determine the evolution of the plastic strain, we can use the Maximum Work Principle [15, 16, 31]. We are not giving a derivation of the following result in this thesis. The interested reader is invited to follow the stated sources for a more in-depth discussion of the origin. There the origin of the term “work” is explained in the context of thermodynamical considerations in the stress–strain space.

Proposition 2.5.1 (Maximum Work Principle)**.** *Consider a stress $\sigma \in \mathbb{S}^d$ with $\phi(\sigma) = 0$. Assume that $\frac{d}{dt}\phi(\sigma) > 0$. Then we have an evolution of the plastic strain $\dot{p} \in \mathbb{S}^d_0$ which satisfies*

$$\dot{p} : \sigma \geq \dot{p} : \tau \qquad \forall \tau \in \mathcal{E}.$$

Recall that in this thesis the elastic region $\mathcal{E}$ is a convex set. Therefore, equivalent characterizations of the Maximum Work Principle are possible using convex analysis (cf. [8]). We recall some tools for the following results.

Definition 2.5.2 (Normal cone)**.** *Let V a vector space, V' its dual, $\mathcal{E} \subseteq V$ a convex subset and $x \in \mathcal{E}$. We call*

$$N_{\mathcal{E}}(x) := \{\mu \in V' \colon \langle \mu, y - x \rangle \leq 0 \ \forall y \in \mathcal{E}\}$$

the normal cone of $\mathcal{E}$ at x.

A direct consequence is that $N_{\mathcal{E}}(x) = \{0\}$ if x is not at the boundary of $\mathcal{E}$. We can reformulate the Maximum Work Principle as

$$\dot{p} \in N_{\mathcal{E}}(\sigma)$$

for any stress $\sigma \in \mathcal{E}$. The normal cone is the outer normal of the boundary if $\mathcal{E}$ is smooth at σ. For the Tresca yield criterion, at one of the corners of $\mathcal{E}$, the cone looks like shown in Figure 2.6.

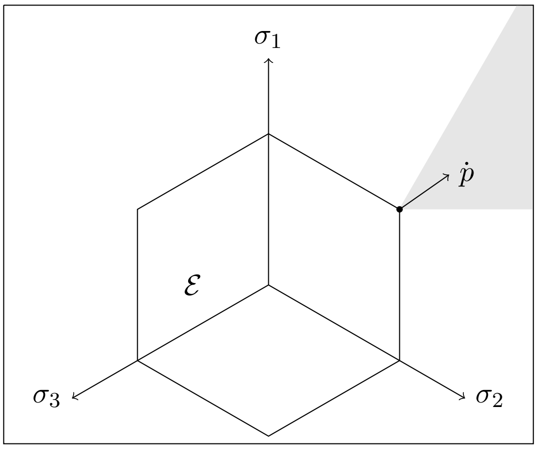


Figure 2.6.: Normal cone (grey) at a nonsmooth boundary point of the elastic region $\mathcal{E}$

Another tool from convex analysis is the subdifferential as an extension of the classical derivative.

Definition 2.5.3 (Convex subdifferential)**.** *Let V a vector space, V' its dual, $\phi \colon V \to \mathbb{R}$ convex and $x \in V$. We call the set*

$$\partial\phi(x) := \{\mu \in V' \colon \phi(y) \geq \phi(x) + \langle \mu, y - x \rangle \ \forall y \in V\}$$

the subdifferential of ϕ at x.

If ϕ is differentiable at x, we have $\partial\phi(x) = \{\nabla\phi(x)\}$. Since the boundary of the elastic region $\mathcal{E}$ is the level set of the yield function ϕ, we can characterize the normal cone of $\mathcal{E}$ directly by the subdifferential of the yield function ϕ. Note that the normal cone is infinite. Hence, we have no information about the size of its elements. Therefore, there is a scalar $\gamma > 0$, called the plastic multiplier, such that

$$\dot{p} \in \gamma \, \partial\phi(\sigma).$$

We have only an evolution in p if we reach a stress σ with $\phi(\sigma) = 0$. Hence, we demand

$$\gamma > 0, \quad \phi(\sigma) \leq 0, \quad \gamma\phi(\sigma) = 0 \tag{2.10}$$

to hold simultaneously.

2.6. Hardening

In the idealized stress–strain curve in Figure 2.3, we are now able to describe the evolution of a plastic strain p after arriving at (B). We need to introduce a concept to describe the appearing plastic loading slope. Moreover, we notice that stresses beyond the yield stress σ_0 are reached, which is not covered by the theory for far. In addition, the value of p seems to depend on the stress. This is contrary to the previously stated assumptions on p.

We introduce a concept of a changing elastic region $\mathcal{E}$ depending on the history of events. This is called *hardening*. In order to track these changes, we need to introduce *internal variables*. Formally, there are internal variables which become part of the stress, and other internal variables which become part of the plastic strain. These variables keep track of the history of plastic effects and of changes of the elastic region.

The detailed definition of the internal variables depends on the underlying hardening model. In literature [16, 31] two main categories of hardening are described.

2.6.1. Kinematic Hardening

Until now, the elastic region was always at a fixed place in the stress space. In the kinematic hardening model we assume that $\mathcal{E}$ can move in the direction of the plastic evolution $\dot{p}$ [31, Section 1.2.4.1]. In mathematical terms there exists an internal stress variable $a \in \mathbb{S}^d$, called back stress. The back stress tracks the movement of the elastic region. Since the internal variables affect the yield function, a modified variant of ϕ is defined. The modified variant ϕ_{kin} for the kinematic stress $\Sigma_{\text{kin}} := (\sigma, a)$ takes the form

$$\phi_{\text{kin}}(\sigma, a) := \phi(\sigma + a)$$

and shows therefore directly the movement of the elastic region.

From thermodynamics [16, Section 2.7], we postulate the existence of a corresponding internal plastic strain variable $\alpha \in \mathbb{S}^d$. We only consider linear connections between a and α [16, 31]

$$\exists k_1 > 0 : \qquad a = -k_1\alpha. \tag{2.11}$$

In theory, however, nonlinear connections are possible as well.

The corresponding kinematic plastic strain is denoted by $P_{\mathrm{kin}} := (p, \alpha)$. In Figure 2.7 an example for $d = 2$ is shown in the principal space. In this picture the stress σ reaches the boundary of the initial elastic region and causes a movement towards a new center at $-a$.

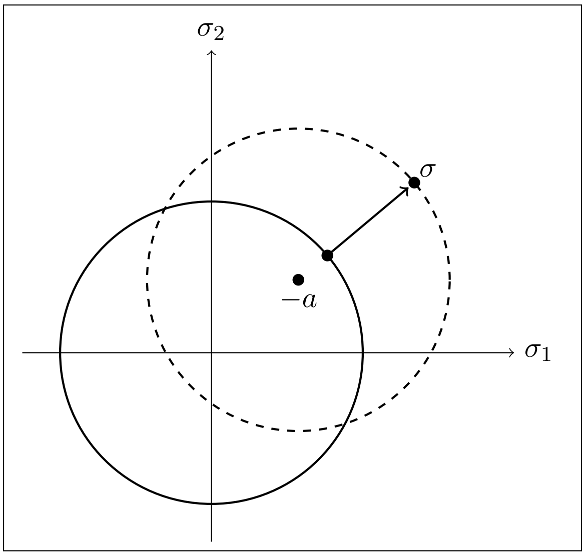


Figure 2.7.: Kinematic hardening in principal stress space for $d = 2$. The initial elastic region is represented by solid lines. The moved elastic region is dashed with new center at $-a$.

2.6.2. Isotropic Hardening

Besides moving the elastic region as in the kinematic hardening above, growth of the elastic region is another possibility to reach larger stress values. We consider for simplicity that the region grows uniformly (isotropically) in all directions. The simplest way to achieve this it to increase the value of σ_0 in the yield functions (2.8) and (2.9). Technically, we add a negative scalar[2] $g \leq 0$ to the yield function to obtain

$$\phi_{\mathrm{iso}}(\sigma, g) := \phi(\sigma) + g.$$

There is again a corresponding scalar internal variable η of the plastic strain. The internal variable η describes the influence of the growth parameter g on the plastic strain. The isotropic stress is defined by $\Sigma_{\mathrm{iso}} := (\sigma, g)$, and, correspondingly, the isotropic plastic strain by the pair $P_{\mathrm{iso}} := (p, \eta)$.

In Figure 2.8, an example for $d = 2$ is shown in principal stress space, where a yield stress is reached and a growth of the elastic region follows.

---

[2] We usually denote scalars by Greek letters, but we stick to the notation of [16].

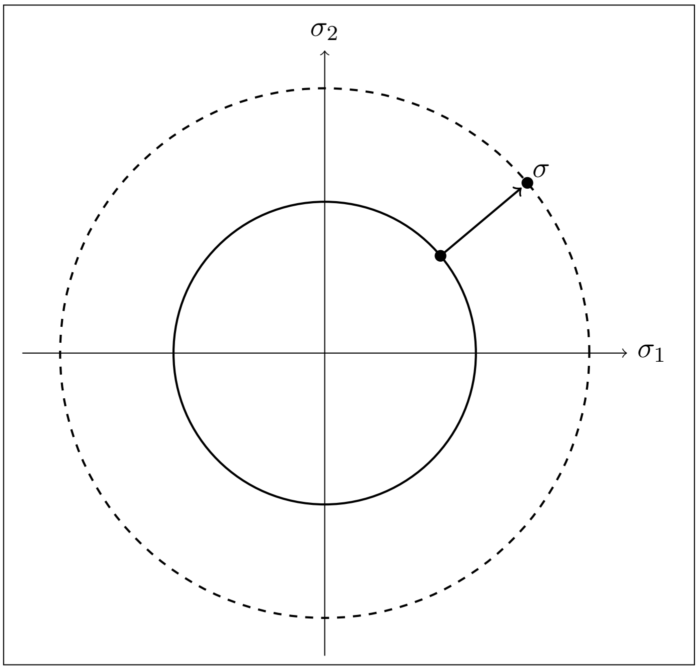


Figure 2.8.: Isotropic hardening in principal stress space for $d = 2$. The initial elastic region is represented by solid lines. The grown elastic region is shown by dashed lines.

For the connections of g and η multiple variants are studied in this thesis. We consider linear connections, e.g., [16, Secion 3.5]

$$\exists k_2 \geq 0 : \qquad g = -k_2\eta. \tag{2.12}$$

In addition, we also consider a more complex nonlinear connection from [32, page 393]

$$g(\eta) = -(k_\infty - \sigma_0)(1 - e^{-\beta\eta})$$

with non-negative constants $k_\infty \geq \sigma_0$ and β.

We want to consider both linear and nonlinear isotropic hardening at the same time. Therefore we define the isotropic hardening relationship

$$g(\eta) = -(k_\infty - \sigma_0)(1 - e^{-\beta\eta}) - k_2\eta. \tag{2.13}$$

Remarks 2.6.1. *The isotropic hardening connection* (2.13) *has the following properties.*

(i) g is continuous and strongly monotone decreasing.

(ii) We have definiteness

$$0 = g(\eta) \Leftrightarrow \eta = 0. \tag{2.14}$$

(iii) We only consider growth of the elastic region, hence $g \leq 0$. From monotonicity and the equation above we conclude that

$$\eta \geq 0. \tag{2.15}$$

2.7. Generalized Formulation

In this thesis we combine both kinematic and isotropic hardening. We define the generalized stress

$$\Sigma := (\sigma, a, g) \in \mathbb{S}^d \times \mathbb{S}_0^d \times \mathbb{R}_{\leq 0}$$

as a tuple of the stress σ, the plastic back-stress a, and the isotropic hardening variable g. Likewise the generalized plastic strain is

$$P := (p, \alpha, \eta) \in \mathbb{S}_0^d \times \mathbb{S}_0^d \times \mathbb{R}_{\geq 0}.$$

The combined yield function in the generalized stress space is

$$\Phi(\Sigma) := \phi(\sigma + a) + g.$$

We define the generalized elastic region

$$\mathbb{E} := \{\Sigma : \ \Phi(\Sigma) \leq 0\}.$$

Since ϕ is a convex function, Φ is a convex function, and hence $\mathbb{E}$ is a convex set.

With the generalized stress Σ and plastic strain P a thermodynamical framework can be developed. This is left out in this thesis, detail are found in [16, Section 3] with the same notation of this thesis. There it is also stated that the Maximum Work Principle still holds for the generalized framework.

Proposition 2.7.1 (Generalized Maximum Work Principle)**.** *Let $\Sigma = (\sigma, a, g)$ be a generalized yield stress, i.e., $\Phi(\Sigma) = 0$. Further let $\frac{d}{dt}\Phi(\Sigma) > 0$. Then we have an evolution of the generalized plastic strain $P = (p, \alpha, \eta)$ characterized by*

$$\dot{P} \diamond \Sigma := \dot{p} : \sigma + \dot{\alpha} : a + \dot{\eta} : g \geq T \diamond \Sigma \qquad \forall T \in \mathbb{E}. \tag{2.16}$$

In Figure 2.9 a two dimensional example of the changing elastic region is given for combined kinematic and isotropic hardening. The ratio between movement and growth is determined by the hardening parameters k_1, k_2, k_∞ and β of the previous sections.

Since $\mathbb{E}$ is convex set, we can again write

$$\dot{P} \in N_{\mathbb{E}}(\Sigma) \tag{2.17}$$

or, equivalently,

$$\dot{P} \in \gamma \cdot \partial\Phi(\Sigma)$$

for the subdifferential $\partial\Phi(\cdot)$ of Φ and a plastic multiplier γ. In detail, we have

$$\dot{P} = \begin{pmatrix} \dot{p} \\ \dot{\alpha} \\ \dot{\eta} \end{pmatrix} \in \gamma \cdot \begin{pmatrix} \partial\phi(\sigma + a) \\ \partial\phi(\sigma + a) \\ \{1\}. \end{pmatrix}$$

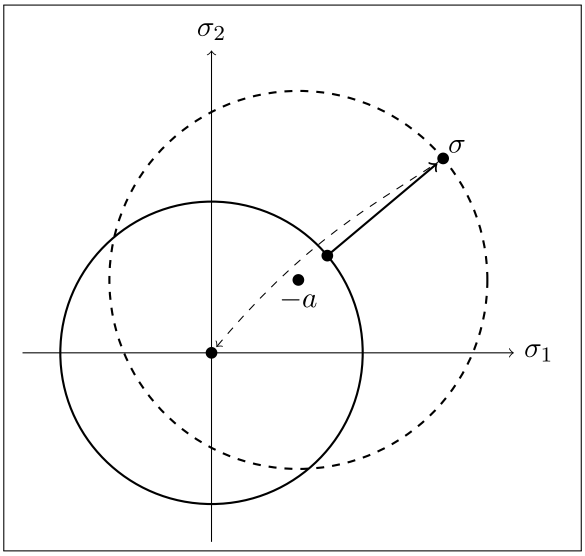


Figure 2.9.: Kinematic and isotropic hardening in principal stress space for $d = 2$. The initial elastic region is represented by solid lines. The new elastic region is dashed with new center at $-a$. The dashed arrow demonstrates that elastic unloading to zero stress is possible in the new elastic region.

From the above, we have the pointwise condition $\dot{p} = \dot{\alpha}$. If we assume that $p = \alpha = 0$ for the initial state at $t = 0$, we can conclude that for all t

$$p = \alpha. \tag{2.18}$$

If we return to the idealized stress–strain curve in Figure 2.3, the plastic loading from (B) to (C), and the elastic unloading path from (C) to (D) have not been discussed yet. The changes in internal variables a and g allow stresses σ to be greater than σ_0, because the elastic region moves and grows. Once we decrease the stress at (C), we are in a situation as displayed in Figure 2.9. The stress-free state is contained in the new elastic region. No further plastic strain evolution takes place while unloading from (C) to (D). Hence, the last part is purely elastic.

3. Variational Formulations of Small-Strain Plasticity

The generalized Maximum Work Principle is the starting point for variational formulations of the small-strain plasticity problem. From here, we have to choose what variables will be the unknowns of the global system of equations. Basically, we have two options.

The first possibility is the *dual formulation*. In this form, we start start directly with the inclusion $\dot{P} \in N_{\mathbb{E}}(\Sigma)$ (2.17). In this form we can eliminate the generalized plastic strain P by other, already known, conditions on Σ. The generalized stress is one of the unknowns of this form.

The other form is the *primal formulation*. Here, we want to eliminate the generalized stress Σ from the model. Therefore, we need to introduce further structures to express Σ directly in terms of $\dot{P}$. Then the generalized plastic strain P is one of the unknowns.

Both formulations exist in literature. One benefit of the dual formulation is that we use the physically measurable stress to describe the unknowns. In contrast, the plastic strain is a model quantity that cannot be measured directly. Nevertheless, the primal formulation leads to series of nonsmooth minimizing problems after a suitable time discretization, while the dual formulation does not.

3.1. Excursus: Dual Formulation of Plasticity

Before we dive directly into the primal formulation of plasticity we give a short overview how the plasticity problem is treated in the dual formulation. The reader can skip the following section since its contents will not play a role in the rest of this thesis.

In this section, no details about the specific function spaces of the solutions are given. In the dual form, the displacement field $\mathbf{u}$, and the global stress fields $\boldsymbol{\sigma}, \mathbf{a}$ and $\mathbf{g}$ are the quantities of interest. For simplicity, only a linear form of kinematic and isotropic hardening is considered. Furthermore, only volumetric forces $\mathbf{b}$ are present, and no surface forces $\mathbf{s}$, as defined in (2.3).

We can express the plastic variables in terms of the stress variables

$$\boldsymbol{\alpha} = -\frac{1}{k_1}\mathbf{a}, \quad \boldsymbol{\eta} = -\frac{1}{k_2}\mathbf{g}$$

with positive coefficients $k_1, k_2 > 0$, cf. (2.11) and (2.12). From the generalized Maximum Work Principle (2.17), we incorporate the definition of the normal cone

$$\langle \dot{\mathbf{P}}, \tilde{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma} \rangle \leq 0 \qquad \forall \tilde{\boldsymbol{\Sigma}} \in \mathbb{E}.$$

For a volumetric force density $\mathbf{b}$ we have the balance of momentum (2.4)

$$-\operatorname{div}(\boldsymbol{\sigma}) = \mathbf{b}.$$

With that in mind, we define the time-dependent linear load functional $\langle \mathbf{l}(t), \cdot \rangle$ by

$$\langle \mathbf{l}(t), \mathbf{u} \rangle := -\int_\Omega \mathbf{b}(t) \cdot \mathbf{u} \, dx.$$

A bilinear form $b(\cdot, \cdot)$ is introduced by

$$b(\mathbf{u}, \boldsymbol{\sigma}) := -\int_\Omega \boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\sigma} \, dx.$$

Using the symmetry of the arguments one can show by integrating by parts[1] that

$$b(\tilde{\mathbf{u}}, \boldsymbol{\sigma}) = \int_\Omega \operatorname{div}(\boldsymbol{\sigma}) \cdot \tilde{\mathbf{u}} = \langle \mathbf{l}(t), \tilde{\mathbf{u}} \rangle \qquad \forall \tilde{\mathbf{u}}.$$

Under the assumption that the Hooke tensor $\mathbf{H}$ from (2.6) is invertible we define the bilinear stress–stress mapping

$$a(\boldsymbol{\sigma}, \tilde{\boldsymbol{\sigma}}) := \int_\Omega \boldsymbol{\sigma} : \mathbf{H}^{-1} \tilde{\boldsymbol{\sigma}} \, dx,$$

and additional bilinear functions for the generalized variables

$$\begin{aligned} c_1(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) &:= \int_\Omega \frac{1}{k_1} \boldsymbol{\alpha} : \tilde{\boldsymbol{\alpha}} \, dx, \\ c_2(\mathbf{g}, \tilde{\mathbf{g}}) &:= \int_\Omega \frac{1}{k_2} \mathbf{g} \cdot \tilde{\mathbf{g}} \, dx. \end{aligned}$$

Altogether, we define the generalized bilinear form

$$A(\boldsymbol{\Sigma}, \tilde{\boldsymbol{\Sigma}}) := a(\boldsymbol{\sigma}, \tilde{\boldsymbol{\sigma}}) + c_1(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) + c_2(\mathbf{g}, \tilde{\mathbf{g}}).$$

The explicit form of (2.17) is

$$\int_\Omega \dot{\mathbf{p}} : (\tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma}) + \dot{\mathbf{a}} : (\tilde{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) + \dot{\boldsymbol{\eta}} \cdot (\tilde{\mathbf{g}} - \mathbf{g}) \, dx \leq 0.$$

Using both

$$\dot{\boldsymbol{\alpha}} = -k_1 \dot{\mathbf{a}} \quad \text{and} \quad \dot{\mathbf{g}} = -k_2 \dot{\boldsymbol{\eta}},$$

and Hooke's law (2.6)

$$\boldsymbol{\epsilon}(\dot{\mathbf{u}}) = \dot{\mathbf{p}} + \dot{\mathbf{e}} = \dot{\mathbf{p}} + \mathbf{H}^{-1} \dot{\boldsymbol{\sigma}} \quad \Rightarrow \quad \dot{\mathbf{p}} = \boldsymbol{\epsilon}(\dot{\mathbf{u}}) - \mathbf{H}^{-1} \dot{\boldsymbol{\sigma}},$$

---

[1] This is done in detail in Section 3.3.

we arrive at

$$\int_\Omega \underbrace{\epsilon(\dot{\mathbf{u}}) : (\tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma})}_{-\mathbf{b}(\dot{\mathbf{u}},\tilde{\boldsymbol{\sigma}}-\boldsymbol{\sigma})} - \underbrace{\mathbf{H}^{-1}\dot{\boldsymbol{\sigma}} : (\tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma})}_{a(\dot{\boldsymbol{\sigma}},\tilde{\boldsymbol{\sigma}}-\boldsymbol{\sigma})} - \underbrace{\frac{1}{k_1}\dot{\boldsymbol{\alpha}} : (\tilde{\boldsymbol{\alpha}} - \boldsymbol{\alpha})}_{c_1(\dot{\boldsymbol{\alpha}},\tilde{\boldsymbol{\alpha}}-\boldsymbol{\alpha})} - \underbrace{\frac{1}{k_2}\dot{\mathbf{g}} \cdot (\tilde{\mathbf{g}} - \mathbf{g})}_{c_2(\dot{\mathbf{g}},\tilde{\mathbf{g}}-\mathbf{g})} \, dx \leq 0.$$

This can be shortened to

$$A(\dot{\boldsymbol{\Sigma}}, \tilde{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}) + b(\dot{\mathbf{u}}, \tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma}) \geq 0 \quad \forall \tilde{\boldsymbol{\Sigma}} \in \mathbb{E}. \tag{3.1}$$

Together with the previously stated linear constraint

$$b(\tilde{\mathbf{u}}, \boldsymbol{\sigma}) = \langle \mathbf{l}(t), \tilde{\mathbf{u}} \rangle \qquad \forall \tilde{\mathbf{u}} \tag{3.2}$$

the dual formulation will result in a classical quadratic inequality on a convex set with a linear constraint.

Further numerical treatment of this variational problem is out of the scope of this thesis and therefore intentionally omitted. A detailed description of the construction of the fully discrete problem and numerical solution strategies is given in in [16, Section 13].

3.2. Support Functions and Dissipation

In this section we want to investigate a necessary tool for the primal formulation of plasticity. We want to express the global system of equations in terms of the generalized plastic strain $\mathbf{P}$ and the displacement $\mathbf{u}$. Hence, we need a description of the generalized plastic strain in terms of the generalized stress $\boldsymbol{\Sigma}$. The results of this section are pointwise results, i.e., we consider $P = \mathbf{P}(x)$, $\Sigma = \boldsymbol{\Sigma}(x)$, etc. for $x \in \Omega$.

We start with polar functions from convex analysis.

Definition 3.2.1 (Polar function)**.** *Let V a vector space, V' its dual. Consider a convex function $\Phi := V \to \mathbb{R} \cup \{\infty\}$. We call*

$$\operatorname{dom}\Phi := \{v \in V \colon \Phi(v) < \infty\}$$

the domain of Φ. Then we can define the polar function

$$\Phi^* \colon V' \to \mathbb{R} \cup \{\infty\}, \quad \Phi^*(v^*) := \sup_{v \in \operatorname{dom}\Phi} \langle v^*, v \rangle - \Phi(v).$$

If we compute the subdifferential for both Φ and Φ^*, we can show an important property.

Theorem 3.2.2 ([8, Corollary 5.2])**.** *Let V a vector space and $\Phi \colon V \to \mathbb{R}$ convex, lower semicontinuous and* $\operatorname{dom}\Phi \neq \emptyset$*. Furthermore, let Φ^* the polar function of Φ. Then we have*

$$x^* \in \partial\Phi(x) \quad \Leftrightarrow \quad x \in \partial\Phi^*(x^*).$$

The proof of this is rather lengthy but it offers the reader a journey through the basics of convex analysis.

We return to plasticity by connecting polar functions to the generalized elastic domain $\mathbb{E}$. Note that $\mathbb{E}$ is closed. We define the indicator function of $\mathbb{E}$ by

$$\chi_{\mathbb{E}}(\Sigma) := \begin{cases} 0 & \Sigma \in \mathbb{E} \\ \infty & \Sigma \notin \mathbb{E}. \end{cases}$$

Note that $\chi_{\mathbb{E}}$ is a convex function if and only if $\mathbb{E}$ is a convex set. The subdifferential of $\chi_{\mathbb{E}}$ is empty outside of $\mathbb{E}$ and $\{0\}$ within $\chi_{\mathbb{E}}$. At the boundary point $\Sigma \in \partial\mathbb{E}$ we have

$$Q \in \partial\chi_{\mathbb{E}}(\Sigma) \quad \Leftrightarrow \quad \chi_{\mathbb{E}}(\tilde{\Sigma}) \geq \chi_{\mathbb{E}}(\Sigma) + \langle Q, \tilde{\Sigma} - \Sigma \rangle$$

for all $\tilde{\Sigma}$ inside or outside of $\mathbb{E}$. If $\tilde{\Sigma}$ is outside of $\mathbb{E}$ the right side is always true. If $\tilde{\Sigma} \in \mathbb{E}$ the equivalence above is equal to

$$0 \geq \langle Q, \tilde{\Sigma} - \Sigma \rangle.$$

This is exactly the generalized flow rule (2.17)

$$Q \in N_{\mathbb{E}}(\Sigma)$$

with $Q = \dot{P}$. Hence, we can conclude that a plastic flow fulfills

$$\dot{P} \in \partial\chi_{\mathbb{E}}(\Sigma). \tag{3.3}$$

We want to apply Theorem 3.2.2 to (3.3) and therefore define the polar function of $\chi_{\mathbb{E}}(\Sigma)$. For indicator functions this polar function is also called support function of $\mathbb{E}$, and in the context of plasticity, dissipation function.

Definition 3.2.3 (Dissipation function)**.** *For a given generalized elastic region $\mathbb{E}$ we define the dissipation function*

$$D : \mathbb{S}_0^d \times \mathbb{S}_0^d \times \mathbb{R}_{\geq 0} \to \mathbb{R} \cup \{\infty\},\ D(P) := \sup_{\Sigma \in \mathbb{E}} P \diamond \Sigma.$$

Note that the definition of the dissipation function is simpler than for general polar functions since the indicator function is zero within $\mathbb{E}$. With the dissipation function at hand, we use Theorem 3.2.2 and formulate the connection of dissipation function and the yield function

$$\boxed{\Sigma \in \partial D(\dot{P}) \quad \Leftrightarrow \quad \dot{P} \in \partial\Phi(\Sigma).} \tag{3.4}$$

We can give explicit formulas of the dissipation functional for both the von Mises and the Tresca flow rule.

Theorem 3.2.4 (Explicit dissipation functions)**.** *Let $\Phi(\Sigma) := \phi(\sigma + a) - g$ denote the generalized yield function with the internal variables a, g, and let $\sigma_0 > 0$. Moreover, consider the generalized plastic strain $P = (p, \eta)$ without α, since α is eliminated by identifying with p (2.18).*

(i) If $\phi = \phi_{VM}$ is the von Mises yield function (2.8)*, then*

$$D_{VM}(P) = D_{VM}(p,\eta) = \begin{cases} \sigma_0 \left\|p\right\|_F & \left\|p\right\|_F \le \eta \\ \infty & \textit{otherwise}, \end{cases} \tag{3.5}$$

where $\left\|\cdot\right\|_F$ denotes the Frobenius norm.

(ii) If $\phi = \phi_{Tr}$ is the Tresca yield function (2.9)*, then*

$$D_{Tr}(P) = D_{Tr}(p,\eta) = \begin{cases} \sigma_0 \left\|p\right\|_2 & \left\|p\right\|_2 \le \eta \\ \infty & \textit{otherwise}, \end{cases} \tag{3.6}$$

where $\left\|\cdot\right\|_2$ denotes the spectral norm. Since p is symmetric, this norm is equivalent to the absolute value of the largest eigenvalue.

Proof. Proofs can be found in [29], but we recall them here for a better understanding of the generalized yield functions.

(i) The generalized elastic region is given by

$$\mathbb{E}_{\mathrm{VM}} = \{\left\|\operatorname{dev}(\sigma + a)\right\|_F - \sigma_0 + g \le 0\}.$$

Note that $g \le 0$, since we only consider growth of the elastic region. Moreover, from (2.18) we have $P \diamond \Sigma = p : (\sigma + a) + \eta g$. Hence, the dissipation function is given by

$$\begin{aligned} D(P) = D(p,\eta) &= \sup_{g\le 0}\left\{ \sup_{\sigma + a \in \mathbb{S}^d} \{p : (\sigma + a) + \eta g,\ \left\|\operatorname{dev}(\sigma + a)\right\|_F - \sigma_0 + g \le 0\}\right\} \\ &= \sup_{g\le 0}\left\{ \sup_{\sigma + a \in \mathbb{S}^d} \{p : \operatorname{dev}(\sigma + a),\ \left\|\operatorname{dev}(\sigma + a)\right\|_F \le \sigma_0 - g\} + \eta g\right\}. \end{aligned}$$

In the second line we use that p is trace-free. Since $\operatorname{dev}(\sigma + a)$ is only restricted by the Frobenius norm, the scalar product with p is maximized if it is parallel to p. The scalar product of parallel arguments is the product of the norms. Hence, the inner supremum satisfies

$$\sup_{\sigma + a \in \mathbb{S}^d} \{p : \operatorname{dev}(\sigma + a),\ \left\|\operatorname{dev}(\sigma + a)\right\|_F \le \sigma_0 - g\} = (g - \sigma_0)\left\|p\right\|_F .$$

This inserted into the outer supremum leads to

$$D(p,\eta) = \sup_{g\le 0}\{(\sigma_0 - g)\left\|p\right\|_F + \eta g\} = \sup_{g\le 0}\{(\left\|p\right\|_F - \eta)g + \sigma_0 \left\|p\right\|_F\}.$$

Then the supremum is given for either $g = 0$ or $g \to -\infty$, depending on the sign of $\left\|p\right\|_F - \eta$.

(ii) The proof for the Tresca yield criterion reuses some techniques from (i). Nevertheless, some rather technical details need to be considered, and therefore we decided to put this into Appendix A.1.

□

Both norms $\|\cdot\|_F$ and $\|\cdot\|_2$ are convex by the triangle inequality and the positive 1-homogeneity ($\|\gamma p\| = \gamma \|p\|$ for $\gamma \geq 0$). Nevertheless, computing derivatives of these norms is different.

Lemma 3.2.5.

(i) The Frobenius norm $\|\cdot\|_F : \mathbb{R}^{d\times d} \to \mathbb{R}\colon A \mapsto \|A\|_F$ *is smooth everywhere, except at* $A = 0$.

(ii) For $A \neq 0$ *we have*

$$\nabla \|A\|_F = \frac{A}{\|A\|_F}.$$

(iii) The subdifferential of the Frobenius norm at $A = 0$ *is*

$$\partial \|0\|_F = \{B \in \mathbb{R}^{d\times d}\colon \ \|B\|_F \leq 1\}.$$

(iv) On the vector space of symmetric matrices, the spectral norm $\|S\|_2$ *is differentiable if and only if the dominant eigenvalue (the eigenvalue with the largest absolute value) of* S *is unique. We call a dominant eigenvalue* σ^* *unique, if it has algebraic multiplicity one.*

Proof. (i) and (ii) are trivially shown by using the definition of the Frobenius norm. (iii) is a direct consequence of the positive 1-homogeneity of the Frobenius norm. (iv) follows from the equivalent definition

$$\|S\|_2 = \max_{i=1,\dots,d} |\sigma_i|$$

for the eigenvalues σ_i of S. It is easy to see that the maximum function

$$(\sigma_1, \dots, \sigma_d) \mapsto \max_{i=1,\dots,d} |\sigma_i|$$

is smooth if and only if the maximum is uniquely determined by one of the σ_i. Moreover let $\mathcal{P}(S, \sigma) = \det(S - \sigma I)$ the characteristic polynomial of S. Each eigenvalue fulfills $\mathcal{P}(S, \sigma_i(S)) = 0$. Since $\mathcal{P}$ is a polynomial in the arguments, we can derive by S and get

$$\nabla_S \mathcal{P}(S, \sigma_i(S)) + \nabla_\sigma \mathcal{P}(S, \sigma_i(S)) \nabla \sigma_i(S) = 0.$$

Therefore, since $\nabla_\sigma \mathcal{P}(S, \sigma_i(S)) \in \mathbb{R}$,

$$\nabla \sigma_i(S) = -\frac{\nabla_S \mathcal{P}(S, \sigma_i(S))}{\nabla_\sigma \mathcal{P}(S, \sigma_i(S))}$$

is defined if and only if $\nabla_\sigma \mathcal{P}(S, \sigma_i(S)) \neq 0$. This is true if and only if the eigenvalue σ_i has algebraic multiplicity one. □

3.3. Variational Formulation of Primal Plasticity

Now we have everything together to state the primal formulation of plasticity in the weak form. In contrast to the quick introduction of the dual formulation in Section 3.1, a more detailed overview is given here. In contrast to the previous section, we consider the global functions such as $\mathbf{u}$ and $\mathbf{P}$ over Ω and use boldface notation again.

First, we need to define suitable function spaces for the unknowns. We consider the displacement $\mathbf{u}$ and the (generalized) plastic strain $\mathbf{P} = (\mathbf{p}, \boldsymbol{\eta})$. All other appearing quantities can be deduced from there for a given point in time t: The stress $\boldsymbol{\sigma}$ is given by Hooke's law (2.6) together with the additive split of the strain (2.5). The internal stress variables $\mathbf{a}$ and $\mathbf{g}$ are given by the hardening relations (2.11) and (2.13), respectively. Note that we identify $\mathbf{p} = \mathbf{a}$ (2.18) in the plastic strain space. Hence, in summary, the generalized stress $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \mathbf{a}, \mathbf{g})$ is explicitly computable from the displacement and generalized plastic strain. Moreover, the evolution of the plastic strain is now defined thanks to the duality of the dissipation function and the yield function (3.4).

In the following we introduce the explicit spaces for the unknowns. The choice is made such that we allow as much numerical flexibility as possible, but are still able to show existence results.

The displacement $\mathbf{u}$ needs to be weakly differentiable, since want to define the strain $\boldsymbol{\epsilon}(\mathbf{u})$. Consider additionally Dirichlet boundary values on a subset Γ of the boundary of Ω. In the terms of continuum mechanics Dirichlet conditions are usually only applied to the displacement field $\mathbf{u}$ with the condition that $\mathbf{u}|_\Gamma \equiv 0$. Together with the Dirichlet boundary an obvious choice for the function space of $\mathbf{u}$ is

$$V := \left(H^1_\Gamma(\Omega)\right)^d = \left\{\mathbf{v} \in H^1(\Omega, \mathbb{R}^d) : \ \mathbf{v}|_\Gamma \equiv 0\right\}. \tag{3.7}$$

In contrast, the resulting stress $\boldsymbol{\sigma}$ only requires integrability for the upcoming scalar products and we choose the simple L^2 space

$$Q := L^2(\Omega, \mathbb{S}^d) \tag{3.8}$$

for the stress.

Similarly, the plastic strain $\mathbf{p}$ takes trace-free symmetric values, but has no regularity assumptions. Therefore, it is taken from Q_0, which is defined by Q restricted to trace-free matrices. The scalar field $\boldsymbol{\eta}$ for isotropic hardening is taken from

$$M := L^2(\Omega).$$

Combined, the product space of the appearing unknowns $\mathbf{w} := (\mathbf{u}, \mathbf{p}, \boldsymbol{\eta})$ is

$$W := V \times Q_0 \times M.$$

Note that W is a Hilbert space with a standard scalar product inherited from the factor Hilbert spaces.

Starting point of the weak formulation is the flow rule in $\dot{\mathbf{P}}$ given in terms of the subdifferential of the dissipation function (3.4)

$$\mathbf{\Sigma} \in \partial D(\dot{\mathbf{P}}),$$

which is understood pointwise in Ω. Explicitly written, it is

$$D(\tilde{\mathbf{P}}) \geq D(\dot{\mathbf{P}}) + \mathbf{\Sigma} \diamond \left(\tilde{\mathbf{P}} - \dot{\mathbf{P}}\right) \qquad \forall \tilde{\mathbf{P}} = (\tilde{\mathbf{p}}, \tilde{\boldsymbol{\eta}}) \in Q_0 \times M$$

using the generalized scalar product $\diamond$ from (2.16). By the identification $\dot{\boldsymbol{\alpha}} = \dot{\mathbf{p}}$ and $\dot{\mathbf{P}} = (\dot{\mathbf{p}}, \dot{\boldsymbol{\eta}})$ we have

$$D(\tilde{\mathbf{p}}, \tilde{\boldsymbol{\eta}}) \geq D(\dot{\mathbf{p}}, \dot{\boldsymbol{\eta}}) + \boldsymbol{\sigma} : (\tilde{\mathbf{p}} - \dot{\mathbf{p}}) + \mathbf{a} : (\tilde{\mathbf{p}} - \dot{\mathbf{p}}) + \mathbf{g} \cdot (\tilde{\boldsymbol{\eta}} - \dot{\boldsymbol{\eta}}) \qquad \forall (\tilde{\mathbf{p}}, \tilde{\boldsymbol{\eta}}) \in Q_0 \times M.$$

Now, we use the hardening laws $\mathbf{a} = -k_1 \boldsymbol{\alpha} = -k_1 \mathbf{p}$ (2.11) and Hooke's law (2.6) $\boldsymbol{\sigma} = \mathbf{H}\mathbf{e} = \mathbf{H}(\boldsymbol{\epsilon}(\mathbf{u}) - \mathbf{p})$ to eliminate the stress variables $\boldsymbol{\sigma}$ and $\mathbf{a}$ from the inequality and arrive at

$$\begin{aligned} D(\tilde{\mathbf{p}}, \tilde{\boldsymbol{\eta}}) \geq D(\dot{\mathbf{p}}, \dot{\boldsymbol{\eta}}) &+ (\mathbf{H}(\boldsymbol{\epsilon}(\mathbf{u}) - \mathbf{p}) - k_1 \mathbf{p}) : (\tilde{\mathbf{p}} - \dot{\mathbf{p}}) \\ &+ \mathbf{g}(\boldsymbol{\eta}) \cdot (\tilde{\boldsymbol{\eta}} - \dot{\boldsymbol{\eta}}) \qquad \forall (\tilde{\mathbf{p}}, \tilde{\boldsymbol{\eta}}) \in Q_0 \times M. \end{aligned} \tag{3.9}$$

A second component of the primal formulation is the balance of momentum $-\operatorname{div}(\boldsymbol{\sigma}) = \mathbf{b}$ (2.4) for the volumetric force vector field $\mathbf{b}$ from (2.3). Multiplication with a displacement field $\tilde{\mathbf{u}} - \dot{\mathbf{u}} \in V$ yields

$$\int_\Omega \mathbf{b} \cdot (\tilde{\mathbf{u}} - \dot{\mathbf{u}})\, dx = \int_\Omega -\operatorname{div}(\boldsymbol{\sigma}) \cdot (\tilde{\mathbf{u}} - \dot{\mathbf{u}})\, dx. \tag{3.10}$$

We want to perform an integration by parts to resolve the div-operator on the stress. Note that the displacement field is a vector field $\mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_d)^T$, and that $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_1 | \dots | \boldsymbol{\sigma}_d)^T$ is a matrix field with columns $\boldsymbol{\sigma}_i$. The divergence is applied columnwise. Hence, we obtain

$$\begin{aligned} \int_\Omega -\operatorname{div}(\boldsymbol{\sigma}) \cdot \mathbf{u}\, dx &= \sum_{i=1}^d \int_\Omega -\operatorname{div}(\boldsymbol{\sigma}_i) \mathbf{u}_i\, dx \\ &= \sum_{i=1}^d \int_\Omega \boldsymbol{\sigma}_i \cdot \nabla \mathbf{u}_i\, dx - \sum_{i=1}^d \int_{\partial\Omega} \mathbf{u}_i \boldsymbol{\sigma}_i \cdot \mathbf{n}\, dS. \end{aligned} \tag{3.11}$$

A first observation from the above is that for a symmetric $\boldsymbol{\sigma}$ the scalar product with $\nabla \mathbf{u}$ only depends on the symmetric part of $\nabla \mathbf{u}$. This is the small strain tensor $\boldsymbol{\epsilon}(\mathbf{u})$:

$$\sum_{i=1}^d \boldsymbol{\sigma}_i \cdot \nabla \mathbf{u}_i\, dx = \frac{1}{2} \boldsymbol{\sigma} : \left(\nabla \mathbf{u} + \nabla \mathbf{u}^T\right) = \boldsymbol{\sigma} : \boldsymbol{\epsilon}(\mathbf{u}). \tag{3.12}$$

Second, the Cauchy stress $\boldsymbol{\sigma}$ (see Proposition 2.2.1) is given such that $\boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{s}$ for the outer normal $\mathbf{n}$. Hence the last term in (3.11) is the outer surface force field acting on the displacement:

$$\sum_{i=1}^d \int_{\partial\Omega} \mathbf{u}_i \boldsymbol{\sigma}_i \cdot \mathbf{n}\, dS = \int_{\partial\Omega} \mathbf{u} \cdot (\boldsymbol{\sigma} \cdot \mathbf{n})\, dS = \int_{\partial\Omega} \mathbf{u} \cdot \mathbf{s}\, dS.$$

The remaining $\boldsymbol{\sigma}$ in (3.12) is eliminated by Hooke's law $\boldsymbol{\sigma} = \mathbf{H}(\boldsymbol{\epsilon}(\mathbf{u}) - \mathbf{p})$. Now we can rewrite (3.10) without stress components as

$$\int_\Omega \mathbf{b} \cdot (\tilde{\mathbf{u}} - \dot{\mathbf{u}})\, dx = \int_\Omega \mathbf{H}(\boldsymbol{\epsilon}(\mathbf{u}) - \mathbf{p}) \cdot (\boldsymbol{\epsilon}(\tilde{\mathbf{u}}) - \boldsymbol{\epsilon}(\dot{\mathbf{u}}))\; dx - \int_{\partial\Omega} (\tilde{\mathbf{u}} - \dot{\mathbf{u}}) \cdot \mathbf{s}\, dS. \tag{3.13}$$

If we add the equation (3.13) to an integrated version of the inequality (3.9), then we obtain an inequality with $\mathbf{w}(t) \coloneqq (\mathbf{u}(t), \mathbf{p}(t), \boldsymbol{\eta}(t)) \in W$ in the form

$$a(\mathbf{w}(t), \tilde{\mathbf{w}} - \dot{\mathbf{w}}(t)) + j(\tilde{\mathbf{w}}) - j(\dot{\mathbf{w}}(t)) \geq \langle \mathbf{l}(t), \tilde{\mathbf{w}} - \dot{\mathbf{w}}(t) \rangle \qquad \forall \tilde{\mathbf{w}} \in W. \tag{3.14}$$

Here, the function a represents the stored energy of the elastic parts $(\boldsymbol{\epsilon}(\mathbf{u}) - \mathbf{p})$ and the plastic hardening and is given by

$$\begin{aligned} &a \colon W \times W \to \mathbb{R}, \\ &a(\mathbf{w}, \tilde{\mathbf{w}}) \coloneqq \int_\Omega \mathbf{H}(\boldsymbol{\epsilon}(\mathbf{u}) - \mathbf{p}) : (\boldsymbol{\epsilon}(\tilde{\mathbf{u}}) - \tilde{\mathbf{p}}) + k_1 \mathbf{p} : \tilde{\mathbf{p}} - \mathbf{g}(\boldsymbol{\eta}) \cdot \tilde{\boldsymbol{\eta}}\, dx. \end{aligned} \tag{3.15}$$

The functional is linear in the second component, and even bilinear if $\mathbf{g}(\boldsymbol{\eta})$ is linear in $\boldsymbol{\eta}$. The nonsmooth dissipation components are given by the functional

$$\begin{aligned} j \colon W &\to \mathbb{R}, \\ j(\mathbf{w}) &\coloneqq \int_\Omega D(\mathbf{p}, \boldsymbol{\eta})\, dx, \end{aligned} \tag{3.16}$$

which only depends on the plastic components. The missing component is the load functional that stores both the volumetric forces $\mathbf{b}$ and the surface forces $\mathbf{s}$ as a linear functional. This functional acts on the displacement

$$\begin{aligned} \langle \mathbf{l}(t), \cdot \rangle &\colon W \to \mathbb{R}, \\ \langle \mathbf{l}(t), \mathbf{w} \rangle &\coloneqq \int_\Omega \mathbf{b}(t) \cdot \mathbf{u}, dx + \int_{\partial\Omega} \mathbf{u} \cdot \mathbf{s}(t)\, dS \end{aligned}$$

which is of course time-dependent, since this functional controls the deformation process of our material.

4. Discretization

Now the primal formulation of small-strain plasticity is complete. The next step is to discretize the function spaces and the time interval in order to make the problem accessible for computers. We start with the time discretization since it enables us to reformulate the inequality as minimization problems.

4.1. Time Discretization

In inequality (3.14), all unknown functions $\mathbf{w} = (\mathbf{u}, \mathbf{p}, \boldsymbol{\eta})$ are space- and time-dependent. Moreover, both the evolution $\dot{\mathbf{w}}$ and the value $\mathbf{w}$ appear in the inequality. We replace this time derivative with a simple backwards difference scheme. Therefore, we divide the time interval $[0, T]$ into equal pieces of length Δt. We set $t_n := n\Delta t$ for the n-th time step. Let $\mathbf{w}_n \approx \mathbf{w}(t_n)$ denote the time-discrete approximation of the solution at time t_n. Now we simply replace the evolution of $\mathbf{w}$ by

$$\dot{\mathbf{w}}(t_n) \approx \frac{\mathbf{w}_n - \mathbf{w}_{n-1}}{\Delta t} =: \frac{\Delta \mathbf{w}_n}{\Delta t}.$$

With these changes the inequality (3.14) becomes

$$a\left(\mathbf{w}_n, \tilde{\mathbf{w}} - \frac{\Delta \mathbf{w}_n}{\Delta t}\right) + j(\tilde{\mathbf{w}}) - j\left(\frac{\Delta \mathbf{w}_n}{\Delta t}\right) \geq \left\langle \mathbf{l}(t), \tilde{\mathbf{w}} - \frac{\Delta \mathbf{w}_n}{\Delta t}\right\rangle \qquad \forall \tilde{\mathbf{w}} \in W. \tag{4.1}$$

Note that $j(\cdot)$, induced by a canonical matrix norm, is 1-homogeneous, and hence

$$j(\tfrac{\Delta \mathbf{w}_n}{\Delta t}) = \tfrac{j(\Delta \mathbf{w}_n)}{\Delta t}.$$

Furthermore, $a(\cdot, \cdot)$ is linear in the second argument, and therefore (4.1) is equal to

$$a(\mathbf{w}_n, \Delta t\tilde{\mathbf{w}} - \Delta \mathbf{w}_n) + j(\Delta t\tilde{\mathbf{w}}) - j(\Delta \mathbf{w}_n) \geq \langle \mathbf{l}(t), \Delta t\tilde{\mathbf{w}} - \Delta \mathbf{w}_n\rangle \qquad \forall \tilde{\mathbf{w}} \in W. \tag{4.2}$$

Since W is a vector space, we can neglect the factor Δt without changing the inequality. We arrive at

$$a\left(\mathbf{w}_n, \tilde{\mathbf{w}} - \Delta \mathbf{w}_n\right) + j(\tilde{\mathbf{w}}) - j\left(\Delta \mathbf{w}_n\right) \geq \langle \mathbf{l}(t), \tilde{\mathbf{w}} - \Delta \mathbf{w}_n\rangle \qquad \forall \tilde{\mathbf{w}} \in W. \tag{4.3}$$

Our next step is to transform this inequality into an equivalent minimization problem. To this end, we apply [8, Chapter II, Proposition 2.2] which states the following.

Theorem 4.1.1. *Let V a vector space and $F : V \to \mathbb{R}$ divided into $F = F_1 + F_2$, such that both F_1 and F_2 are lower semi-continuous and convex. Moreover, we assume that F_1 is Gâteaux-differentiable with differential F_1'. Then for $v \in V$ the two statements are equivalent:*

(i) v is a minimizer of F,

(ii) $\langle F_1'(v), \tilde{v} - v\rangle + F_2(\tilde{v}) - F_2(v) \geq 0 \qquad \forall \tilde{v} \in V.$

Since both $a(\mathbf{w}_n, \cdot)$ and $\langle \mathbf{l}(t), \cdot\rangle$ are linear, they will form the linear functional (ii) in the theorem above with $v = \Delta\mathbf{w}_n$. To achieve this, we have to enforce the direct dependence of $a(\mathbf{w}_n, \cdot)$ on $\Delta\mathbf{w}_n$. This is done by shifting $a(\cdot, \cdot)$ in the second component. Since $\mathbf{w}_{n-1}$ is assumed to be known at time t_n, we define

$$a_n(\mathbf{w}, \tilde{\mathbf{w}}) := a(\mathbf{w} + \mathbf{w}_{n-1}, \tilde{\mathbf{w}}). \tag{4.4}$$

If we use $\mathbf{w}_n = \mathbf{w}_{n-1} + \Delta\mathbf{w}_n$ we can rewrite (4.3) as

$$a_n\left(\Delta\mathbf{w}_n, \tilde{\mathbf{w}} - \Delta\mathbf{w}_n\right) + j(\tilde{\mathbf{w}}) - j\left(\Delta\mathbf{w}_n\right) \geq \langle \mathbf{l}(t), \tilde{\mathbf{w}} - \Delta\mathbf{w}_n\rangle \qquad \forall \tilde{\mathbf{w}} \in W. \tag{4.5}$$

Next, we want to define the objective functional from (i) in Theorem 4.1.1 and then we will compute its differential. Let us introduce a bilinear form $z : W \times W \to \mathbb{R}$ given by

$$z(\mathbf{w}, \tilde{\mathbf{w}}) := \int_\Omega \mathbf{H}(\boldsymbol{\epsilon}(\mathbf{u}) - \mathbf{p}) : (\boldsymbol{\epsilon}(\tilde{\mathbf{u}}) - \tilde{\mathbf{p}}) + k_1 \mathbf{p} : \tilde{\mathbf{p}}\, dx. \tag{4.6}$$

Note that $z(\cdot, \cdot)$ and $a(\cdot, \cdot)$ from (3.15) coincide if we do not consider isotropic hardening and hence, no terms of $\mathbf{g}$ and $\boldsymbol{\eta}$. For a fixed time step t_n let $L_n\colon W \to \mathbb{R}$ given by

$$L_n(\Delta\mathbf{w}_n) := \frac{1}{2} z(\Delta\mathbf{w}_n, \Delta\mathbf{w}_n) - \int_\Omega \mathbf{G}_n(\Delta\boldsymbol{\eta}_n)\, dx + j(\Delta\mathbf{w}_n) - \langle \mathbf{l}_n, \Delta\mathbf{w}_n\rangle + z(\mathbf{w}_{n-1}, \Delta\mathbf{w}_n)$$

with the isotropic hardening integral

$$\mathbf{G}_n(\boldsymbol{\eta}) = \int_0^{\boldsymbol{\eta} + \boldsymbol{\eta}_{n-1}} \mathbf{g}(s)\, ds, \tag{4.7}$$

which is to be understood pointwise in Ω. The differential with respect to the function $\boldsymbol{\eta}$ is given by

$$D\mathbf{G}_n(\boldsymbol{\eta}) = \mathbf{g}(\boldsymbol{\eta} + \boldsymbol{\eta}_{n-1}).$$

For Theorem 4.1.1, the objective function is given by $F = L_n$ with the splitting $F_2 = j$ and $F_1 = L_n - j$. The variables in our case are $v = \Delta w_n$ and $\tilde{v} = \tilde{w}$. In order to compute the Gâteaux-derivative, we briefly recall its definition.

Definition 4.1.2. *Let V a vector space, V' is dual, and $F\colon V \to \mathbb{R}$. We call $F'(u; v)$ the directional derivative of F at $u \in V$ in the direction $v \in V$ if*

$$F'(u; v) = \lim_{h \to 0^+} \frac{F(u + hv) - F(u)}{h},$$

provided the limit exists. Let $u \in V$. If there is a $u' \in V'$ with

$$\langle u', v\rangle = F'(u; v) \quad \forall v \in V$$

then we call u' the Gâteaux-differential of F at u and $\langle F'(u), \cdot\rangle := \langle u', \cdot\rangle$ the Gâteaux-derivative.

Let us compute the directional derivative of $F_1 = L_n - j$ at $\Delta\mathbf{w}_n$ in an arbitrary direction $\tilde{\mathbf{w}}$. We make use of the (bi-)linearity of the components and get

$$\begin{aligned}
&\lim_{h\to 0^+} \frac{F_1(\Delta\mathbf{w}_n + h\tilde{\mathbf{w}}) - F_1(\Delta\mathbf{w}_n)}{h} \\
&= z(\Delta\mathbf{w}_n, \tilde{\mathbf{w}}) - \int_\Omega \mathbf{g}(\boldsymbol{\eta}_{n-1} + \Delta\boldsymbol{\eta}_n)\tilde{\boldsymbol{\eta}}\,dx - \langle \mathbf{l}_n, \tilde{\mathbf{w}}\rangle + z(\mathbf{w}_{n-1}, \tilde{\mathbf{w}}) \\
&= a_n(\Delta\mathbf{w}_n, \tilde{\mathbf{w}}) - \langle \mathbf{l}_n, \tilde{\mathbf{w}}\rangle.
\end{aligned}$$

Note that the integral over the $\mathbf{g}$ term is included in the definition of the bilinear form $a(\cdot,\cdot)$ (3.15), and the shift by $\boldsymbol{\eta}_{n-1}$ is incorporated by $a_n(\cdot,\cdot)$ (4.4). Hence, it follows that

$$\langle F_1'(\Delta\mathbf{w}_n), \tilde{\mathbf{w}}\rangle = a_n(\Delta\mathbf{w}_n, \tilde{\mathbf{w}}) - \langle \mathbf{l}_n, \tilde{\mathbf{w}}\rangle.$$

Plugging this into (4.5) is precisely (ii) in Theorem 4.1.1 with $F_2 = j$. Hence, have the equivalent formulation of the small-strain primal plasticity problem with discretized time at t_n given by

$$\boxed{\text{Minimize} \quad L_n(\Delta\mathbf{w}_n)} \tag{4.8}$$

among all $\Delta\mathbf{w} \in W$. The next $\mathbf{w}_n$ is given by applying the increment to the previous time step

$$\mathbf{w}_n := \mathbf{w}_{n-1} + \Delta\mathbf{w}_n.$$

Lemma 4.1.3. *If we have both linear kinematic hardening with $k_1 > 0$ and isotropic hardening $\mathbf{g}$ with the assumptions from Section 2.6.2, then the functional L_n is strictly convex and coercive.*

Proof. Since $\mathbf{g}(0) = 0$ (2.14), $\mathbf{g} \leq 0$ and strongly monotone decreasing (pointwise), we can conclude that $-\mathbf{G}$ (4.7) is strictly convex and coercive. The nonsmooth component j inherits convexity and coercivity directly from the matrix norm. The linear terms of L_n have no impact on the claim. Hence, the last term of interest is $\frac{1}{2}z(\mathbf{w}, \mathbf{w})$ from (4.6). This is indeed V-elliptic [16, Lemma II.7.2], i.e., there is an $\alpha > 0$ with

$$z(\mathbf{w}, \mathbf{w}) \geq \alpha \left\|\mathbf{w}\right\|_V^2.$$

Since z is symmetric and bilinear, it follows that it is also strictly convex and coercive. □

Remark 4.1.4. *In the case of no isotropic hardening, we eliminate $\mathbf{g}$ and $\boldsymbol{\eta}$ from our equations and have a smaller functional space $W = V \times Q_0$ (see [30]). The lemma above still holds ins this setting.*

From the Lemma 4.1.3, we can directly conclude the theorem below.

Theorem 4.1.5. *The minimization problem* (4.8) *has a unique solution.*

4.2. Eliminating Internal Variables

We state the time discrete primal problem of small-strain plasticity in the form of a minimization problem in the increment of the unknowns $\Delta\mathbf{w}_n = (\Delta\mathbf{u}_n, \Delta\mathbf{p}_n, \Delta\boldsymbol{\eta}_n)$ at time t_n. From Section 3.2, we know that $\|\Delta\mathbf{p}_n\| \leq \Delta\boldsymbol{\eta}_n$ for the minimizer of L_n (with $\|\cdot\|$ either the Frobenius or the spectral norm). As it turns out, the connection is even stronger in the time-discrete setting.

Theorem 4.2.1. *Let* $\Delta\mathbf{w} = (\Delta\mathbf{u}, \Delta\mathbf{p}, \Delta\boldsymbol{\eta})$ *minimize* L_n. *Then*

$$\|\Delta\mathbf{p}\| = \Delta\boldsymbol{\eta}$$

holds almost everywhere.

Proof. The proof follows the argumentation of [2, Example 5]. If $\Delta\mathbf{w}$ is the minimizer of L_n, then $\Delta\boldsymbol{\eta}$ is the minimizer for fixed $(\Delta\mathbf{u}, \Delta\mathbf{p})$. The contribution of $\Delta\boldsymbol{\eta}$ in L_n is only given by the term $\int_\Omega -\mathbf{G}_n(\Delta\boldsymbol{\eta})\,dx$ with the constraint $\Delta\boldsymbol{\eta} \geq \|\Delta\mathbf{p}\|$. Since $-\mathbf{G}_n$ is monotone increasing for $\Delta\boldsymbol{\eta} \geq 0$, the minimizer obtains the smallest admissible value if $\Delta\boldsymbol{\eta} = \|\Delta\mathbf{p}\|$ almost everywhere. □

Theorem 4.2.1 allows us to replace

$$(\mathbf{u}_n, \mathbf{p}_n, \boldsymbol{\eta}_n) \coloneqq (\mathbf{u}_{n-1}, \mathbf{p}_{n-1}, \boldsymbol{\eta}_{n-1}) + \arg\min L_n(\Delta\mathbf{u}, \Delta\mathbf{p}, \Delta\boldsymbol{\eta})$$

by the reduced step

$$\begin{aligned} (\mathbf{u}_n, \mathbf{p}_n) &\coloneqq (\mathbf{u}_{n-1}, \mathbf{p}_{n-1}) + \arg\min \tilde{L}_n(\Delta\mathbf{u}, \Delta\mathbf{p}) \\ \boldsymbol{\eta}_n &\coloneqq \boldsymbol{\eta}_{n-1} + \|\Delta\mathbf{p}_n\| \end{aligned} \tag{4.9}$$

for the reduced functional

$$\tilde{L}_n(\Delta\mathbf{u}, \Delta\mathbf{p}) \coloneqq L_n(\Delta\mathbf{u}, \Delta\mathbf{p}, \|\Delta\mathbf{p}\|).$$

Tracking the iterates $\boldsymbol{\eta}_n$ cannot be discarded, since $\boldsymbol{\eta}_n$ appears in the next increment functional L_{n+1}, see (4.7). Eliminating $\boldsymbol{\eta}$ does not lead to a minimization problem that is more difficult than the original one.

Lemma 4.2.2. *Let* L_n *be strictly convex and coercive. Then* $\tilde{L}_n$ *is also strictly convex and coercive.*

Proof. Since $-\mathbf{G}_n$ is strictly convex and monotone increasing for positive values and the spectral norm is also convex, the composition $-\mathbf{G}_n(\|\cdot\|)$ is convex [1, Proposition 8.21]. The strict convexity of $\tilde{L}_n$ is preserved since the $z(\cdot,\cdot)$ term (4.6) in L_n is strictly convex (see the proof of Lemma 4.1.3) and independent of $\Delta\boldsymbol{\eta}$. The coercivity follows directly from the coercivity of L_n. □

4.3. Space Discretization

We are interested in computing one increment step

$$\mathbf{w}_n := \mathbf{w}_{n-1} + \Delta\mathbf{w}_n$$

for a fixed point in time t_n. The isotropic variable $\Delta\boldsymbol{\eta}_n$ is eliminated. For sake of simplicity we abuse the notation from before and write from now on

$$\boxed{\mathbf{w} := (\Delta\mathbf{u}_n, \Delta\mathbf{p}_n) \in W := V \times Q_0}$$

for the unknown of our minimization problem.

A standard finite element discretization is used. Therefore, let $\mathcal{T}$ be a triangulation of the initial domain Ω.

Displacement The function space of displacements $\mathbf{u}_n$

$$V = \left(H^1_\Gamma(\Omega)\right)^d$$

is approximated by first-order Lagrangian elements with values in $\mathbb{R}^d$. A basis of the displacement space is given by the nodal basis $\{\boldsymbol{\phi}_i(x)\}_{i=1,\dots,N_u}$ on N_u grid nodes with the Lagrangian property

$$\boldsymbol{\phi}_i(x_k) = \delta_{ik}$$

for the grid nodes $x_k \in \mathcal{T}$. Using this basis, the discrete approximation of the deformation field is

$$\mathbf{u}_h(x) := \sum_{i=1}^{N_u}\sum_{j=1}^{d} u_{ij}\boldsymbol{\phi}_i(x)\, e_j \tag{4.10}$$

with canonical unit vectors $e_j \in \mathbb{R}^d$. It is characterized by the real-valued coefficients $u_{ij} \in \mathbb{R}$.

Plastic Strain The plastic strain $\mathbf{p}$ is a function from

$$Q_0 = L^2(\Omega, \mathbb{S}^d_0).$$

The plastic strain $\mathbf{p}$ is approximated by zero-order (element-wise constant) elements with values in $\mathbb{S}^d_0$, since no derivatives of $\mathbf{p}$ are to be computed. Let N_p be the number of grid elements and let $\{\boldsymbol{\theta}_i\}_{i=1}^{N_p}$ denote the set of scalar zero-order basis functions with the property

$$\boldsymbol{\theta}_i|_{T_k} \equiv \delta_{ik}$$

for the grid elements T_k of $\mathcal{T}$.

For the plastic strain component a basis of the vector space $\mathbb{S}^d_0$ of trace-free symmetric matrices needs to be chosen. Different choices are possible but we stick to the choice done in [30] since it offers some numerical advantages as seen later.

Remark 4.3.1. *The dimension of* $\mathbb{S}_0^d$ *is given by*

$$d_p := \frac{d^2+d}{2} - 1.$$

For $d = 2$ the space $\mathbb{S}_0^2$ is only two-dimensional. An obvious basis choice is an orthonormal basis (with respect to the Frobenius inner product)

$$B_1 := \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \quad \text{and} \quad B_2 := \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}.$$

One can easily verify the orthonormality of this basis. For $d = 3$ the space $\mathbb{S}_0^3$ is by the lemma above 5-dimensional. Hence we need five basis matrices, for example

$$B_1 := \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}, \quad B_2 := \frac{1}{\sqrt{6}}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -2\end{pmatrix},$$

$$B_3 := \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}, \quad B_4 := \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix}, \quad B_5 := \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}.$$

This also forms an orthonormal basis of $\mathbb{S}_0^3$ under the Frobenius inner product. Therefore, for both $d = 2$ and $d = 3$, it also defines an isometry between $\mathbb{R}^{d_p}$ equipped with the Euclidean norm $\|\cdot\|_2$ and $\mathbb{S}_0^d$ with the Frobenius norm.

Lemma 4.3.2. *Let* $\{B_1, \ldots, B_{d_p}\}$ *an orthonormal basis of* S_0^d *with respect to the Frobenius inner product. Then, for each* $a \in \mathbb{R}^{d_p}$*, we have*

$$\left\| \sum_{j=1}^{d_p} a_j B_j \right\|_F = \|a\|_2 . \tag{4.11}$$

Proof. This follows directly form the orthonormality of the basis:

$$\left\| \sum_{j=1}^{d_p} a_j B_j \right\|_F^2 = \left\langle \sum_{i=1}^{d_p} a_i B_i, \sum_{j=1}^{d_p} a_j B_j \right\rangle = \sum_{i=1}^{d_p} \sum_{j=1}^{d_p} a_i a_j \delta_{ij} = \sum_{j=1}^{d_p} a_j^2 = \|a\|_2^2 .$$

□

Remark 4.3.3. *We have no equivalent expression of* $\left\| \sum_{j=1}^{d_p} a_j B_j \right\|_2$ *in terms of a norm of* a*.*

The global discrete plastic strain function is given by

$$\mathbf{p}_h(x) := \sum_{i=1}^{N_p} \sum_{j=1}^{d_p} p_{ij} \boldsymbol{\theta}_i(x)\, B_j.$$

It is characterized by scalar coefficients $p_{ij} \in \mathbb{R}$.

Although the isotropic hardening variable is $\boldsymbol{\eta}$ is eliminated from the minimization problem, we still need to track finite element approximations for the update step (4.9). For compatibility with the discrete plastic strain, we use zero-order elements with values in $\mathbb{R}$.

Assembly of the Algebraic Problem For a better handling we view the coefficients u_{ij} and p_{ij} not as matrices in $\mathbb{R}^{N_u \times d}$, or $\mathbb{R}^{N_p \times d_p}$, respectively, but rather as a (block) vector

$$u \in \mathbb{R}^{N_u d}, \quad p \in \mathbb{R}^{N_p d_p}$$

where we address by u_i the i-th block of size d, and by p_i the i-th block of size d_p. Then we can represent the plastic strain on the element T_i by

$$B(p_i) \coloneqq \sum_{j=1}^{d_p} p_{ij} B_j.$$

Sometimes we refer to *the* coefficient vector which is the composition of both u and p as

$$w \in \mathbb{R}^{N_u d + N_p d_p}, \quad w \coloneqq \begin{pmatrix} u \\ p \end{pmatrix}.$$

We assemble the algebraic representation (in terms of the coefficient vector w) of the reduced increment functional $\tilde{L}_n$. We abuse the notation and also denote the algebraic form of the objective functional by $\tilde{L}_n$. The quadratic part of $\tilde{L}_n$, namely the bilinear form z (4.6) is assembled into a matrix $A \in \mathbb{R}^{(N_u d + N_p d_p) \times (N_u d + N_p d_p)}$ with a 2×2 block structure

$$A = \begin{pmatrix} E & C \\ C^T & P \end{pmatrix}. \tag{4.12}$$

The $N_u d \times N_u d$ displacement part, partitioned into $d \times d$ blocks is given by

$$(E_{ij})_{kl} \coloneqq \int_\Omega \mathbf{H}(\boldsymbol{\epsilon}(\boldsymbol{\phi}_i e_k) : (\boldsymbol{\epsilon}(\boldsymbol{\phi}_j e_l)\, dx,$$

where $i, j = 1, \ldots, N_u$ select the nodal basis functions and $k, l = 1, \ldots, d$ the space directions. For the Hooke tensor $\mathbf{H}$ we only consider an isotropic material (2.6). Therefore we have

$$\mathbf{He} : \tilde{\mathbf{e}} = \lambda \operatorname{tr}(\mathbf{e}) \operatorname{tr}(\tilde{\mathbf{e}}) + 2\mu \mathbf{e} : \tilde{\mathbf{e}}. \tag{4.13}$$

The plastic strain part of size $N_p d_p \times N_p d_p$, divided into $d_p \times d_p$ blocks is block-diagonal, since the zero-order basis functions have disjoint support. The stiffness matrix on the element T_i is given by

$$(P_{ii})_{kl} \coloneqq \int_\Omega \mathbf{H}(\boldsymbol{\theta}_i B_k) : (\boldsymbol{\theta}_i B_l) + k_1 \boldsymbol{\theta}_i^2 B_k : B_l \, dx.$$

Moreover, all matrices B_i are trace-free, and hence we have with (4.13)

$$(P_{ii})_{kl} = \int_\Omega (2\mu + k_1) \boldsymbol{\theta}_i^2 \underbrace{B_k : B_l}_{=\delta_{kl}} \, dx. \tag{4.14}$$

Hence, P is diagonal, since the basis $\{B_k\}$ is orthonormal. The mixed displacement–plastic strain coupling matrix C of size $N_u d \times N_p d_p$ is defined by

$$(C_{ij})_{kl} \coloneqq \int_\Omega \mathbf{H}(\boldsymbol{\epsilon}(\boldsymbol{\phi}_i e_k) : (\boldsymbol{\theta}_i B_l)\, dx.$$

The two linear components of $\tilde{L}_n$ are assembled into a vector

$$b \in \mathbb{R}^{N_u d + N_p d_p}, \quad b =: \begin{pmatrix} b^u \\ b^p \end{pmatrix}. \tag{4.15}$$

The dissipation function $j(\cdot)$ and the isotropic hardening term $\mathbf{G}_n$ of $\tilde{L}_n$ depend on $\mathbf{p}$ only. We can simply rewrite the integral over Ω by a sum of the integrals over the grid elements T_i, since we have a zero-order basis. The resulting algebraic functional is

$$\tilde{L}_n(w) = \frac{1}{2} w^T A w - b^T w + \sum_{i=1}^{N_p} \int_{T_i} -\mathbf{G}_n(\|B(p_i)\|) + \sigma_0 \|B(p_i)\| \, dx.$$

Since p_i is constant on T_i, we can replace the integral by a factor ω_i, which is the volume of T_i, i.e.,

$$\omega_i := \int_\Omega \boldsymbol{\theta}_i \, dx = \int_{T_i} dx. \tag{4.16}$$

The final form of the algebraic minimization functional is

$$\boxed{L(w) = \frac{1}{2} w^T A w - b^T w + \sum_{i=1}^{N_p} \omega_i \varphi(p_i)} \tag{4.17}$$

with

$$\varphi : \mathbb{R}^{d_p} \to \mathbb{R}, \quad \varphi(p) := -\mathbf{G}_n(\|B(p)\|) + \sigma_0 \|B(p)\|. \tag{4.18}$$

Remarks 4.3.4.

(i) From (4.7) *and* (2.13)*, the explicit form of $-\mathbf{G}_n(\boldsymbol{\eta})$ is*

$$-\mathbf{G}_n(\boldsymbol{\eta}) = (k_\infty - \sigma_0)\left(\boldsymbol{\eta} + \boldsymbol{\eta}_{n-1} + \frac{1}{\beta} e^{-\beta(\boldsymbol{\eta} + \boldsymbol{\eta}_{n-1})}\right) + \frac{1}{2} k_2 (\boldsymbol{\eta} + \boldsymbol{\eta}_{n-1})^2.$$

(ii) From convexity and coercivity of both $-\mathbf{G}_n$ and the matrix norm, we conclude that φ is convex, coercive and lower semi-continuous.

(iii) However, due to the norm we also conclude that φ is not differentiable.

4.4. Excursus: Primal Plasticity with Gradient Regularization

The presented model of small-strain primal plasticity does not contain any gradient terms of the plastic strain. However, in certain applications this may be an important ingredient. For instance, plasticity problems at finite strains, as given in the second part of this thesis, require gradient regularization for the existence theory.

To demonstrate the implementation of gradient regularization on the easier problem of small-strain plasticity, a small overview is given in this section.

Formally, a gradient-regularized model has gradient terms of the internal variables in the global form $a(\cdot,\cdot)$ from (3.15). In [16, Sec. 7.3.2] the Gurtin model that contains a gradient of the plastic strain field $\mathbf{p}$ is given. If we consider a quadratic case with kinematic hardening only, the bilinear form in becomes

$$\tilde{a}(\mathbf{w},\tilde{\mathbf{w}}) := \int_\Omega \mathbf{H}(\boldsymbol{\epsilon}(\mathbf{u})-\mathbf{p}) : (\boldsymbol{\epsilon}(\tilde{\mathbf{u}})-\mathbf{p}) + (R\mathbf{p}) \;\;:\;\; \tilde{\mathbf{p}} + \mathbf{S}(\nabla\mathbf{p}) \vdots \nabla\tilde{\mathbf{p}}\, dx.$$

In that form $\mathbf{R}$ and $\mathbf{S}$ are forth-order or sixth-order tensors, respectively, and

$$A \vdots B := \sum_{i,j,k=1}^{d} A_{ijk}B_{ijk}$$

is the contraction over three indices. The simplest model is given by $\mathbf{R} \equiv k_1\,\mathrm{id}$, and $\mathbf{S} \equiv k_2\,\mathrm{id}$. In that case $\tilde{a}$ is a symmetric bilinear form. The gradient-regularized model is set in the function space

$$W := V \times \tilde{Q}_0.$$

We reuse the H^1-space V (3.7) for the displacement field $\mathbf{u}$, and the extend the L^2-space Q_0 (3.8) to an H^1-space

$$\tilde{Q}_0 := \{\mathbf{p} \in H^1(\Omega, \mathbb{S}^d_0),\ \mathbf{p} = 0 \text{ on } \Gamma\},$$

where Γ is a nonempty part of $\partial\Omega$. After a time discretization, as in Section 4.1, we arrive at a minimization problem

$$\text{Minimize} \quad L(\Delta\mathbf{w}_n) := \frac{1}{2}\tilde{a}(\Delta\mathbf{w}_n, \Delta\mathbf{w}_n) + j(\Delta\mathbf{w}_n) + \text{linear terms}$$

with the nonsmooth dissipation functional j from (3.16) that only depends on $\Delta\mathbf{p}_n$. The finite element discretization of $\Delta\mathbf{w}_n$ from Section 4.3 is not suitable here, since we need at least continuous finite elements to approximate the plastic strain field $\Delta\mathbf{p}_n$. Therefore, we use Lagrange elements for both the displacement field and the plastic strain. As a result, we arrive at an algebraic form similar to (4.17). The quadratic and linear parts are assembled into a matrix $\tilde{A}$ and a vector $\tilde{b}$, such that

$$L(w) := \frac{1}{2}w^T\tilde{A}w - \tilde{b}^T w + \int_\Omega D\left(\sum_{i=1}^{N_u}\sum_{j=1}^{d_p} p_{ij}\boldsymbol{\phi}(x)B_j\right) dx.$$

We use the first-order finite element basis $\boldsymbol{\phi}_i$ and the number of nodal points N_u from the displacement (4.10). Since the finite element functions are not piecewise constant we cannot simply write the dissipation term as a sum, in contrast to (4.17). Therefore, we approximate L by a lumped sum

$$\tilde{L}(w) := \frac{1}{2}w^T\tilde{A}w - \tilde{b}^T w + \int_\Omega \sum_{i=1}^{N_u}\boldsymbol{\phi}(x)\,dx \cdot D\left(\sum_{j=1}^{d_p} p_{ij}B_j\right).$$

Since the finite elements of the plastic strain are of first-order, they have no disjoint support. This leads to a sparse, but non-diagonal part $\tilde{P}$ in the 2×2 block structure

$$\tilde{A} = \begin{pmatrix} \tilde{E} & \tilde{C} \\ \tilde{C} & \tilde{P} \end{pmatrix},$$

cf (4.12) and (4.14).

Later in the numerical treatment of small-strain plasticity problems without gradient regularization, the diagonal structure of P will be advantageous, as seen in Section 5.3, especially (5.8). Nevertheless, we can adopt the numerical scheme to the sparse $\tilde{P}$. The detailed treatment of this case is presented in the second part of this thesis in Section 10.6.2.

5. Minimization Algorithms for Small-Strain Primal Plasticity

The algebraic minimization problem (4.8) is nonsmooth. Hence, special solvers are needed to find the minimizer numerically. In this section, two candidates are presented.

5.1. Predictor–Corrector Methods

In engineering, the classical approach to solve a minimization problem of the kind (4.17) is to consider the two main components u and p of the solution separately and minimize the objective functional alternating between these components [16, 31]. Minimization with respect to the displacement field is usually called a *predictor* step, while minimization with respect to the plastic strain is called the *corrector*. For better performance, the predictor step can have additional information of the plastic strain included. Therefore, we consider different predictor strategies.

In the most simple case the predictor step is the so-called elastic predictor and minimizes the objective functional only with respect to the displacements. All plastic strain components are fixed in this stage. Since all nonlinear and nonsmooth terms depend only on plastic strains, the resulting functional is quadratic. This can be minimized by solving a linear system.

The result of the elastic predictor needs most likely a correction of the plastic strain, since a purely elastic result was computed. This correction is called the corrector step. In this step we minimize the objective functional with respect to the plastic strain. In the corrector step the displacement is fixed. Sure, from occurring nonlinearities and non-differentiabilities, this part is more involved, but we can make use of the special structure of the problem (4.17). Since the coupling matrix P from (4.12) is block-diagonal, and the dissipation is given by a sum of the block-coefficients p_i, the blocks of p are independent of each other. Hence, minimization of the plastic strain can be done independently for each of the blocks.

For each block component p_i the minimization problem is of the form

$$\tilde{L}(p_i) := \frac{1}{2} p_i^T P_{ii} p_i - r_i^T p_i + \varphi_i(p_i)$$

for a linear part r_i resulting form the restriction to the plastic strain component. How the minimization of this nonsmooth problem is done efficiently is given in detail in Section 5.3.

The beauty of these predictor–corrector methods is their simplicity. In both steps, the minimization problems are easy to implement and analyze. Nevertheless, up to the knowledge of the author, no global convergence results of this alternating componentwise minimization strategy exists. Moreover, by decoupling u and p in the predictor and corrector steps, the method turns out to be not as efficient as other methods that consider $w = (u, p)$ as a whole.

In order to improve this simple predictor–corrector scheme, various modifications, especially for the elastic predictor step, are found in literature. The more sophisticated predictor–corrector approach from [16, Section 12.2] uses a tangent predictor. This means that we enrich the elastic predictor from before with second order information of the plastic strain. To this end, at iteration step k, we identify the blocks p_i^k in which the dissipation functional $\varphi(p_i^k)$ is smooth, and replace $\varphi = \varphi(p_i^k + \delta p)$ by

$$\bar{\varphi}_i^k(\delta p) := \begin{cases} \varphi(p_i^k) + \nabla\varphi(p_i^k)^T\delta p + \frac{1}{2}\delta p^T\nabla^2\varphi(p_i^k)\delta p & \text{if } \varphi \text{ is smooth at } p_i, \\ 0 & \text{else.} \end{cases}$$

For the quadratic part of $L(w)$ from (4.17) we also formulate the second Taylor expansion in w^k. The new quadratic increment objective functional is

$$\bar{L}^k(\delta w) := \frac{1}{2}\delta w^T A \delta w - (b - Aw^k)^T\delta w + \sum_{i=1}^{N_p}\omega_i\bar{\varphi}_i^k(\delta p_i).$$

Minimizing $\bar{L}^k(\delta w)$ is equal to computing a Newton-like step on L. Hence the name tangent predictor. The correction step, which corrects all plastic blocks (including the nonsmooth blocks) is the same as before.

Convergence in the functional value can be shown [16] if the tangent predictor does not create an energy increase. This can be guaranteed by applying a line search to the predictor. To this end, after each predictor step with $\bar{L}^k$, let δw^k denote the unique minimizer of the tangent predictor step. Instead of simply updating $w^k \leftarrow w^k + \delta w^k$, we compute a $\gamma^k > 0$, such that

$$L(w^k + \gamma^k\delta w^k) \leq L(w^k)$$

and update $w^k \leftarrow w^k + \gamma^k\delta w^k$ instead. This modified step is called consistent tangent predictor. Still this algorithm has two downsides:

- No global convergence proof in the argument w is found in literature.
- The quadratic predictor functionals $\bar{L}^k$ need to be solved exactly.

Therefore, to overcome these problems, the TNNMG algorithm is presented in the next section.

5.2. The TNNMG Algorithm

The Truncated Nonsmooth Newton Multigrid Method (TNNMG) is a generic method for solving block-separable, convex minimization problems. It was introduced in [11] and works well for small-strain plasticity problems with smooth elastic regions [30]. In this thesis the case of nonsmooth elastic regions, induced by the Tresca yield function (2.9), is also covered. TNNMG extracts second order information from the objective functional and tries to perform a Newton-like iteration step

$$w \leftarrow w + L''(w)^{-1}L'(w)$$

on the whole space of unknowns in order to approach the minimizer. Since L is not differentiable, we cannot compute gradients and hessians. Therefore, we need a modified Newton step.

Newton methods are generally good if we are already close to the minimizer, but we have poor performance (or even no convergence) if we are far from the minimizer. TNNMG is able to overcome this problem and we can even provide a global convergence result, as seen later. In order to apply TNNMG to our minimization problem (4.8), we start by quickly describing the general procedure.

TNNMG assumes a block-separable structure of the problem. For this we need two properties of the problem. First, the solution space is partitioned into orthogonal subspaces

$$\mathbb{R}^N = \prod_{i=1}^{m} \mathbb{R}^{N_i}, \tag{5.1}$$

where, $N_i \geq 0$ and $\sum_{i=1}^{m} N_i = N$. In order to address the i-th subspace we use the canonical restriction operator

$$R_i : \mathbb{R}^N \to \mathbb{R}^{N_i}.$$

We call W_i the subspace of $\mathbb{R}^N$ that corresponds to the i-th block $\mathbb{R}^{N_i}$. This simply means that all elements of W_i have zeroes at entries not belonging to the i-th block. The second property we demand is that the objective functional is separated in the nonsmooth parts according to the partition above. This means that our problem takes the form

$$L(w) = J_0(w) + \sum_{i=1}^{m} \varphi_i(R_i w) \tag{5.2}$$

with a “smooth” part J_0 and (possibly) nonsmooth parts. In detail, we demand that $J_0 : \mathbb{R}^N \to \mathbb{R}$ is twice differentiable. The separated parts $\varphi_i : \mathbb{R}^{N_i} \to \mathbb{R} \cup \{\infty\}$ are supposed to be proper, convex and lower semi-continuous, but not necessarily differentiable. All in all, the objective functional L has to be strictly convex and coercive to ensure the existence and uniqueness of a solution.

5.2.1. TNNMG Stages

TNNMG is an iterative solver. Let ν be the iteration counter and $w^\nu \in \mathbb{R}^N$ the current iterate. In order to get the next iterate $w^{\nu+1}$, we perform a TNNMG step, that consists of four major stages:

(1) Nonlinear presmoothing

(2) Inexact truncated linear correction

(3) Projection

(4) Line search

(1) Nonlinear Presmoothing At this first stage we perform the following steps

1. Set $\tilde{w}^0 = w^\nu$
2. For $i = 1, \dots, m$ do:

$$\text{compute } \tilde{w}^i = \underset{\tilde{w} \in \tilde{w}^{i-1} + W_i}{\arg\min} \; L(\tilde{w}) \tag{5.3}$$

3. Set $w^{\nu+\frac{1}{2}} := \tilde{w}^m$

Line 2 iterates through all subspaces W_i, defined by the partition of the solution space (5.1). The minimization problems of line 2 are low-dimensional. Since L is strictly convex and coercive, each of these subproblems always has a unique solution.

In the convergence proof given later it turns out this first stage of TNNMG alone is enough to achieve global convergence, and the other stages are used for acceleration.

(2) Inexact Truncated Linear Correction At this stage the Newton-like step is performed on L. It can be described as follows.

1. Determine a subspace $W_\nu \subset \mathbb{R}^N$, such that $L|_{W_\nu}$ is twice differentiable at $w^{\nu+\frac{1}{2}}$
2. Compute an inexact Newton correction $c^\nu \in W_\nu$

$$c^\nu \approx -\left(L''(w^{\nu+\frac{1}{2}})|_{W_\nu \times W_\nu}\right)^{-1} L'(w^{\nu+\frac{1}{2}})|_{W_\nu} \tag{5.4}$$

We will try to find a maximal subspace W_ν in which the functional L is indeed twice differentiable. Usually, we can easily determine which degrees of freedom of the solution space are currently “stuck” at a non-differentiable position and simply exclude them from the subspace W_ν.

Inexactness of the Newton step is understood as follows: Solving the linear correction problem (5.4) can be done directly with numerical exactness, or inexactly by using iterative approximation schemes. In the latter case we will use a geometric multigrid step to compute an approximation (5.4). Details about such a multigrid step are presented in Appendix A.2.

(3) Projection The domain of L (the subset of $\mathbb{R}^N$ where $L < \infty$) may be a real subset of $\mathbb{R}^N$. For instance, in the dissipation functions (3.6) and (3.5) we have the constraint $\eta \leq \|p\|$ if we do not eliminate η from the system. The Newton correction added to the current iterate $w^{\nu+\frac{1}{2}} + c^\nu$ may point outside the domain of L. Hence, we need to project into the domain of L.

1. Compute the projection $c^\nu_{\text{pr}} = P_{\operatorname{dom} L - w^{\nu+\frac{1}{2}}}(c^\nu)$, i.e., compute c^ν_{pr} such that $w^{\nu+\frac{1}{2}} + c^\nu_{\text{pr}}$ is closest to $w^{\nu+\frac{1}{2}} + c^\nu$ in $\operatorname{dom} L$ in a given norm.

The norm used in the projection step is usually the Euclidean norm. Nevertheless, for our reduced problem (4.17) with eliminated η, this stage can be skipped.

(4) Line Search Finally, we further compute a suitable step size of the projected Newton correction c^ν_{pr}, such that we have no increase in the value of L.

1. Compute $\rho^\nu \in [0, \infty)$, such that $L(w^{\nu+\frac{1}{2}} + \rho^\nu c^\nu_{pr}) \leq L(w^{\nu+\frac{1}{2}})$
2. Set $w^{\nu+1} = w^{\nu+\frac{1}{2}} + \rho^\nu c^\nu_{pr}$

This can be computed quickly as it is just a one-dimensional optimization problem with a strictly convex L.

5.2.2. Convergence of TNNMG

In summary, repeated application of all four stages of TNNMG leads to a sequence $(w^\nu)_{\nu\in\mathbb{N}}$ of iterates. We have a global convergence result for this method.

Theorem 5.2.1. *Let $L\colon \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ be strictly convex, coercive, proper, lower semi-continuous, and block-separably nonsmooth. Then the sequence w^ν produced by the TNNMG algorithm will converge to the unique minimizer of L.*

This is the convergence statement as it applies to the algorithm discussed in this manuscript. The proof in [11] actually holds under weaker assumptions on L. It also allows for the local minimization problems (5.3) to be solved inexactly under certain conditions. This can save additional run-time when solving (5.3) exactly is expensive (see, e.g., [10]), but will not be pursued here.

Remarks 5.2.2. *The TNNMG algorithm is closely related to the consistent tangent predictor–corrector (ctPC) approach from Section 5.1. The differences are :*

- *TNNMG loops over all degrees of freedom in the nonlinear smoothing stage.*
- *The linear correction stage is not necessarily solved exactly in TNNMG.*
- *ctPC cannot handle restricted domains of L.*

Otherwise, the general procedure, including the line search in the end are basically the same. The predictor step of ctPC corresponds to the last three stages of TNNMG while the corrector step is closely related to the first TNNMG stage.

5.3. Application to Primal Small-Strain Plasticity Problems

Applying the TNNMG algorithm to a particular problem now means constructing a decomposition (5.1) such that the objective functional takes the block-separable form (5.2), and specializing the four steps for the particular functional. In this paper the subspaces W_i are chosen as follows. Let N_u denote the number of grid vertices and N_p the number of grid elements. For the displacement coefficients we construct N_u blocks of size d, i.e., $N_i = d$ for $i = 1, \dots, N_u$. For the plastic strain we construct similarly N_p blocks of size d_p, i.e., $N_i = d_p$ for $i = N_u + 1, \dots, N_u + N_p$. Using this partition each plastic strain coefficient block p_i is in its own subspace W_i in terms of (5.1). Therefore, the nonsmooth parts are separated in the objective function L from (4.17).

The first stage of a TNNMG iteration consists of solving minimization problems in local displacement and plastic strain subspaces. While TNNMG convergence theory allows for certain forms of inexact smoothers [11], we show here how the local minimization problems for both von Mises and Tresca dissipation increment functionals can be solved numerically exactly. For the occurring matrices in the following recall the definitions at (4.12).

5.3.1. Minimizing in Displacement Subspaces

In case of minimization for a displacement tuple $u_i \in \mathbb{R}^d$, one has to minimize a quadratic functional

$$L_i^{\text{disp}}(u_i) := \frac{1}{2} u_i^T E_{ii} u_i - r_i^T u_i + \text{const}, \tag{5.5}$$

with constant terms depending on the plastic strain. The linear term $r_i \in \mathbb{R}^d$ is

$$r_i := b_i^u - \sum_{i \neq j} E_{ij} u_j - \sum_{i \neq j} C_{ij} p_j,$$

with b_i^u from (4.15). The functional (5.5) has a unique minimizer u_i^*, because it the restriction of a strictly convex and coercive functional. Hence, the matrix E_{ii} is positive definite and therefore invertible. Since E_{ii} is symmetric, the minimizer u_i^* can be computed by solving the linear system

$$E_{ii} u_i^* = r_i.$$

Since this is only a $d \times d$ system, we can it directly without drawbacks in runtime or stability.

5.3.2. Minimizing in Plastic Strain Subspaces

Minimizing in a plastic strain subspaces is more difficult than in the displacement subspaces. The main problem is in the nonsmooth dissipation and hardening functional

$$\varphi_i(p_i) := -\mathbf{G}_n(\|B(p_i)\|) + \sigma_0 \|B(p_i)\|$$

derived in (4.18). The restricted functional for the i-th plastic strain subspace is

$$L_i^{\text{plast}}(p_i) := \frac{1}{2} p_i^T P_{ii} p_i - r_i^T p_i + \omega_i \varphi_i(p_i) + \text{const}, \tag{5.6}$$

with constant terms depending on the displacement, and $\omega_i := \int_\Omega \boldsymbol{\theta}_i \, dx$. Since P is block-diagonal, the linear term r_i here is just

$$r_i := b_i^p - \sum_{i \neq j} C_{ji} u_j \tag{5.7}$$

with b_i^p taken from (4.15). Recall (4.13), i.e., that we have

$$p_i^T P_{ii} p_i = \omega_i (2\mu + k_1) \|p_i\|_2^2 . \tag{5.8}$$

The local functional L_i^{plast} is continuous, strictly convex and coercive, and thus a unique minimizer $p_i^* \in \mathbb{R}^{d_p}$ exists. For the further treatment of the local minimizers, we distinguish between von Mises and the Tresca dissipation function.

Von Mises Dissipation

In this section we consider $\|\cdot\| = \|\cdot\|_F$. For better readability we use the isometry (4.11) and identify

$$P^* := B(p_i^*), \quad R := B(r_i) \in \mathbb{S}_0^d.$$

Lemma 5.3.1. *If we use von Mises dissipation, the minimizer of* (5.6) *is* $P^* = 0$ *if and only if*

$$\|r_i\|_2 = \|R\|_F \le \sigma_0 - g(\eta_{n-1}).$$

Proof. We show that each direction $S \in \mathbb{S}^d$ is an ascending direction in P^*. Recall that $DG(0) = g(\eta_{n-1})$ from (4.7). Moreover, from the positive 1-homogeneity of the norm we have

$$\nabla \|\cdot\|_F (0)[S] = \|S\|_F .$$

The directional derivative is then given by

$$\nabla L_i^{\text{plast}}(0)[S] = -R : S + (\sigma_0 - g(\eta_{n-1})) \|S\|_F \ge - \|R\|_F \|S\|_F + (\sigma_0 - g(\eta_{n-1})) \|S\|_F \ge 0.$$

If $\|R\|_F > \sigma_0 - g(\eta_{n-1})$, we have $\nabla \|\cdot\|_F (0)[R] < 0$ and hence, $P^* \neq 0$. □

From Lemma 5.3.1, we can determine a priori whether the minimizer is zero or not. Hence, let us characterize the minimizer for the case that it is not zero.

Lemma 5.3.2. *If* $\|R\|_F > \sigma_0 - g(\eta_{n-1})$*, the functional* L_i^{plast} *is smooth around the minimizer* $P^* \neq 0$*. Then, the minimizer satisfies*

$$\begin{aligned} 0 = \frac{1}{\omega_i} \nabla L_i^{plast}(P^*) &= (2\mu + k_1) P^* - R + \Big[\sigma_0 - g\big(\|P^*\|_F + \eta_{n-1}\big)\Big] \frac{P^*}{\|P^*\|_F} \\ &= \left(2\mu + k_1 + \frac{\sigma_0 - g(\|P^*\|_F + \eta_{n-1})}{\|P^*\|_F} \right) P^* - R. \end{aligned}$$

Since in the lemma above the term in parenthesis is positive, we can conclude that $P^* = \gamma R$ for a positive γ. To determine γ, we have to solve the nonlinear equation

$$(2\mu + k_1)\gamma + \frac{\sigma_0 - g(\gamma \left\|R\right\|_F + \eta_{n-1})}{\left\|R\right\|_F} = 1. \tag{5.9}$$

We note that the left hand side is strictly monotone increasing in γ. Moreover, for $\gamma = 0$, the left hand side is smaller than one, since $\left\|R\right\|_F > \sigma_0 - g(\eta_{n-1})$. On the other hand, for $\gamma \to \infty$, the left and side goes also to infinity. We can conclude that there is a unique γ since the left hand side is continuous in γ.

The reduction to a one dimensional problem can be interpreted as a closest point projection onto the elastic region, as described in, e.g., [31]. Since in the von Mises case this region is point-symmetric, the projection direction is given directly by the direction to the center of the elastic region. In the case of linear isotropic hardening ($g(\eta) = -k_2\eta$), we can determine γ directly by

$$(2\mu + k_1 + k_2) \left\|R\right\|_F \gamma = \left\|R\right\|_F - k_2\eta_{n-1} \;\Rightarrow\; \gamma = \frac{\left\|R\right\|_F - k_2\eta_{n-1}}{(2\mu + k_1 + k_2) \left\|R\right\|_F}.$$

This also covers the case of no isotropic hardening ($k_2 = 0$) and no kinematic hardening ($k_1 = 0$).

The general nonlinear case (5.9) is solved by a damped Newton method.

Tresca Dissipation

The case of Tresca dissipation is more involved, since the dissipation function includes the spectral norm. We consider the relevant space dimensions $d = 2$ and $d = 3$ separately.

For the case $d = 2$, note that for any $p \in \mathbb{S}_0^2$ with eigenvalues $\lambda_1(p), \lambda_2(p) = -\lambda_1(p)$ we have

$$\left\|p\right\|_2 = \max\left\{|\lambda_1(p)|, |\lambda_2(p)|\right\} = |\lambda_1(p)| = \frac{1}{\sqrt{2}}\sqrt{\lambda_1(p)^2 + \lambda_2(p)^2} = \frac{1}{\sqrt{2}} \left\|p\right\|_F , \tag{5.10}$$

and therefore the Tresca dissipation functional is proportional to the von Mises dissipation. Hence, it is already covered in the previous section. For the case $d = 3$ the five-dimensional problem can be reduced to two dimensions, because the eigenvectors of $B(p_i^*)$ can be computed from r_i alone. This has been inspired from a similar result in [28].

Lemma 5.3.3. *Assume that the material is isotropic* (2.7)*. Let $r_i \in \mathbb{R}^5$ be the linear factor of* (5.7) *and $R := B(r_i) \in \mathbb{S}_0^3$ the corresponding trace-free symmetric matrix. Then the matrix representation of the minimizer $B^* := B(p_i^*)$ has the same eigenvectors as R.*

Proof. Using that $B : \mathbb{R}^{d_p} \to \mathbb{S}_0^d$ is an isometry, we can express L_i^{plast} in terms of matrices

$$\hat{L}_i^{\text{plast}} : \mathbb{S}_0^3 \to \mathbb{R} \cup \{\infty\}, \quad \hat{L}_i^{\text{plast}}(P) := \frac{\omega_i}{2}\mathbf{H}P : P - R : P + \omega_i\hat{\varphi}_i(P) + \text{const}, \tag{5.11}$$

where $\hat{\varphi}_i(P) \coloneqq \varphi_i(B^{-1}(P))$. If the Hooke tensor $\mathbf{H}$ describes an isotropic material the functional $\hat{L}_i^{\text{plast}}$ takes the form

$$\hat{L}_i^{\text{plast}}(P) = \frac{(2\mu + k_1)\omega_i}{2} P : P - R : P + \omega_i \hat{\varphi}_i(P) + \text{const}, \tag{5.12}$$

and its subdifferential is

$$\partial \hat{L}_i^{\text{plast}}(P) = (2\mu + k_1)\omega_i P - R + \omega_i \partial \hat{\varphi}_i(P). \tag{5.13}$$

Since $\hat{\varphi}_i$ only depends on the spectral norm, we can express it by a lower semi-continuous, convex scalar function

$$\bar{\varphi}_i : \mathbb{R} \to \mathbb{R},\ \bar{\varphi}_i(\nu) \coloneqq -\mathbf{G}_n(\nu) + \sigma_0 \nu, \tag{5.14}$$

such that

$$\hat{\varphi}_i(P) = \bar{\varphi}_i(\|P\|_2).$$

Then we can express the subdifferential by

$$\partial \hat{\varphi}_i(P) = \{\alpha\, \partial \left\|\cdot\right\|_2(P) \ :\ \alpha \in \partial \bar{\varphi}(\cdot)(\|P\|_2)\}.$$

See, e.g., [1] for a proof. Hence, each element of the subdifferential of $\hat{\varphi}$ is a scalar multiple of an element of the subdifferential of the spectral norm. Let $\lambda : \mathbb{S}^3 \to \mathbb{R}^3$, $\lambda(A) \coloneqq (\lambda_1(A), \lambda_2(A), \lambda_3(A))^T$ be the eigenvalue mapping. Then $(\left\|\cdot\right\|_\infty \circ \lambda)(A) = \|A\|_2$ holds. We can apply [1, Proposition 24.63] since $\left\|\cdot\right\|_\infty$ is symmetric, i.e., independent of the order of the arguments, and convex. We get

$$\partial \left\|\cdot\right\|_2(P) = \left\{ V \operatorname{diag}(y) V^T :\ y \in \partial \left\|\lambda(P)\right\|_\infty,\ V \in \mathrm{SO}(3) :\ P = V \operatorname{diag}(\lambda(P)) V^T \right\}.$$

If P is the minimizer of (5.12) we have $0 \in \partial \hat{L}_i^{\text{plast}}(P)$. Then, (5.13) yields

$$R \in (2\mu + k_1)\omega_i P + \omega_i \partial \hat{\varphi}_i(P).$$

Hence R has the same principal axes as P. ☐

With the eigenvectors of the minimizer P^* known, we can reformulate the minimization problem (5.11) in the three-dimensional space of eigenvalues $\lambda_1(P), \lambda_2(P), \lambda_3(P)$. As P^* is trace-free we can further eliminate $\lambda_3(P)$ by setting $\lambda_3(P) = -\lambda_1(P) - \lambda_2(P)$. To write the functional $\hat{L}$ in terms of two eigenvalues of P we compute

$$\|B(p)\|_2 = \max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\} = \max\{|\lambda_1|, |\lambda_2|, |\lambda_1 + \lambda_2|\}$$

and

$$\|B(p)\|_F = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2} = \sqrt{2\lambda_1^2 + 2\lambda_2^2 + 2\lambda_1\lambda_2}.$$

Let ν_i denote the eigenvalues of R and λ_i the eigenvalues of P. As P and R have the same principal axes the scalar product reduces to

$$P : R = \operatorname{tr}(P^T R) = \underbrace{(\nu_1 - \nu_2)}_{=:\tilde{\nu}_1} \lambda_1 + \underbrace{(\nu_2 - \nu_3)}_{=:\tilde{\nu}_2} \lambda_2.$$

Using the equations above, the resulting objective functional in two variables is given by

$$\begin{aligned}\tilde{L}_i(\lambda_1,\lambda_2) \coloneqq \omega_i &\left[\frac{2\mu+k_1}{2}(2\lambda_1^2+2\lambda_2^2+2\lambda_1\lambda_2)+\bar{\varphi}_i(m(\lambda_1,\lambda_2))\right] \\ &-\tilde{\nu}_1\lambda_1-\tilde{\nu}_2\lambda_2+\text{const},\end{aligned} \tag{5.15}$$

with

$$m(\lambda_1,\lambda_2) \coloneqq \max\{|\lambda_1|,|\lambda_2|,|\lambda_1+\lambda_2|\}. \tag{5.16}$$

The scalar function $\bar{\varphi}_i$ (5.14) is smooth The nondifferentiability of L_i^{plast} only originates from the function m. Nevertheless, $\tilde{L}_i$ is piecewise smooth on the decomposition shown in Figure 5.1.

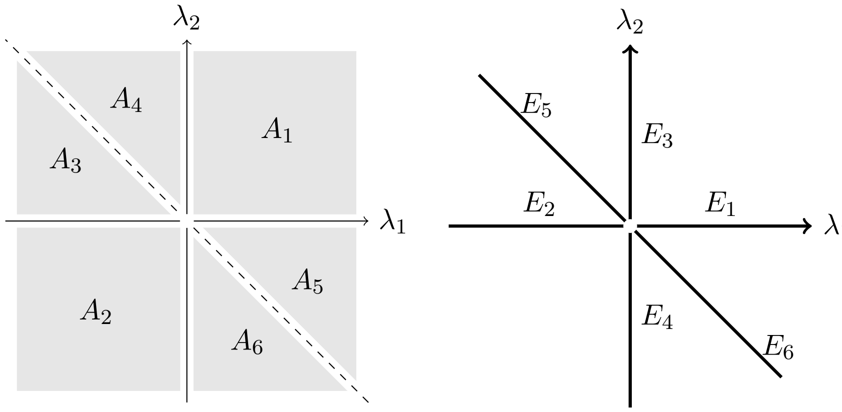


Figure 5.1.: Six open areas $A_1,\dots,A_6 \subset \mathbb{R}^2$ and six open edges $E_1,\dots,E_6 \subset \mathbb{R}^2$

On each of the six open areas $A_1,\dots,A_6$ and edges $E_1,\dots,E_6$ the maximum term can be expressed as

$$m(\lambda_1,\lambda_2)|_C = \alpha_1^C\lambda_1+\alpha_2^C\lambda_2 \tag{5.17}$$

with the following coefficients:

Case	A_1	A_2	A_3	A_4	A_5	A_6	E_1	E_2	E_3	E_4	E_5	E_6
α_1^C	1	-1	-1	0	1	0	1	-1	0	0	0	1
α_2^C	1	-1	0	1	0	-1	0	0	1	-1	1	0

The set $\mathcal{P} \coloneqq \{\{0\},A_1,\dots,A_6,E_1,\dots,E_6\}$ is a disjoint partition of $\mathbb{R}^2$. Since $\tilde{L}_i$ has a unique minimizer, the minimizer is contained in exactly one $C \in \mathcal{P}$. The zero point is a special case, because we can check a priori whether the minimizer is zero.

Lemma 5.3.4. *Consider Tresca dissipation and let $\tilde{L}_i$ be as in* (5.15) *with the linear factor $\tilde{\nu}=(\tilde{\nu}_1,\tilde{\nu}_2)^T$. If*

$$\tilde{\nu} \in \omega_i \cdot \partial\left[\bar{\varphi}_i(m(\cdot,\cdot))\right](0,0), \tag{5.18}$$

then $\lambda_1=\lambda_2=0$ is the unique minimizer of $\tilde{L}_i$.

Proof. $\tilde{L}_i$ is strictly convex and coercive. Therefore, we have existence of a unique minimizer. The subdifferential of $\tilde{L}$ at zero is

$$\partial \tilde{L}_i(0,0) = \omega_i \cdot \partial \left[\bar{\varphi}_i(m(\cdot,\cdot))\right](0,0) - \tilde{\nu}. \tag{5.19}$$

Moreover, $0 \in \partial \tilde{L}_i(0,0)$ is a sufficient condition for the minimizer. Hence, if $\tilde{\nu} \in \omega_i \cdot \partial \left[\bar{\varphi}_i(m(\cdot,\cdot))\right](0,0)$, the claim follows. ☐

With the lemma above, we can summarize the computation of the minimizer by the following procedure. We have to carefully choose the correct order of the 13 elements of the set $\mathcal{P} := \{\{0\}, A_1, \dots, A_6, E_1, \dots, E_6\}$. Therefore, we will consider the case zero, the open areas A_i, and the line segments E_i separately.
For $\tilde{L}_i$ as in (5.15), consider the following steps:

(1) Use Lemma 5.3.4) and check whether the minimizer is zero:
  - Check whether
  $$\tilde{\nu} \in \omega_i \cdot \partial \left[\bar{\varphi}_i(m(\cdot,\cdot))\right](0,0).$$
  - If it holds: **Return** $(0,0)$.

(2) For each $C \in \{A_1, \dots, A_6\}$:
  - Restrict $\tilde{L}_i$ to C using (5.17).
  - Extend $\tilde{L}_i|_C$ canonically to $\mathbb{R}^2$.
  - Compute the unique minimizer λ^C of the smooth extension.
  - If $\lambda^C \in C$: **Return** λ^C.

(3) Set $S := \emptyset$. For each $C \in \{E_1, \dots, E_6\}$:
  - Restrict $\tilde{L}_i$ to C using (5.17).
  - Extend $\tilde{L}_i|_C$ canonically to the infinite line in $\mathbb{R}^2$ containing C.
  - Compute the unique minimizer λ^C of the smooth 1D extension[1].
  - If $\lambda^C \in C$: Add λ^C to the set S.

(4) **Return** the optimal candidate $\lambda \in S$.

Remarks 5.3.5. *If we found a minimizer λ^C in* (2) *we can stop immediately and skip* (3) *and* (4), *since λ^C is contained in the interior of C, and the restriction $\tilde{L}_i|_C$ is strictly convex. In* (3) *multiple candidates may appear, as shown in Figure 5.2. Therefore we have to consider all 6 edges and choose the optimal point in the end.*

In the following we will show the explicit application of the algorithm above for certain scenarios of small-strain plasticity with hardening. The following characterization of $\partial m(0,0)$ follows directly from [1, Theorem 18.5].

---

[1] To be precise here: On the edges E_i the maximum functional m (5.17) depends only on one variable. Formally, we replace $\tilde{m}(\lambda) := m(\lambda,\lambda)|_C$ and compute the unique minimizer of $\tilde{m}$

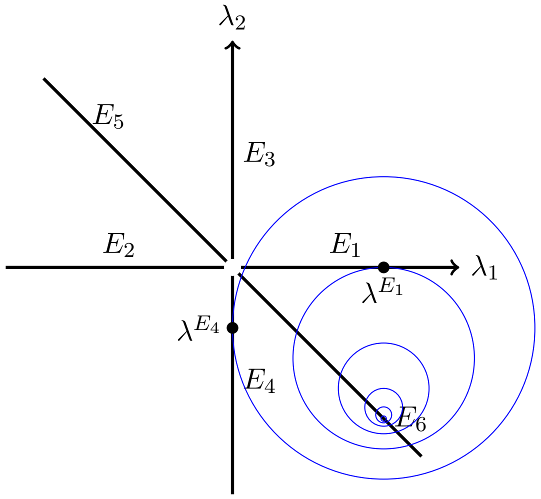


Figure 5.2.: Level sets of a convex function with global minimum on E_6 and local minima $\lambda^{E_1}, \lambda^{E_4}$ on the restrictions E_1 and E_4, respectively

Lemma 5.3.6. *For m from* (5.16)*, we have*

$$\partial m(0,0) = \operatorname{conv}\{\pm e_1, \pm e_2, \pm(e_1+e_2)\}$$

for the canonical unit vectors $e_1, e_2 \in \mathbb{R}^2$.

Example 5.3.7 (Linear hardening)**.** For the case of linear kinematic and linear isotropic hardening (2.12) the nonlinearity $\mathbf{G}_n$ is given by

$$-\mathbf{G}_n(\Delta\eta) = k_2\eta_{n-1}\Delta\eta + \frac{k_2}{2}\Delta\eta^2 + \text{const}$$

for a positive scalar $k_2 > 0$. After eliminating $\Delta\eta$ as described in Section 4.2, the objective functional reduces to

$$\begin{aligned}\tilde{L}(\lambda_1,\lambda_2) = \omega_i \Bigg[& \frac{2\mu+k_1}{2}(2\lambda_1^2 + 2\lambda_2^2 + 2\lambda_1\lambda_2) - \tilde{\nu}_1\lambda_1 - \tilde{\nu}_2\lambda_2 \\ & + \frac{k_2}{2} m(\lambda_1,\lambda_2)^2 \\ & + (k_2\eta_{n-1} + \sigma_0)\, m(\lambda_1,\lambda_2) \Bigg] + \text{const}.\end{aligned}$$

For step (1) of the algorithm above, use Lemma 5.3.6 and compute

$$\partial\left[\bar{\varphi}_i(m(\cdot,\cdot))\right](0,0) = (k_2\eta_{n-1} + \sigma_0)\operatorname{conv}\{\pm e_1, \pm e_2, \pm(e_1+e_2)\}.$$

By Lemma 5.3.4 the condition

$$\tilde{\nu} \in (k_2\eta_{n-1} + \sigma_0)\operatorname{conv}\{\pm e_1, \pm e_2, \pm(e_1+e_2)\}$$

holds if and only if the minimizer is at zero. If it is not, restrict $\tilde{L}$ to the 12 cases shown in Figure 5.1. For each of these cases, the problem is quadratic:

$$\begin{aligned}\tilde{L}_i(\lambda_1, \lambda_2)|_C = \omega_i \Big[&\frac{2\mu + k_1}{2}(2\lambda_1^2 + 2\lambda_2^2 + 2\lambda_1\lambda_2) + \frac{k_2}{2}(\alpha_1^C\lambda_1 + \alpha_2^C\lambda_2)^2 \\ &+ (k_2\eta_{n-1} + \sigma_0)(\alpha_1^C\lambda_1 + \alpha_2^C\lambda_2)\Big] \\ &- \tilde{\nu}_1\lambda_1 - \tilde{\nu}_2\lambda_2 + \text{const}\end{aligned}$$

with $C \in \{A_i, E_i \colon i \in \{1, \dots, 6\}\}$. Therefore, we can compute the steps (2)–(4) of the algorithm above directly.

Example 5.3.8 (Nonlinear saturated isotropic hardening)**.** If we extend the previous example by nonlinear saturated hardening as given in (2.13), the nonlinearity $\mathbf{G}_n$ takes the form

$$-\mathbf{G}_n(\Delta\eta) = k_2\eta_{n-1}\Delta\eta + \frac{k_2}{2}\Delta\eta^2 + (k_\infty - \sigma_0)\left(\Delta\eta + \frac{\theta_n}{\beta}e^{-\beta\Delta\eta}\right) + \text{const}$$

with positive scalars $k_\infty \geq \sigma_0, k_2, \beta$ and $\theta_n := e^{-\beta\eta_{n-1}}$. For using Lemma 5.3.4, we compute

$$\partial\left[\bar{\varphi}_i(m(\cdot,\cdot))\right](0,0) = \left(k_2\eta_{n-1} + \sigma_0 + (k_\infty - \sigma_0)(1 - \theta_n)\right)\, \mathrm{conv}\{\pm e_1, \pm e_2, \pm(e_1 + e_2)\}$$

and can easily check whether the minimizer is zero with Lemma 5.3.4. For the other cases, the restriction is given by

$$\begin{aligned}\tilde{L}(\lambda_1, \lambda_2)|_C = \omega_i \Big[&\frac{2\mu + k_1}{2}(2\lambda_1^2 + 2\lambda_2^2 + 2\lambda_1\lambda_2) + \frac{k_2}{2}(\alpha_1^C\lambda_1 + \alpha_2^C\lambda_2)^2 \\ &+ (k_\infty - \sigma_0)\frac{\theta_n}{\beta}e^{-\beta(\alpha_1^C\lambda_1 + \alpha_2^C\lambda_2)} + (k_2\eta_{n-1} + k_\infty)(\alpha_1^C\lambda_1 + \alpha_2^C\lambda_2)\Big] \\ &- \tilde{\nu}_1\lambda_1 - \tilde{\nu}_2\lambda_2 + \text{const}\end{aligned}$$

with $C \in \{A_i, E_i \;\; i \in \{1, \dots, 6\}\}$. In each case, the minimization problem is convex and can be solved by globally converging damped Newton steps. In practice, this converges in a few steps to the unique minimizer.

5.3.3. Determining Maximal Differentiable Subspaces

In the TNNMG linear correction stage, for the linear correction step (5.4), we have to determine the maximal subspace W_ν in which the functional is differentiable. For von Mises dissipation this is easy, since each component is non-differentiable only in the zero point.

However, the spectral norm of the Tresca dissipation has indeed non-trivial subspaces of non-differentiability. The spectral norm of symmetric matrices is differentiable if and only if the dominant eigenvalue (in terms of the absolute value) is unique. Only the relevant cases $d = 2, 3$ are considered here. Recall that for $d = 2$ Tresca dissipation is equal to von Mises dissipation (5.10). For $d = 3$, the eigenvalues of trace-free matrices satisfy $\lambda_1 + \lambda_2 + \lambda_3 = 0$. We can divide all configurations of the eigenvalues into three disjoint cases.

(i) one dominant eigenvalue $|\lambda_1| > \max\{|\lambda_2|, |\lambda_3|\}$

(ii) two dominant eigenvalues $\lambda_1 = -\lambda_2 \neq 0$, and $\lambda_3 = 0$

(iii) all eigenvalues are zero

In case (i), the spectral norm is differentiable. In case (ii) we conclude that the current plastic strain component P in matrix has the spectrum

$$\{\lambda, -\lambda, 0\}.$$

Since symmetric matrices have orthogonal eigenvectors there are two unit vectors $a, b \in \mathbb{R}^3$ with $a \perp b$ such that

$$P = \lambda(aa^T - bb^T).$$

Nevertheless, the spectral norm restricted to the one-dimensional space

$$\Big\{\gamma \cdot (aa^T - bb^T) : \ \gamma \in \mathbb{R}\Big\}$$

is differentiable in P. For the case (iii) there is no non-trivial subspace, in which the spectral norm is differentiable.

5.4. Efficient Implementation of Norm Derivatives

For a start, we need fast and robust methods to compute the value, gradient and hessian of the Frobenius norm or the spectral norm of trace-free symmetric matrices in order to compute the Newton-like linearization step (5.4). The basis of $\mathbb{S}^d_0$ was chosen such that for a $p \in \mathbb{R}^{d_p}$, which induces a trace-free and symmetric matrix $B = B(p) := \sum_{i=1}^{d_p} B_i p_i$, we always have the isometry $\|B\|_F = \|p\|_2$. Deriving the Frobenius norm is relatively easy and can be done directly by

$$\nabla \|B(\cdot)\|_F(p) = \frac{p}{\|B(p)\|_F}, \qquad \nabla^2 \|B(\cdot)\|_F(p) = \frac{\mathbf{I}}{2\,\|B(p)\|_F} - \frac{pp^T}{\|B(p)\|_F^3}. \tag{5.20}$$

This is well-defined as long as $p \neq 0$, which is the only non-differentiable point. Moreover, in two dimensions we further have the equivalence of von Mises and Tresca dissipation since $\|B(p)\|_F = \sqrt{2}\,\|B(p)\|_2$. Therefore, we can use the same formula above to obtain the gradient and the hessian of the spectral norm in p. In the three-dimensional case we actually need to compute the eigenvalues of $B(p)$ first.

Lemma 5.4.1. *For a symmetric, trace-free matrix $B \in \mathbb{S}_0^3$ the eigenvalues are explicitly given by*

$$\lambda_i(B) = s_i \sqrt{\frac{2}{3}} \, \|B\|_F \cos\left(\frac{1}{3} \arccos\left(- \det(B) \frac{\sqrt{54}}{\|B\|_F^3} \right) + r_i \right) \tag{5.21}$$

for $i = 1, 2, 3$ with $s = (-1, 1, -1)$ and $r = (\frac{\pi}{3}, 0, -\frac{\pi}{3})$.

Proof. We have

$$\|B\|_2 = \max\{|\lambda_1(B)|, |\lambda_2(B)|, |\lambda_3(B)|\} \tag{5.22}$$

for the eigenvalues $\lambda_1(B), \lambda_2(B), \lambda_3(B)$ of B. These values are the roots of characteristic polynomial

$$\mathcal{B}(\lambda, B) := \det(\lambda I - B) = \lambda^3 - \frac{1}{2} \|B\|_F^2 \lambda - \det(B) = 0, \tag{5.23}$$

The roots of this cubic equation can be computed directly by Cardano's method. This common result can be found in monographs like [7]. No second order terms appear in the characteristic polynomial since B is trace-free. This saves us the generally necessary transformation step of Cardano's method in order to achieve the reduced form

$$\lambda^3 + \alpha\lambda + \beta = 0$$

of the third order polynomial. ☐

Deriving the eigenvalues with respect to the matrix is possible if and only if the dominant eigenvalue is unique. To check this, we can compute the eigenvalues using the lemma above. It is not wise to directly attack the equation (5.21). It will lead to a lengthy chain rule application which is error-prone and difficult to implement. Instead, we use the implicit formula (5.23) for $B = B(p)$ and derive both sides by p to obtain

$$3\lambda^2 \nabla_p \lambda - \lambda p - \frac{1}{2} \|B(p)\|_F^2 \nabla_p \lambda - \nabla_p \det(B(p)) = 0,$$

where λ is the dominant eigenvalue. Recall that $\nabla_p \left(\frac{1}{2} \|B(p)\|_F^2\right) = \nabla_p \left(\frac{1}{2} \|p\|_2^2\right) = p$. Moreover, $\det(B(p))$ is a third-order polynomial in p and therefore easy to derive. Hence,

$$\nabla_p \lambda = \frac{\nabla_p \det(B(p)) + \lambda p}{3\lambda^2 - \frac{1}{2} \|B(p)\|_F^2}$$

is a direct formula for the gradient of the spectral norm. We have to switch the sign of the result, if the dominant eigenvalue is negative. For completeness, the denominator is never zero, since λ is the dominant eigenvalue and hence,

$$3\lambda^2 > \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = \|B\|_F^2 > \frac{1}{2} \|B\|_F^2 .$$

For the hessian of the dominant eigenvalue, we derive (5.23) once more to obtain

$$6\lambda \nabla_p \lambda (\nabla_p \lambda)^T + 3\lambda^2 \nabla_p^2 \lambda - 2\nabla_p \lambda p - \lambda I - \frac{1}{2} \|B(p)\|_F^2 \nabla_p^2 \lambda - \nabla_p^2 \det(B(p)) = 0,$$

which can be solved for $\nabla_p^2 \lambda$ as long as $3\lambda^2 \neq 0$. If λ is the unique dominant eigenvalue it cannot be zero. If the dominant eigenvalue is negative, we have to switch the sign of the hessian of the eigenvalue to get the hessian of the spectral norm.

Another numerically critical part is to stably determine whether the dissipation functionals (3.5) and (3.6) with the terms

$$\|B(p)\|_F \quad \text{and} \quad \|B(p)\|_2$$

are smooth around $p \in \mathbb{R}^{d_p}$. The case of the Frobenius norm is rather easy, since only $p = 0$ is a non-differentiable point. Numerically, the condition $p \neq 0$ is not stable, since for $p \to 0$ the hessian in (5.20) becomes infinite large. Therefore, a criterion $\|p\|_2 > 10^{-10}$ is used in practice to determine differentiability.

For the spectral norm differentiability is equivalent to the uniqueness of the dominant eigenvalue. From the zero trace of $B(p)$ we know that

$$\lambda \text{ is not uniquely dominant} \quad \Leftrightarrow \quad \operatorname{spectrum}(B(p)) = \{\lambda, 0, -\lambda\}.$$

We use this property by computing the eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ of $B(p)$ by Cardanos's method and implement the check whether $|\lambda_2| > 10^{-10}$ in order to determine differentiability numerically.

6. Numerical Results

In this section we test the implementation of the solution algorithms from Chapter 5 applied to the algebraic minimization problem (4.17).

Several test scenarios are considered. A first model choice is the dissipation function. In this part of the thesis, the von Mises dissipation (3.5) and the Tresca dissipation (3.6) are studied. Tests for the von Mises dissipation are performed in two and three dimensions. We restrict the tests for the Tresca dissipation to three dimensions, since in two dimensions both dissipation functions are equivalent, cf. Lemma 2.4.3 and (5.10).

The considered solvers for the resulting minimization problems are TNNMG and the predictor–corrector method with a consistent tangent predictor (ctPC). From Remarks 5.2.2 both algorithms are related, and therefore suitable for a comparison. While the algorithmic parameters of ctPC are fixed, we have some freedom in the application of TNNMG. Each TNNMG iteration uses one application of the local smoother, which is in this case an exact block Gauss–Seidel method, as presented in Section 5.3. We have to choose the solver for the linear correction steps (5.4). We will solve these problems inexactly with a single geometric multigrid step (Appendix A.2) for an inexact solution. On the coarsest grid level, the direct solver CHOLMOD [3] is used.

The material parameters, as well as the domain, boundary conditions and load are given for each test separately.

6.1. Von Mises Dissipation with Kinematic Hardening

The results of this section are taken from [30]. We give an overview of the aspects, which are the most relevant for this thesis. That article treats only the case of primal small-strain plasticity with linear kinematic hardening. For convenience, we therefore recall the objective functional after space and time discretization. Since there is no isotropic hardening present, we set $\mathbf{G} \equiv 0$ in (4.7). We eliminate the isotropic hardening variable completely from our system. Hence, the discrete objective functional (4.17) takes the form

$$L(w) = \frac{1}{2} w^T A w - b^T w + \sum_{i=1}^{N_p} \omega_i \sigma_0 \left\| p_i \right\|_2 \tag{6.1}$$

for the yield stress σ_0, cf. Section 2.4, and the shape function integrals ω_i (4.16). Note that we used the isometry $\|B(p)\|_F = \|p\|_2$ between the matrix valued $B(p) \in \mathbb{S}_0^d$ and the vector $p \in \mathbb{R}^{d_p}$ (4.11).

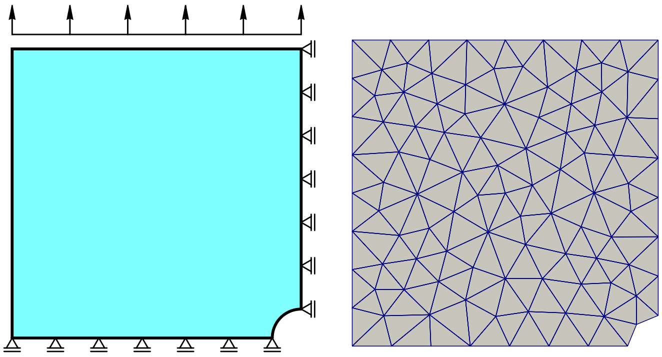
6.1.1. Two-Dimensional Test

The first numerical test in [30] is in two dimensions. All length specifications are measured in millimeters (mm). For a test body we consider a square $[-10, 10]^2$, measured in mm, with a circular hole in the center with a radius of one millimeter. We apply a symmetric pulling force on both the upper and the lower edge. Since we investigate only isotropic materials the resulting displacement and plastic strain will be point-symmetric to the center of the square. Hence, we can consider only the upper left quarter and set appropriate boundary values. To this end, we consider the domain $\Omega = [0, 10]^2 \setminus C$, where C is the circle with radius one millimeter at the center. To justify the reduction to the upper left quarter, we state the sliding boundary values on the displacement field $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2)$

$$\mathbf{u}_1(10, x_2) = 0 \quad \forall x_2 \in [1, 10], \quad \mathbf{u}_2(x_1, 0) = 0 \quad \forall x_1 \in [0, 9]$$

on the bottom and right edge.

The initial grid covering Ω consists of 176 triangular grid elements. For the multigrid solvers we will create a grid hierarchy of several uniform global refinements (levels) of this grid. In Figure 6.1 the domain with sliding boundary conditions on the displacement field, as well as the initial coarse grid are visualized. Moreover, a table of the number of grid elements and grid vertices are given for several uniform grid refinements. The maximal number of grid levels is 8 since it is the maximal level to fit into the memory of the test machine (32 GB). We choose an isotropic material for this test. The material parameters below are measured in N/mm^2. The Lamé parameters (2.7) are



levels	elements	vertices
1	176	105
2	704	385
3	2 816	1 473
4	11 264	5 761
5	45 056	22 785
6	180 224	90 625
7	720 896	361 473
8	2 883 584	1 443 841

Figure 6.1.: From [30]. Left: The domain Ω including sliding boundary conditions. Middle: Initial (coarse) grid covering Ω. Right: Number of grid elements and vertices of several uniform grid refinements

$$\lambda = 6.5 \cdot 10^6, \quad \mu = 10^7.$$

The yield stress σ_0 and the kinematic hardening parameter k_1 (2.11) are set to

$$\sigma_0 = 450, \quad k_1 = 3 \cdot 10^6.$$

We apply a surface force to the top edge of the square ($\Gamma := [0, 10] \times \{10\}$) by the time-dependent load functional

$$\langle \mathbf{l}(t), \mathbf{u} \rangle := 100t \int_\Gamma \mathbf{u}_2 \, dS$$

that acts on the second component of the displacement field.

In total 20 time steps are computed. The discrete time points are $t_i := i$ for $i \in \{1, 2, \ldots, 20\}$.

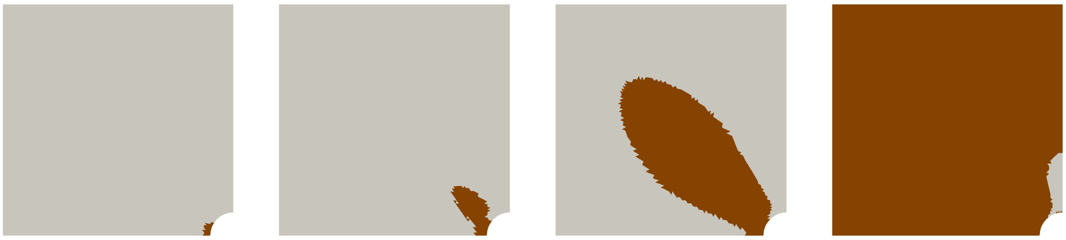


Figure 6.2.: From [30]. Plastic strain result for the time steps 3, 4, 5 and 6

The resulting evolution of a plastic strain is displayed in Figure 6.2. The initial value is zero. One can clearly see that the plastic strain starts to evolve at the hole in the center and then rapidly covers almost the complete domain. Before time step 3 the material behaves completely elastic, and after time step 6 it is completely plastic.

A relative stop criterion is used for both algorithms during the simulations. In each time step t_i the solver iteration stops once the energy norm $\|\cdot\|_A$ (with the matrix A (4.12)) of the current correction correction drops below 10^{-7} of the norm of the first correction.

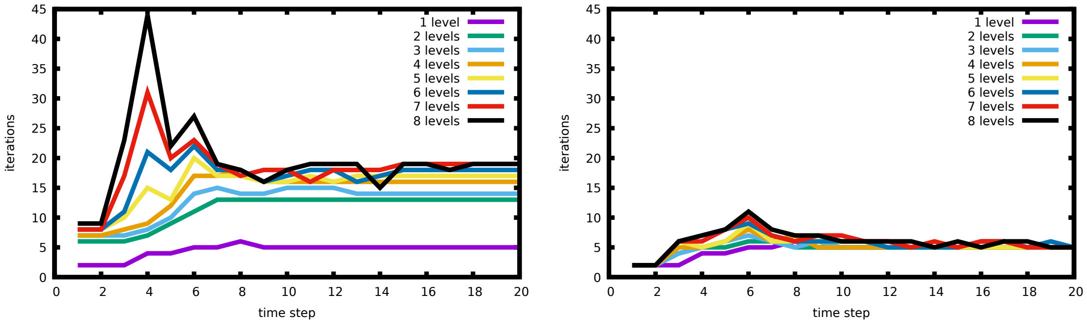


Figure 6.3.: From [30]. Number of iterations until convergence for the 20 time steps. Left: TNNMG. Right: Predictor–Corrector

In Figure 6.3 for each time step the number of necessary steps to achieve the relative convergence criterion are given for both solution algorithms. The different lines indicate the number of grid levels of the multigrid hierarchy, thus the resolution of the solution.

For the TNNMG method one can see that the number of iterations seems to be bounded from above for increasing grid levels for most of the time steps. For the first time steps (purely elastic) the number is at most 10, followed by a peak at time step 4 with 45 iterations on 8 grid levels. This step indicates the transmission to an elastoplastic process, hence the active set (the degrees of freedom in which the objective functional is nonsmooth) changes. From time step 5 on the iterations are again bounded by 20. For the ctPC method the number are significantly smaller. With a peak at time step 6 with 11 iterations the numbers are generally below 10 for all other time steps. Note that on one level TNNMG degenerates to an exact solver of a Predictor–Corrector type, and hence the iteration numbers for this level are equal in both diagrams.

We conclude that the ctPC method yields about half of the required iteration numbers of TNNMG for plastic processes after time step 5.

We also investigate the required wall time for both solvers since a single ctPC step is likely more time consuming than an inexact multigrid step of TNNMG. To this end, the wall times of both algorithms for the time steps are presented in Figure 6.4.

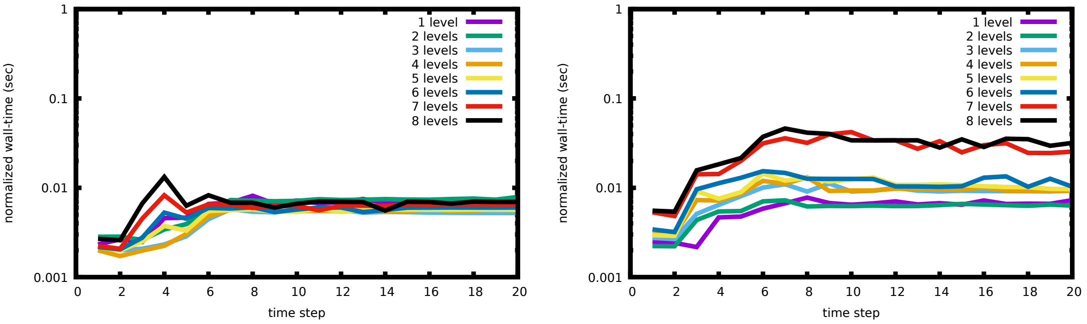


Figure 6.4.: From [30]. Normalized wall time for the computation of the 20 time steps. Left: TNNMG. Right: Predictor–Corrector

For a better comparability we display the *normalized* wall-time: On one grid level this is the original wall time. On all other levels it is scaled inversely proportional to the number of degrees of freedom. This is a good tool to recognize a linear dependence of the wall time depending on the number of degrees of freedom. In this case all lines for the different numbers of grid levels coincide. It seems to be the case for the TNNMG algorithm. Apart from the first 5 time steps (elastic–plastic transition), the normalized wall times coincide. Hence, we conclude that TNNMG has a wall time linearly dependent on the number of degrees of freedom, as expected for a multigrid solver. On the other hand, the ctPC solver does show a superlinear time dependence. For one grid level both algorithms are comparable. However, for the other grid levels, the necessary wall time for the ctPC solver is much higher for TNNMG. On the finest grid level 8 the required wall time of the ctPC algorithm is about five times higher than the one of TNNMG. We have no mathematical explanation for the jump of the wall times between 6 and 7 grid levels. We suppose it arises from caching effects in the hardware.

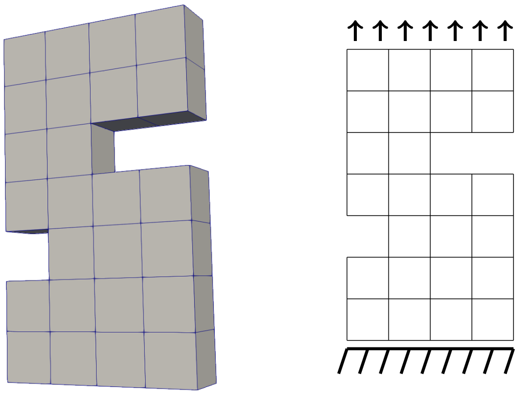
In summary, we conclude that TNNMG has superior performance to the ctPC solver for large grids.

In [30] additional tests are presented: First of all, the memory consumption of TNNMG is about half of the one of ctPC, which is 5 GB vs. 10 GB on 8 grid levels. Direct solvers are very memory consuming for large problems. To overcome this issue, a combination of ctPC and repeated multigrid steps is considered and tested. Even then TNNMG turned out to be the fastest solver.

Another test is done for a lower values of the hardening parameters k_1. It turned out that the minimization problem becomes ill-conditioned in this case causing TNNMG to require an unbounded number of iterations for increasing grid levels, while the direct solvers of ctPC are not affected by small values of k_1. To improve this situation for TNNMG, a Schur complement technique is applied resulting in better condition numbers and hence, better performance.

6.1.2. Three-Dimensional Test

A three dimensional test with the same material properties as in the previous Section 6.1.1 is also considered in [30] and quickly presented here.



levels	elements	vertices
1	25	168
2	200	942
3	1 600	6 108
4	12 800	43 512
5	102 400	327 408
6	819 200	2 537 952

Figure 6.5.: From [30]. Left/Middle: Geometry, boundary conditions and initial grid of the 3D test object. Right: Number of grid entities for grid refinements

The test body is given in a bounding box $[0,4] \times [0,7] \times [0,1]$ (measured in millimeters) with two notches, as seen in Figure 6.5. The initial coarse grid containing 25 cubical elements as well as the number of grid elements and grid vertices for various uniform grid refinements are given in this figure.

Note that due to the curse of dimensionality the number of grid entities, and therefore the degrees of freedom of the discrete problem are growing much quicker than in the two dimensional case. Hence, we consider only up to 6 grid levels, since another refinement did not fit into the memory of the test machine.

Boundary conditions are given by sticky Dirichlet conditions at the bottom face

$$\mathbf{u}(x_1, 0, x_3) = 0 \quad \forall (x_1, x_3) \in [0,4] \times [0,1]$$

and a surface load term is applied to the top face $\Gamma = [0,4] \times \{7\} \times [0,1]$ by the time-dependent load functional

$$\langle \mathbf{l}(t), \mathbf{u} \rangle := 20t \int_\Gamma \mathbf{u}_2 \, dS$$

acting on the second (upward facing) component of the displacement.

Again, we consider 20 time steps $t_i := i$ and use the same solver options as in the previous test in Section 6.1.1.

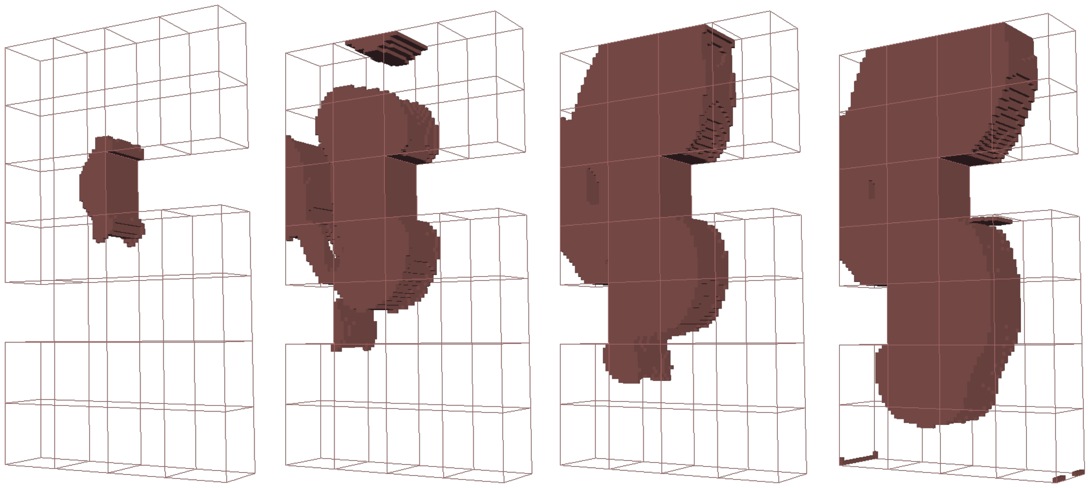


Figure 6.6.: From [30]. Plastic strain result for the time steps 5, 10, 15 and 20

In Figure 6.6 the evolution of the plastic strain in the test body is shown for the time steps 5, 10, 15 and 20. Starting from time step 5 a plastic strain starts to develop at the narrowest part of the test body. Until the last time step 20 a majority of the test body has a nonzero plastic strain. Displacement is not shown since it is negligible in this small-strain test.

In Figure 6.7 the numbers of necessary iteration steps for both solution algorithms for solving the problem in the 20 time steps are shown. First of all, we were not able to compute solutions of the ctPC solver for 6 levels due to insufficient memory on the test machine for the direct solver. We notice that the number are again comparable for one grid level since both solvers are of the same type in this case. For an increasing number of grid levels TNNMG shows increasing iteration numbers that seem to converge to an upper bound little above 30 iterations. Moreover, the numbers tend to become large with increasing time steps. This happens most likely because the number of activated components of the plastic strain increases faster with expanding plastic strain. On the other hand, the iteration numbers for the ctPC solver are almost unaffected by the grid fineness and the time step. With at most 7 steps the total number is much lower than on the TNNMG side.

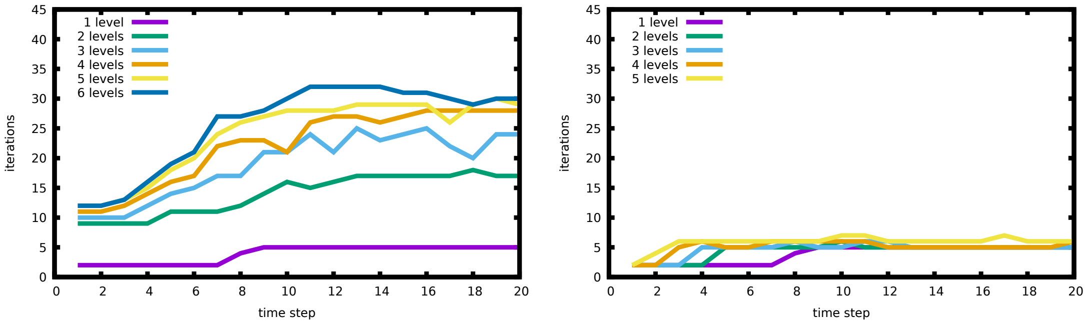


Figure 6.7.: From [30]. Number of iterations until convergence for the 20 time steps. Left: TNNMG. Right: Predictor–Corrector

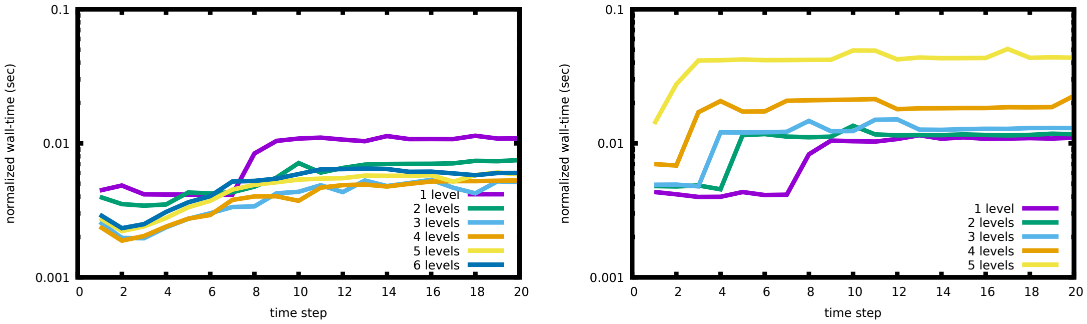


Figure 6.8.: From [30]. Normalized wall time for the computation of the 20 time steps. Left: TNNMG. Right: Predictor–Corrector

The normalized wall times for the 20 time steps are given in Figure 6.8. For TNNMG we see again a linear time dependence on the wall time on the number of degrees of freedom. On the contrary, ctPC shows a superlinear dependence. As a result, for 5 grid levels, the TNNMG method is about 10 times faster than the ctPC method. Moreover, the memory consumption at 5 grid levels of ctPC is about 6700 MB compared to only 729 MB of TNNMG [30].

This demonstrates that direct solvers such as CHOLMOD have a huge time consumption applied to solving discretized problems of three dimensional PDEs on fine grids. Even on 6 grid levels TNNMG consumes about only 3500 MB. Overall, we can show the efficiency of TNNMG both in the wall time and in the memory consumption.

6.2. Three-Dimensional Tresca Dissipation

In this test, we consider only kinematic hardening for better comparability with the previous von Mises tests. Hence, we still set formally $\mathbf{G} \equiv 0$ in (4.7), and therefore do not consider any isotropic hardening variable $\boldsymbol{\eta}$ at all. The objective functional (4.17) takes now the form

$$L(w) = \frac{1}{2} w^T A w - b^T w + \sum_{i=1}^{N_p} \omega_i \sigma_0 \left\| B(p_i) \right\|_2 \tag{6.2}$$

with the spectral norm $\|\cdot\|_2$.

For best comparison to the numerical test in Section 6.1.2 we reuse the domain, the sticky boundary conditions and initial grid, as shown in Figure 6.5.

However, material parameters in this test are chosen differently from Section 6.1.2. We consider a yield stress $\sigma_0 = 450$, Lamé parameters (2.7) $\mu = 8 \cdot 10^5$ and $\lambda = 1.2 \cdot 10^6$, as well as a kinematic hardening paramter $k_1 = 41080$. The mentioned parameters are measured in $\mathrm{N/mm^2}$.

External forces are again given as surface forces at the top face Γ of the shape ($x_2 = 7$). We consider the time-dependent load functional

$$\langle \mathbf{l}(t), \mathbf{u} \rangle := 300t \int_\Gamma \mathbf{u}_2 \, dS.$$

We compute the solutions for 20 time steps $t_i := i$ for i from 1 to 20. The loading force and the time steps are chosen such that we have a comparable material behavior as in the three dimensional von Mises test case.

In detail, we have again about 4 elastic steps with no appearing plastic strain, followed by a plastic evolution. Finally, at time step 20, a majority of the domain has a nonzero plastic strain.

Apart from the slightly different material model the solver-specific options are equal to the ones from Section 6.1.2.

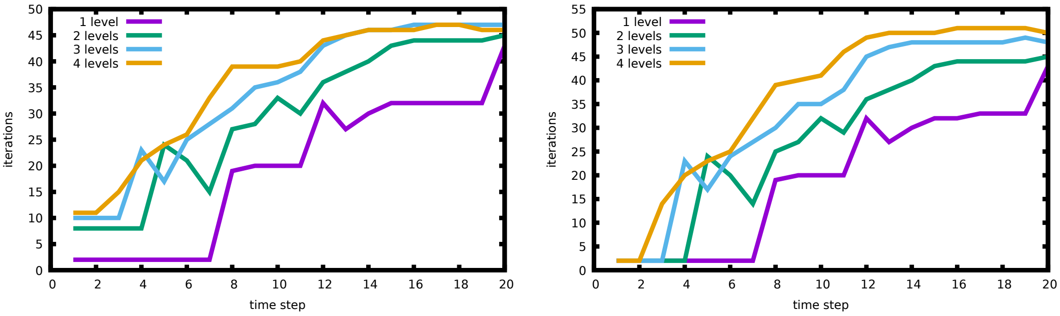


Figure 6.9.: Number of iterations until convergence for the 20 time step on difference grid resolutions Left: TNNMG. Right: Predictor–Corrector

The necessary iteration steps for TNNMG, and the ctPC solver to compute the solutions for the 20 time steps are displayed in Figure 6.9.

In contrast to the equivalent results with von Mises dissipation in Figure 6.7 the number are almost the same on all grid levels for both algorithms.

The reason is the spectral norm $\|B(p_i)\|_2$ in the discrete objective functional (6.2). While the Frobenius norm has exactly one nondifferentiable point at zero, the spectral norm is nonsmooth at every matrix with non-unique dominant eigenvalue. This increases the degree of nonsmoothness. Hence, minimization with nonsmooth Newton methods, like TNNMG and ctPC, is more difficult.

Recall that TNNMG computes anyway only inexact nonsmooth Newton steps. Hence, comparably smooth cases, like von Mises dissipation, the total iteration numbers are already high. In contrast, the exact and direct Newton steps of the ctPC method require a low number of iterations for convergence (cf. Figure 6.7). However, for Tresca dissipation, there is no advantage of computing exact Newton steps, since the Newton corrections will likely "cross" a nondifferentiability of the objective functional. Therefore, the minimization of the objective functional relies more on the nonsmooth local solvers from the first stage of TNNMG and the corrector step of ctPC from Section 5.3.

In the von Mises case, TNNMG is already superior in the wall time measurements compared to the ctPC method, as seen in Figure 6.8. We expect this gap to be even larger in the Tresca case. The *normalized* wall times for Tresca dissipation are shown in Figure 6.10.

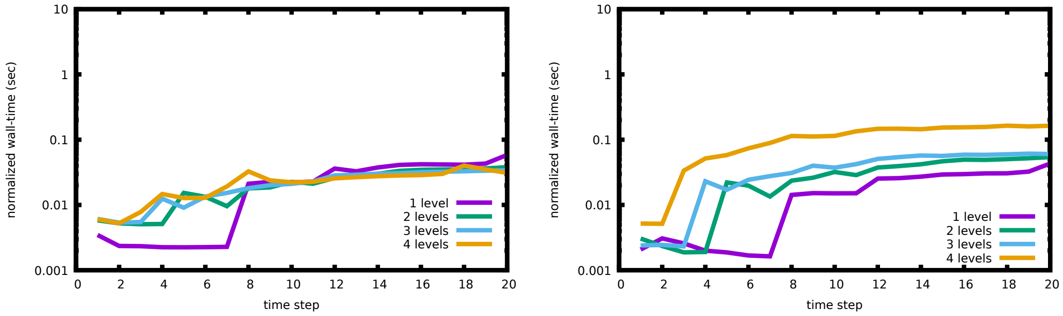


Figure 6.10.: Normalized wall-time for the computation of the 20 steps solutions for different refinement levels. Left: TNNMG. Right: Predictor–Corrector

The TNNMG solver shows again a linear wall time dependence on the number of degrees of freedom as all lines tend to coincide. The ctPC with the direct solver CHOLMOD starts with a linear wall time behavior for the first 3 levels. This is followed by a quadratic wall time dependency from level 3 to 4. The reason for the linear start is simply that direct solvers are very efficient for small linear problems. The wall time is dominated by the overhead of the grid and solver structures for coarse grids. And grid-dependent routines are usually linear in the grid size, and hence in the degrees of freedom of the problem.

Clearly, TNNMG is the faster algorithm with a gap of factor 8 compared to the Predictor–Corrector at 4 grid levels. Computations with more than 4 grid levels exceed the possible measurable wall time on the test machine. We conjecture that the gap in the wall times of TNNMG and ctPC will be even larger for finer grids.

Additional extensions of the ctPC method to a numerically exact multigrid solver for the elastic predictor step (as in [30] for von Mises dissipation) are not considered here. It will most likely reduce the quadratic wall time dependence of CHOLMOD to a linear one, but as seen in the number of iterations (Figure 6.9), there is no advantage in solving the Newton steps exactly. Hence, this improvement strategy for the ctPC solver cannot outperform TNNMG.

6.3. Isotropic Hardening

The last numerical test extends the two-dimensional example from [30]. We want to investigate the impact of isotropic hardening on the wall time of the presented solvers. This test covers both Tresca and von Mises dissipation, since they are equivalent in two dimensions (Lemma 2.4.3).

We reuse the two dimensional grid from Section 6.1.1. Boundary values as well as the grid hierarchy are unchanged.

For a better investigation of the influence of isotropic hardening we choose a load functional such that we consider two complete hysteresis loops. This is achieved by the load functional

$$\langle \mathbf{l}(t), \mathbf{u} \rangle := 450 \cdot \sin(4\pi t) \int_\Gamma \mathbf{u}_2 \, dS.$$

on the top edge Γ of the domain for time t in $[0, 1]$. The time interval $[0, 1]$ is divided into 100 equal parts. In this benchmark, we use either kinematic or isotropic hardening only. In the isotropic case we use the non-linear isotropic hardening functional

$$g(\eta) = (\sigma_0 - k_\infty)(1 - e^{-\beta\eta}) - k_2\eta,$$

as from (2.13). For a better comparison to the kinematic case we use $k_2 = k_\infty = 41080$ N/mm, $\beta = 1$ and reuse the other material parameters from Section 6.2.

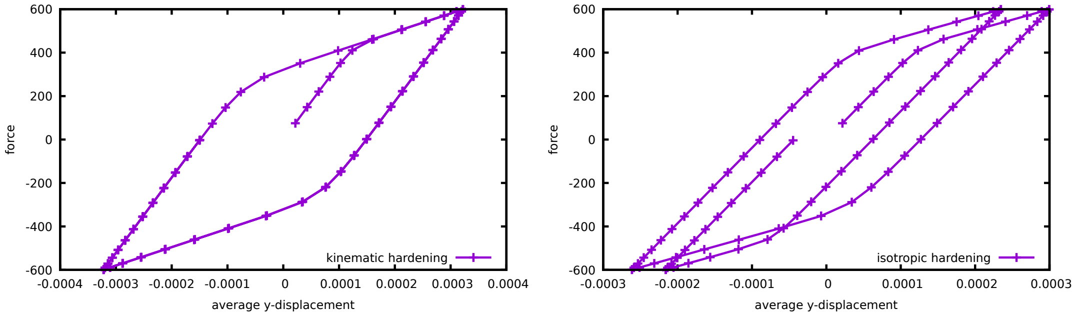


Figure 6.11.: Hysteresis for two force loops. Left: kinematic hardening. Right: isotropic hardening

In Figure 6.11 the resulting hysteresis curves for pure kinematic and pure isotropic hardening are given. We compute the average displacement in the x_2 component over the domain as a drop-in of the usually used strain measurement of these hysteresis curves.

For the kinematic hardening test (with $k_1 = 41080$ N/mm) the hysteresis curve is the expected parallelogram. The isotropic result shows the development of a growing elastic region after plastic work occurred, visible as the shrinking width of the parallelogram.

For both simulations we use the TNNMG solver with the same solver options as given in Section 6.2. The total time for all 100 homotopy steps does not depend on the chosen hardening functional, as seen in Figure 6.12.

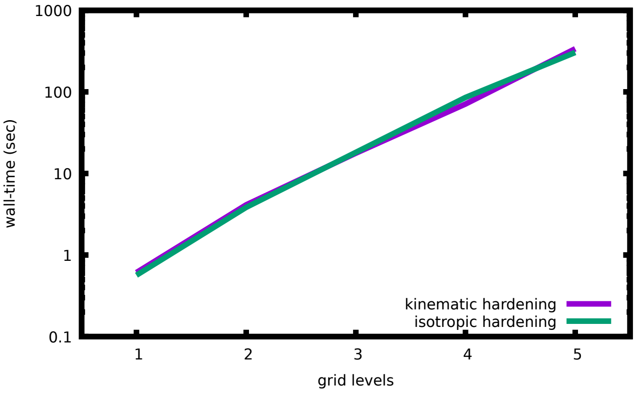


Figure 6.12.: Total run time for the hysteresis loop with 100 homotopy steps

6.4. Conclusion

First of all, we highlight that the TNNMG algorithm outperforms the Predictor–Corrector method in all test cases. This was already demonstrated in [30] for the von Mises case. In this thesis we are able to show that the performance gap between the TNNMG and ctPC solver is even larger.

In summary, TNNMG turns out to be a simple and easy to use algorithm with high efficiency for highly nonsmooth objective functionals. Another advantage of TNNMG is its form of an algebraic black box solver. There are only few parameters to set, especially no problem-depending control-parameters.

Another result addresses isotropic hardening. As seen in the last test case the plastic material behavior depends heavily on the considered hardening rule. Nevertheless, the solvers are not affected by this. For all considered hardening rules the wall times and the number of iterations are comparable. The reason is most likely that isotropic hardening rules lead to smooth parts of the objective functional. The critical parts for the quality of the solvers are the nonsmooth components arising from the dissipation functionals.

Part II.

An Efficient and Globally Convergent Minimization Algorithm for Finite-Strain Plasticity Problems

7. Introduction to Finite-Strain Mechanics

The restriction to small strains may not be appropriate in certain situations. Therefore, in this section we present a rigorous continuum mechanics model of finite-strain plasticity. Although the structure of this chapter is similar to Chapter 2 in the first part of this thesis, each section introduces new mathematical aspects since we drop the assumption of small strains.

7.1. Kinematics

7.1.1. Deformation

Similar to the first part of this thesis, the kinematics are formulated in the physically meaningful space dimensions two and three. Hence, the we have $d \in \{2, 3\}$. The domain $\Omega \subset \mathbb{R}^d$ is again closed and bounded. In the theory of small strains the displacement field

$$\mathbf{u} : \Omega \to \mathbb{R}^d$$

is the kinematic quantity of interest, since in the first part of this thesis we have the fundamental assumption that $\nabla \mathbf{u}$ is in some sense *small*. In the second part of the thesis, this assumption is dropped. Hence, there is no practical difference between using either the displacement field $\mathbf{u}$ and directly considering the deformation field

$$\mathbf{y} : \Omega \to \mathbb{R}^d, \ \mathbf{y}(x) = \mathbf{u}(x) + x.$$

For an illustration of $\mathbf{y}$ and $\mathbf{u}$ consider Figure 2.1 in the first part of this thesis. The theory of finite-strain mechanics is usually formulated in terms of the deformation field $\mathbf{y}$, and we will follow this path. Since we do not consider the effects of rigid body motions, only spatial derivatives of the deformation field will contribute to the *strain* of the material. The spatial gradient of the deformation field is given by

$$\mathbf{F} \coloneqq \nabla \mathbf{y} = \nabla \mathbf{u} + \mathbf{I},$$

and is directly connected to $\nabla \mathbf{u}$ by a unit matrix shift.

7.1.2. Strain Tensors

In the first part there is already a definition of the strain tensor given, cf. Definition 2.1.1. For finite strains, multiple definitions of *strains* are found in the literature. The one introduced before is the so-called Green–St. Venant strain tensor and it is repeated here for completeness.

Definition 7.1.1 (Green–St. Venant Strain)**.** *Let* $\mathbf{F} : \Omega \to \mathbb{R}^{d\times d}$ *deformation gradient field. Then we call*

$$\mathbf{E} := \frac{1}{2}\left(\mathbf{F}^T\mathbf{F} - \mathbf{I}\right) \tag{7.1}$$

the Green–St. Venant strain tensor.

Another frequently found strain tensor is the (right) Cauchy–Green strain tensor.

Definition 7.1.2 (right Chauchy–Green Strain)**.** *Let* $\mathbf{F} : \Omega \to \mathbb{R}^{d\times d}$ *deformation gradient field. Then we call*

$$\mathbf{C} := \mathbf{F}^T\mathbf{F} \tag{7.2}$$

the (right) Cauchy–Green strain tensor.

Remark 7.1.3. *The Green–St. Venant strain tensor is zero for rigid motions* $\mathbf{y} = Q\mathbf{y} + q$ *with an orthogonal* $Q \in \mathrm{SO}(d)$ *and* $q \in \mathbb{R}^d$*.*

7.1.3. Decomposition of the Deformation Gradient

In the first part of this thesis we introduced the plastic strain variable by splitting the small-strain tensor

$$\boldsymbol{\epsilon} = \frac{1}{2}\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right)$$

additively into the elastic part $\mathbf{e}$ and the plastic part $\mathbf{p}$, such that,

$$\boldsymbol{\epsilon} = \mathbf{e} + \mathbf{p}. \tag{7.3}$$

For finite strains, the decomposition into elastic and plastic parts is not applied to the Green–St. Venant strain tensor $\mathbf{E}$, but rather directly to the deformation gradient $\mathbf{F}$. Moreover, we do not consider an additive split, but a multiplicative splitting

$$\mathbf{F} = \mathbf{F}_{el}\mathbf{P}. \tag{7.4}$$

This multiplicative split was introduced by [20]. It is used throughout this part of the thesis. At a first glance, this looks to be in complete contrast to the additive split (7.3). In the following we will clarify that this is actually just an extension of the additive split used in the small-strain case. In the small-strain case, we assume that both $\mathbf{e}$ and $\mathbf{p}$ are pointwise small, i.e., there is a $\mathbf{h}\colon \Omega \to \mathbb{R}$ with small values and

$$\max\{\|\mathbf{e}\|, \|\mathbf{p}\|\} \le \mathbf{h}. \tag{7.5}$$

Next, we want to express $\mathbf{F}_{\mathrm{el}}$ and $\mathbf{P}$ in terms of $\mathbf{e}$ and $\mathbf{p}$. The required tool for this is the exponential function of matrices. For a matrix $A \in \mathbb{R}^{d\times d}$ the matrix exponential is given by

$$\exp(A) := \sum_{i=0}^{\infty} \frac{A^i}{i!}.$$

It is well-known (e.g., [18]) that the matrix exponential is well-defined for all matrices with quadratic shape. If the norm of A is small, then the exponential series above decays quickly. Hence, we have

$$\exp(A) = I + A + \mathcal{O}(\|A\|^2) \approx I + A \quad \text{for } \|A\| \to 0.$$

Now we use the matrix exponential to set

$$\mathbf{F}_{el} := \exp(\mathbf{e}), \quad \mathbf{P} := \exp(\mathbf{p}).$$

If we assume (7.5), then we have

$$\mathbf{F} = \mathbf{F}_{el}\mathbf{P} = \mathbf{e} + \mathbf{p} + \mathbf{I} + \mathcal{O}(\mathbf{h}^2),$$

and hence,

$$\mathbf{E} = \frac{1}{2}\left(\mathbf{F}^T\mathbf{F} - \mathbf{I}\right) = \mathbf{e} + \mathbf{p} + \mathcal{O}(\mathbf{h}^2).$$

In summary, the additive split of the small-strain tensor $\boldsymbol{\epsilon}$ can be transferred into multiplicative split of $\mathbf{F}$ (7.4) by using the matrix exponential. Note that both $\mathbf{F}_{\mathrm{el}}$ and $\mathbf{P}$ are in general allowed to be non-symmetric.

7.2. Ranges

7.2.1. Matrix Manifolds

We use matrix groups for the ranges of the elastic strain $\mathbf{F}_{\mathrm{el}}$ and the plastic strain $\mathbf{P}$. This is more advanced than the vector space approach in small-strain plasticity. However, this allows us to use the multiplicative split of the deformation gradient $\mathbf{F} = \mathbf{F}_{\mathrm{el}}\mathbf{P}$ without leaving the matrix group structure.

The deformation field $\mathbf{y}$ describes the deformed domain $\Omega^t = \mathbf{y}(\Omega)$. Hence, from the integral transformation formula

$$\int_{\Omega(\mathbf{y})} d\hat{x} = \int_{\Omega} |\det(\nabla\mathbf{y})| \, dx$$

we conclude that $\det(\mathbf{F}) = \det(\nabla\mathbf{y})$ is the local volume transformation density. From a mechanical point of view, we do not allow a degenerate material with vanishing volume, i.e., we postulate $\det(\mathbf{y}) > 0$. The set of matrices equipped with this property forms the multiplicative group

$$\mathrm{GL}(d)^+ := \{A \in \mathbb{R}^{d\times d} : \ \det(A) > 0\}, \tag{7.6}$$

which forms a subgroup of the *generalized linear group*, where the constraint $\det(A) \neq 0$ is given. Furthermore, we assume that plastic effects do not change the volume, cf. [22, Chapter 4.2.1]. Hence, the range of the plastic strain is equipped with the stronger constraint $\det(\nabla\mathbf{y}) = 1$. Again, since the determinant is multiplicative, this forms another matrix group

$$\mathrm{SL}(d) := \{A \in \mathbb{R}^{d\times d} : \ \det(A) = 1\}, \tag{7.7}$$

called the *special linear group*. Obviously, $\mathrm{SL}(d)$ is a subgroup of $\mathrm{GL}(d)^+$. In summary, we consider $\mathrm{GL}(d)^+$ for the range of the elastic strain

$$\mathbf{F}_{\mathrm{el}}(x) \in \mathrm{GL}(d)^+ \quad \text{for almost all } x \in \Omega,$$

whereas the plastic strain takes values in $\mathrm{SL}(d)$

$$\mathbf{P}(x) \in \mathrm{SL}(d) \quad \text{for almost all } x \in \Omega.$$

Note that the "for almost all" statement is already placed here to comply with the embedding into Sobolev spaces later on.

7.2.2. Plastic Spin

Plastic spin has different meanings and definitions along the history on finite-strain continuum mechanics. An excessive overview is given in [5]. In this thesis, whenever we refer to *plastic spin* we indicate to the rotational parts of the plastic range. These rotational parts can be introduced formally by the orthogonal part of the *polar decomposition* of the values of $\mathbf{P}$.

Lemma 7.2.1 ([18, Theorem 8.1])**.** *For any regular matrix $P \in \mathbb{R}^{d\times d}$ there is a unique polar decomposition into an orthogonal matrix L and a unique symmetric, positive definite matrix V, such that*

$$P = LV.$$

Remark 7.2.2. *By applying the polar decomposition to the transpose $P^T = LV$, we get immediately another decomposition*

$$P = VR$$

with the orthogonal matrix $R := L^T$ on the right side.

From the global pointwise polar decomposition $\mathbf{P} = \mathbf{LV}$, we associate the rotational part $\mathbf{L}$ with the plastic spin. In this thesis we cover both the general plastic range $\mathrm{SL}(d)$, and the case of spinless plasticity. For the latter, we introduce the set

$$\mathrm{SL}(d)^+_{\mathrm{sym}} := \{A \in \mathrm{SL}(d) : \ A \text{ is symmetric and positive definite}\} \tag{7.8}$$

for the range of the plastic strain without spin. Then, for each $\mathbf{P}$ with values in $\mathrm{SL}(d)^+_{\mathrm{sym}}$, we have can identify $\mathbf{L} \equiv I$ in the polar decomposition. Note that $\mathrm{SL}(d)^+_{\mathrm{sym}}$ is not a matrix group under matrix multiplication since the product of two symmetric matrices is in general not symmetric. Further numerical treatment of this set is given in Section 9.2.2. For now, we will use the general case of the plastic strain with spin until further notice.

7.3. Stress Tensors

At small strains we postulate the existence of the Cauchy stress tensor field $\boldsymbol{\sigma} : \Omega \to \mathbb{S}^d$, such that,

$$-\operatorname{div}(\boldsymbol{\sigma}) = \mathbf{b}$$

for the volumetric force density field $\mathbf{b}$ (2.3). The existence of the Cauchy stress tensor on the initial domain requires the gradient of the displacement field, and hence the strain $\mathbf{E}$, to be small [16, Chapter 2.2]. At finite strains, we have to distinguish whether we postulate the physical quantities on the initial domain Ω, or on the deformed domain $\Omega^{\mathbf{y}} := \mathbf{y}(\Omega)$. From the mathematical point of view it is does not matter to consider the initial domain Ω all the time. However, in the real world, forces are in fact applied to the deformed body, since the initial state may not be available. We revisit Proposition 2.2.1 from the first part of this thesis without the assumption of a small strain tensor. This form is found in, e.g., [4, Chapter II.2.3].

Proposition 7.3.1 (Cauchy Stress at finite strains)**.** *Let* $\mathbf{y} : \Omega \to \mathbb{R}^3$ *be a smooth deformation field, and set* $\Omega^{\mathbf{y}} := \mathbf{y}(\Omega)$ *for the deformed domain. Moreover, let* $\mathbf{b}^{\mathbf{y}} : \Omega^{\mathbf{y}} \to \mathbb{R}^d$ *a volumetric force density field on the deformed domain, and let* $\mathbf{s}^{\mathbf{y}} : \partial\Omega^{\mathbf{y}} \to \mathbb{R}^d$ *a surface force density field on the boundary of the deformed domain, such that the total force is given by*

$$F^{\mathbf{y}} := \int_{\Omega^{\mathbf{y}}} \mathbf{b}^{\mathbf{y}} \, dx^{\mathbf{y}} + \int_{\partial\Omega^{\mathbf{y}}} \mathbf{s}^{\mathbf{y}} \, dS.$$

Then there is a differentiable and symmetric matrix field $\boldsymbol{\sigma} : \Omega^{\mathbf{y}} \to \mathbb{R}^{d\times d}$ *such that*

$$\mathbf{s}^{\mathbf{y}}(x^{\mathbf{y}}) = \boldsymbol{\sigma}(x^{\mathbf{y}}) \cdot \mathbf{n}(x^{\mathbf{y}})$$

for the outer normal $\mathbf{n}(x^{\mathbf{y}})$ *of* $\partial\Omega^{\mathbf{y}}$ *at* $x^{\mathbf{y}}$*, and furthermore*

$$F^{\mathbf{y}} = \int_{\Omega^{\mathbf{y}}} \operatorname{div}(\boldsymbol{\sigma}) + \mathbf{b}^{\mathbf{y}} \, dx^{\mathbf{y}}.$$

We call the field $\boldsymbol{\sigma} : \Omega^{\mathbf{y}} \to \mathbb{R}^{d\times d}$ the Cauchy stress tensor.

Remark 7.3.2. *Note that all appearing quantities in the proposition above are time-dependent. For notational simplicity the dependency on the time is not explicitly written.*

From a computational point of view, the Cauchy stress is not convenient because it is defined on the deformed domain, which is usually unknown beforehand. To resolve this problem, the *Piola transformation* comes into play. Without going into detail here, this transformation allows us to back-trace functionals on the deformed domain back to the initial domain. An excessive introduction can be found in [4, Chapter II.2.5]. In our case, we define a new stress field $\mathbf{T} : \Omega \to \mathbb{R}^{d\times d}$, which acts as the counterpart of $\boldsymbol{\sigma}$ on $\Omega^{\mathbf{y}}$.

Definition 7.3.3 (First Piola–Kirchhoff Stress)**.** *Let* $\mathbf{y} : \Omega \to \mathbb{R}^3$ *be a smooth deformation field,* $\Omega^{\mathbf{y}} := \mathbf{y}(\Omega)$ *the deformed domain. Furthermore, let* $\boldsymbol{\sigma} : \Omega^{\mathbf{y}} \to \mathbb{R}^{d\times d}$ *be the Cauchy stress tensor field. We define the first Piola–Kirchhoff stress tensor field by*

$$\mathbf{T} : \Omega \to \mathbb{R}^{d\times d}, \ \mathbf{T}(x) := \det(\nabla\mathbf{y}(x))\, \boldsymbol{\sigma}(\mathbf{y}(x))\, \nabla\mathbf{y}(x)^{-T}.$$

Remark 7.3.4. *For small strains, we can assume that* $\nabla \mathbf{u} \approx 0$*, and hence* $\nabla \mathbf{y} \approx \mathbf{I}$*. Therefore, we have* $\mathbf{T}(x) \approx \boldsymbol{\sigma}(\mathbf{y}(x))$*. From Dirichlet data on parts of* $\partial\Omega$ *we can furthermore assume that* $\mathbf{y}(x) \approx x$*, and therefore* $\mathbf{T}(x) \approx \boldsymbol{\sigma}(x)$ *can be identified.*

Now we have again a stress quantity $\mathbf{T}$ directly given on the initial domain Ω. A disadvantage is the lack of symmetry. While the Cauchy stress tensor is symmetric, the first Piola–Kirchhoff stress tensor is generally not. To overcome this issue, we simply enforce symmetry by a multiplication from the left with $\nabla \mathbf{y}^{-1}$ and arrive at the second Piola–Kirchhoff stress tensor.

Definition 7.3.5 (Second Piola–Kirchhoff Stress)**.** *Let* $\mathbf{T} : \Omega \to \mathbb{R}^{d\times d}$ *be the first Piola–Kirchhoff stress tensor field induced by the smooth deformation field* $\mathbf{y} : \Omega \to \mathbb{R}^d$*. Then the symmetric second Piola–Kirchhoff stress tensor field is defined by*

$$\boldsymbol{\Sigma} : \Omega \to \mathbb{S}^d, \;\; \boldsymbol{\Sigma}(x) = \nabla \mathbf{y}^{-1}\mathbf{T}(x).$$

Remark 7.3.6. *At small strains, where* $\nabla \mathbf{y} \approx \mathbf{I}$*, the first and second Piola–Kirchhoff stress tensors can be identified.*

7.4. Material Laws

7.4.1. Hyperelastic Materials

Next, we introduce elastic material models that interact with the just defined Piola–Kirchhoff stresses. In the case of small-strain elasticity, we assume the existence of a symmetric fourth-order Hooke tensor $\mathbf{H}$ (2.6). Translated into the notation of this part of the thesis, $\mathbf{H}$ linearly connects the second Piola–Kirchhoff stress tensor $\boldsymbol{\Sigma}$ to the Green–St. Venant strain tensor $\mathbf{E}$ (7.1), or alternatively, to the Cauchy–Green strain tensor $\mathbf{C}$ (7.2), by

$$\boldsymbol{\Sigma} = \mathbf{H}\,\mathbf{E} = \frac{1}{2}\mathbf{H}(\mathbf{C} - \mathbf{I}). \tag{7.9}$$

To extend this linear case to more general ones, we postulate an elastic energy density function

$$W_{\mathrm{el}} : \Omega \times \mathbb{R}^{d\times d} \to \mathbb{R}$$

which depends on the actual point $x \in \Omega$ for possibly heterogeneous material properties, and on the deformation gradient $\nabla \mathbf{y}(x) = \mathbf{F}(x)$, from which we can compute the strains $\mathbf{E}(x)$ and $\mathbf{C}(x)$. We call a material *hyperelastic*, if we have the connection to the first Piola–Kirchhoff stress

$$\mathbf{T}(x) = \frac{\partial}{\partial F} W_{\mathrm{el}}(x, \mathbf{F}(x)) \tag{7.10}$$

for almost all $x \in \Omega$. A fundamental axiom of elastic materials is *frame-indifference* [4, Chapter 3]. In short, this means that rotations of the coordinate system (not rotations of the domain Ω) do not change the behavior of the material. This can be mathematically modeled by

$$W_{\mathrm{el}}(x, F) = W_{\mathrm{el}}(x, QF)$$

for all orthogonal matrices Q. From the polar decomposition $F = LV$ into an orthogonal L and a symmetric, positive definite V, we conclude that W_{el} does only depend on the s.p.d. part V. Since the square root of s.p.d. matrices is always well-defined we can avoid the polar decomposition by introducing another elastic energy density $\hat{W}_{\mathrm{el}}$, with the property

$$W_{\mathrm{el}}(x, F) = \hat{W}_{\mathrm{el}}(x, F^T F)$$

for almost all $x \in \Omega$ and $F \in \mathbb{R}^{d \times d}$. Note that the Cauchy–Green strain tensor is precisely $C = F^T F$. A consequence from the definition of hyperelastic materials (7.10) is that

$$\mathbf{\Sigma}(x) = 2 \frac{\partial}{\partial C} \hat{W}_{\mathrm{el}}(x, \mathbf{C}(x)) \tag{7.11}$$

for the second Piola–Kirchhoff stress $\mathbf{\Sigma}$. Back to (7.9), the corresponding hyperelastic material that combines the hyperelastic scheme and the linear Hooke's law (7.9) is the *St. Venant–Kirchhoff material.*

Definition 7.4.1 (St. Venant–Kirchhoff Material)**.** *Let* $\mathbf{H} \in \mathbb{R}^{(d\times d)\times(d\times d)}$ *a fourth-order tensor. Moreover, we assume that* $\mathbf{H}$ *fulfills the symmetry conditions*

$$\mathbf{H}A : B = \mathbf{H}B : A = \mathbf{H}B^T : A = \mathbf{H}B^T : A^T$$

for any matrices $A, B \in \mathbb{R}^{d\times d}$*. Additionally, we demand that* $\mathbf{H}A : A > 0$ *for all* $A \neq 0$*, i.e.,* $\mathbf{H}$ *is positive definite. The elastic energy density of a St. Venant–Kirchhoff material is given by*

$$W_{SVK}(x, F) := \frac{1}{2}\mathbf{H}E : E = \frac{1}{8}\mathbf{H}(C - I) : (C - I) =: \hat{W}_{SVK}(x, C)$$

for the strain tensors $C = F^T F$ *and* $E = \frac{1}{2}(C - I)$*.*

The St. Venant–Kirchhoff material model forms a quadratic energy in terms of the strain.

Lemma 7.4.2. *If we assume a linear connection of the stress and the strain* (7.9)*, then the St. Venant–Kirchhoff material model fulfills*

$$2 \frac{\partial}{\partial C} \hat{W}_{SVK}(x, C) = \Sigma,$$

as well as

$$\frac{\partial}{\partial F} W_{SVK}(x, F) = T.$$

Proof. The first statement follows directly from the symmetry of $\mathbf{H}$, such that,

$$\frac{\partial}{\partial C} \hat{W}_{\mathrm{SVK}}(x, C) = \frac{1}{4}\mathbf{H}(C - I) = \frac{1}{2}\mathbf{H}E = \frac{1}{2}\Sigma.$$

The second property is not trivial, since we have to carefully derive the strain E by F. To this end, it is always worth to take a look at the directional derivative of W_{SVK} in the F-component in an arbitrary direction D. Note that for $E(F) = \frac{1}{2}(F^TF - I)$ we have $E(F + hD) = E(F) + \frac{h}{2}(D^TF + F^TD + hD^TD)$. Then we get

$$\begin{aligned}
&\nabla_F W_{\mathrm{SVK}}(x, F)[D] \\
&= \lim_{h\to 0} \frac{W_{\mathrm{SVK}}(x, F + hD) - W_{\mathrm{SVK}}(x, F)}{h} \\
&= \lim_{h\to 0} \frac{\mathbf{H}E(F + hD) : E(F + hD) - \mathbf{H}E(F) : E(F)}{2h} \\
&= \lim_{h\to 0} \frac{h\mathbf{H}E(F) : (D^TF + F^TD + hD^TD)}{2h} \\
&= \mathbf{H}E(F) : F^TD,
\end{aligned}$$

where the symmetry properties of $\mathbf{H}$ are used. Since for the Frobenius inner product $(AB) : C = B : (A^TC)$ holds for any $A, B, C \in \mathbb{R}^{d\times d}$, we arrive at

$$\nabla_F W_{\mathrm{SVK}}(x, F)[D] = (F\,\mathbf{H}E(F)) : D.$$

Hence,

$$\frac{\partial}{\partial F} W_{\mathrm{SVK}}(x, F) = F\,\mathbf{H}E(F)) = F\,\Sigma = T.$$

□

The *isotropic* variant of a St. Venant–Kirchhoff material is described by only two degrees of freedom, such that,

$$W_{\mathrm{SVK}}(x, F) = \frac{\lambda}{2}\,\mathrm{tr}(E)^2 + \mu\,\|E\|_F^2 \tag{7.12}$$

for the *Lamé parameters* $\mu, \lambda > 0$ and the Green–St. Venant strain tensor E. This corresponds to (2.7) for small strains in the first part of this thesis.

7.4.2. Polyconvex Materials

In this thesis we also include polyconvex materials, which provide nonlinear stress-strain relations. Besides the deformation gradient F, we also consider the determinant $\det(F)$, and the cofactor matrix $\mathrm{cof}(F) := \det(F)F^{-T}$.

Definition 7.4.3 (Polyconvex Material)**.**

(i) In two dimensions we call a material W_{el} polyconvex, if

$$W_{el}(x, F) = W(x, F, \det(F))$$

for a convex $W : \Omega \times \mathbb{R}^{2\times 2} \times \mathbb{R} \to \mathbb{R}$.

(ii) In three dimensions we call a material W_{el} polyconvex, if

$$W_{el}(x, F) = W(x, F, \mathrm{cof}(F), \det(F))$$

for a convex $W : \Omega \times \mathbb{R}^{3\times 3} \times \mathbb{R}^{3\times 3} \times \mathbb{R} \to \mathbb{R}$.

Remarks 7.4.4.

(i) There are definitions of polyconvex materials for $d \notin \{2, 3\}$ *[22, page 238].*

(ii) St. Venant–Kirchhoff materials are not polyconvex [22, page 239].

We will investigate two families of polyconvex materials. Both of them are homogeneous materials, i.e., they only depend on the deformation gradient F and not on $x \in \Omega$. The first one is the three dimensional Ciarlet variant of the Mooney-Rivlin material [4, page 189]

$$W_{\mathrm{MR}}(F) := a\,\|F\|_F^2 + b\,\|\mathrm{cof}(F)\|_F^2 + c\det(F)^2 - d\ln(\det(F)) + e \tag{7.13}$$

with positive parameters $a, b, c, d \geq 0$ and $e \in \mathbb{R}$. This obviously matches the definition of polyconvexity. In two dimensions we set $b = 0$. Mooney–Rivlin materials are nonlinear. Nevertheless, we can choose the parameters according to a given St. Venant–Kirchhoff material, s.t., they correspond at small strains with given Lamé parameters λ and μ (7.12). We can approximate a Mooney–Rivlin energy by the system of equations

$$\begin{aligned} 3a + 3b + c + e &= 0 \\ 2a + 4b + 2c - d &= 0 \\ 2b + 2c &= \frac{\lambda}{2} \\ -2b - 2c + d &= \mu, \end{aligned} \tag{7.14}$$

which is taken from [4, p. 186]. Then, it is shown that

$$W_{\mathrm{MR}}(F) = W_{\mathrm{SVK}}(F) + \mathcal{O}(\|E\|_F^3)$$

for $E \to 0$. There is still one degree of freedom left open in the constants a, b, c, d, e. This degree of freedom can be used to control the weight of the term $c\det(F)^2 - d\ln(\det(F))$ that acts as a penalty term for $\det(F)$. By the second line $2a + 4b + 2c - d = 0$ it is assured that the minimum of $W_{\mathrm{MR}}(x, \gamma I)$ is obtained for $\gamma = 1$, hence the energy penalizes non-volume preserving behavior.

Another polyconvex material considered in this thesis is an *Ogden* material in the form (cf. [4, 22])

$$W_{\mathrm{Ogden}}(F) := \sum_{i=1}^{N} \alpha_i \,\mathrm{tr}\,(C^{p_i}) + \sum_{j=1}^{M} \beta_j \,\mathrm{tr}\,\mathrm{cof}\,(C^{q_j}) + V(\det(F)) \tag{7.15}$$

with $\alpha_i, \beta_j > 0$, and $p_i, q_j \geq 2$ for all appearing i, j, and the Cauchy–Green strain tensor $C = F^T F$. Moreover, $V : (0, \infty) \to \mathbb{R}$ is a convex function. Ogden materials are very flexible because a variety of factors and exponents can be selected. Furthermore, Ogden materials play a crucial role in the existence theory of solutions for finite-strain plasticity problems later in Section 8.6. Note that the Ciarlet variant of the Mooney–Rivlin material (7.13) is a special case of an Odgen material (7.15) with $M = N = 1$, as well as $p_1 = q_1 = 2$, and $V(\xi) = c\xi^2 - d\ln(\xi) + e$.

All so far considered elastic material models have in common that they can be expressed in terms of the principal values of the elastic strain. Without going into detail on this property we formulate another, weaker result.

Lemma 7.4.5. *All considered materials* (7.12)*,* (7.13) *and* (7.15) *are isotropic, i.e.,*

$$W_{el}(x, F) = W_{el}(x, FQ)$$

for any orthogonal matrix $Q \in \mathrm{SO}(d)$*.*

Proof. This is easily verified by the invariance of orthogonal similarity transformations of the Frobenius norm, the determinant and the trace of a matrix. ☐

8. Energetic Formulation of Rate-Independent Finite-Strain Plasticity

8.1. Plastic Energy and Back Stress

In this section we will take a closer look at the multiplicative decomposition $\mathbf{F} = \mathbf{F}_{\mathrm{el}}\mathbf{P}$ (7.4) and at the corresponding dual variables of the components. For the sake of simplicity we do not consider hardening in this section. The influence of kinematic hardening is discussed in Section 8.2. Moreover, dependence of the energy densities on $x \in \Omega$ is not considered. First of all, we recall that the hyperelastic materials from Section 7.4.1 depend on the elastic component $\mathbf{F}_{\mathrm{el}} = \mathbf{F}\mathbf{P}^{-1}$. To formally track the components, we define a plastic energy density

$$W_{\mathrm{pl}} : \mathrm{GL}(d)^{+} \times \mathrm{SL}(d) \to \mathbb{R},\ W_{\mathrm{pl}}(F, P) \coloneqq W_{\mathrm{el}}(FP^{-1}) = W_{\mathrm{el}}(F_{\mathrm{el}}).$$

Again a main assumption is frame-indifference (cf. Section 7.4.1), such that, there is a $\hat{W}_{\mathrm{pl}} : \mathrm{GL}(d)^{+} \times \mathrm{SL}(d) \to \mathbb{R}$ with

$$W_{\mathrm{pl}}(F, P) = \hat{W}_{\mathrm{pl}}(F^T F, P) = \hat{W}_{\mathrm{pl}}(C, P).$$

From $C = F^T F = P^T F_{\mathrm{el}}^T F_{\mathrm{el}} P$ we can recover the elastic Green strain tensor $C_{\mathrm{el}} \coloneqq F_{\mathrm{el}}^T F_{\mathrm{el}}$ by

$$C_{\mathrm{el}} = P^{-T} C P^{-1}.$$

With this at hand we have

$$\hat{W}_{\mathrm{pl}}(C, P) = \hat{W}_{\mathrm{el}}\Big(P^{-T} C P^{-1}\Big) = \hat{W}_{\mathrm{el}}(C_{\mathrm{el}})$$

for the frame-indifferent elastic energy $\hat{W}_{\mathrm{el}}$.

From $\hat{W}_{\mathrm{pl}}$ several quantities can be deduced. Since plastic deformation is not related to stress, the second Piola–Kirchhoff stress (7.11) is given by

$$\Sigma = 2\,\frac{\partial}{\partial C}\hat{W}_{\mathrm{pl}}(C, P). \tag{8.1}$$

To express Σ in terms of derivatives of $\hat{W}_{\mathrm{el}}$ we consider the directional derivative in C in the direction D, which fulfills

$$\begin{aligned}\langle \Sigma, D\rangle &= 2\,\nabla\hat{W}_{\mathrm{el}}\Big(P^{-T} C P^{-1}\Big)[P^{-T} D P^{-1}] \\ &= 2\ \nabla\hat{W}_{\mathrm{el}}\Big(P^{-T} C P^{-1}\Big) : \Big(P^{-T} D P^{-1}\Big) \\ &= 2\ \Big(P^{-1}\nabla\hat{W}_{\mathrm{el}}\Big(P^{-T} C P^{-1}\Big)P^{-T}\Big) : D.\end{aligned}$$

Hence, we identify

$$\Sigma = 2\, P^{-1}\nabla\hat{W}_{\mathrm{el}}(C_{\mathrm{el}})P^{-T}. \tag{8.2}$$

The corresponding negative derivative of $\hat{W}_{\mathrm{pl}}$ in P is called the *back stress*

$$Q \coloneqq -\frac{\partial}{\partial P}\hat{W}_{\mathrm{pl}}(C,P). \tag{8.3}$$

To rewrite this in the elastic energy $\hat{W}_{\mathrm{el}}$ we first note that for the matrix inverse $i(P) \coloneqq P^{-1}$ we have in direction D

$$I = i(P)P \;\Rightarrow\; 0 = \nabla i(P)[D]P + i(P)D \;\Rightarrow\; i(P)[D] = -P^{-1}DP^{-1}. \tag{8.4}$$

Then we have by symmetry of C and $\nabla\hat{W}_{\mathrm{el}}$

$$\begin{aligned}
\langle Q, D\rangle &= -\,\nabla\hat{W}_{\mathrm{el}}(C_{\mathrm{el}})\left[\nabla i(P)[D]^T CP^{-1} + P^{-T}C\nabla i(P)[D]\right]\\
&= -\,\nabla\hat{W}_{\mathrm{el}}(C_{\mathrm{el}}) : \left(-P^{-T}D^T C_{\mathrm{el}} - C_{\mathrm{el}}DP^{-1}\right)\\
&= -\,\nabla\hat{W}_{\mathrm{el}}(C_{\mathrm{el}}) : \left(-2\, C_{\mathrm{el}}DP^{-1}\right)\\
&= \left(2C_{\mathrm{el}}\nabla\hat{W}_{\mathrm{el}}(C_{\mathrm{el}})P^{-T}\right) : D.
\end{aligned}$$

Therefore, we identify

$$Q = 2C_{\mathrm{el}}\nabla\hat{W}_{\mathrm{el}}(C_{\mathrm{el}})P^{-T}. \tag{8.5}$$

In summary, we have an explicit form of the second Piola–Kirchhoff stress tensor Σ associated to the strain tensor C, and a corresponding back-stress tensor Q associated to the plastic strain P. Both of these stress-like variables play an important role in the formulation of yield stresses, which lead to plastic effects.

8.2. Kinematic Hardening

To derive an easy model of kinematic hardening similar to the small-strain case is not trivial and requires a careful identification of the used deformation mappings [15, Section 5.5]. To avoid these technical challenges, we rely on the formulation given in [22, Example 4.2.2]. There, kinematic hardening is given by an energetic formulation as a hardening energy density

$$W_{\mathrm{hd}} : \mathrm{SL}(d) \times \mathbb{R}^{d\times d\times d} \to \mathbb{R},\;\; (P,Z) \mapsto W_{\mathrm{hd}}(P,Z)$$

on the space of plastic strains, and a tensor-valued (possibly neglected) second argument. For a rigorous existence theory pointwise evaluation of the plastic strain are not sufficient since spatial gradient terms of the global plastic strain function $\mathbf{P} : \Omega \to \mathrm{SL}(d)$ are necessary (see Theorem 8.6.1). Hence, we will use the global formulation of a hardening energy

$$\mathcal{W}_{\mathrm{hd}}(\mathbf{P}) \coloneqq \int_\Omega W_{\mathrm{hd}}(\mathbf{P}(x), \nabla\mathbf{P}(x))\, dx,$$

where $(\nabla \mathbf{P})_{ijk} \coloneqq \partial_{x_k} \mathbf{P}_{ij}$ describes the componentwise derivative. In this thesis we will follow [22, Example 4.2.2] and use

$$W_{\text{hd}}(P, Z) \coloneqq k_1 \left\| P \right\|_F^{p_1} + k_2 \left\| Z \right\|_F^{p_2}$$

for non-negative hardening parameters k_1, k_2, and exponents $p_1, p_2 > 1$. The Frobenius norm of the 3-tensor Z is defined componentwise by

$$\| Z \|_F \coloneqq \left(\sum_{i,j,k=1}^{d} Z_{ijk}^2 \right)^{\frac{1}{2}} .$$

Note that for $k_2 = 0$ and $p_1 = 2$ we have the linear kinematic hardening functional from the small-strain case, cf. the middle term in (3.15). This global hardening term enriches the plastic strain energy from Section 8.1, such that, we consider the global plastic energy

$$\mathcal{W}_{\text{pl}}(\mathbf{F}, \mathbf{P}) \coloneqq \mathcal{W}_{\text{el}}(\mathbf{F}\mathbf{P}^{-1}) + \mathcal{W}_{\text{hd}}(\mathbf{P}). \tag{8.6}$$

The hardening term changes the nature of the plastic back-stress Q from Section 8.1. Since the hardening depends on the plastic strain $\mathbf{P}$, an additional plastic strain information enters Q. We do not want to formalize the effect of the plastic strain on Q, since it requires a technical investigation of what we understand by the partial derivative $\frac{\partial}{\partial P} W_{\text{hd}}$ in a pointwise sense, while the hardening term depends globally on the spatial gradient $\nabla \mathbf{P}$. Nevertheless, in the small-strain setting ($k_2 = 0$ and $p_1 = 2$) we get an additional term $-k_1 P$ for Q, which is consistent to the linear kinematic hardening in the yield function, cf. (2.11).

Regarding a yield function: Kinematic hardening is the movement of the elastic region $\mathbb{E}$ of admissible stresses. In both [15, Section 5.2] and [21, Section 2] a yield function[1]

$$\Phi : \mathbb{R}^{d \times d} \to \mathbb{R}, \ S \mapsto \Phi(S)$$

is proposed, such that the convex elastic region is (as in the small-strain case) given by

$$\mathbb{E} \coloneqq \{ S \in \mathbb{R}^{d \times d} : \ \Phi(S) \leq 0 \}.$$

For now it not clear what the variable S is precisely. The idea is to use a stress-like quantity that does only depend on elastic quantities, and, furthermore, is related to a partial derivative of the plastic energy W_{pl} with respect to P. This enables us to formulate a plastic strain driven energetic formulation later. In [15, Section 2.2] it is suggested that we consider

$$S \coloneqq QP^T$$

for the back-stress Q given in (8.3)[2]. This choice of S can be justified by two reasons. First, note that from (8.5) we see that in the case of no hardening ($W_{\text{hd}} \equiv 0$) that

$$QP^T = 2C_{\text{el}} \nabla \hat{W}\text{el}(C_{\text{el}})$$

---

[1] The domain $\mathbb{R}^{d \times d}$ is used for simplification. See [15, Section 2.2] for a general statement.

[2] However, in [15] Q is defined slightly different, but equivalent, such that $S = P^T Q$.

depends only on elastic quantities. This is reasonable since stresses depend only on elastic processes as long as the elastic region is fixed. The second reason is of a more technical nature: Yield stresses are usually determined by properties of the eigenvalues of S. If we consider S as a linear operator between a domain and a range space, then eigenvalues are only meaningful if both of these spaces are the same. The deformation field $\mathbf{y}$ is a function from the reference domain Ω to the deformed domain $\Omega^{\mathbf{y}}$.

We follow the ideas of [15, Section 5] and call $\mathcal{A}$ the range manifold of the deformation in Ω, and $\mathcal{B}$ the range manifold of the deformation in $\Omega^{\mathbf{y}}$. Of course, we have in both cases $\mathcal{A} = \mathcal{B} = \mathbb{R}^d$ in our setting, but for technical reasons we stick here to the general variant. Therefore, for each $x \in \Omega$ the deformation gradient $F = \nabla \mathbf{y}(x)$ is a linear mapping $T\mathcal{A} \to T\mathcal{B}$ between the tangential spaces. By the multiplicative decomposition $F = F_{\text{el}}P$ this scheme is enriched by an intermediate manifold $\mathcal{C}$, such that $P : T\mathcal{A} \to T\mathcal{C}$, and $F_{\text{el}} : T\mathcal{C} \to T\mathcal{B}$. In this setting the back-stress Q (8.3) is a mapping $T\mathcal{A}^* \to T\mathcal{C}^*$ in the dual spaces. Hence, to make domain and range of S equal, we need another mapping $T\mathcal{C}^* \to T\mathcal{A}^*$. This is a property of the adjoint operator P^T. Then $S = QP^T$ is an operator of the type $T\mathcal{C}^* \to T\mathcal{C}^*$.

In summary, we have a yield function depending on the back stress Q and the plastic strain P, such that $S = QP^T$ is a stress-like elastic variable that determines the admissible stresses. The explicit form of the yield function is not stated in this thesis for the finite-strain case, since it opens the door for technical complications which are not discussed here. However, we need the yield function for analytical purposes later.

8.3. Plastic Flow

Similar to the small-strain case, the plastic evolution $\dot{\mathbf{P}}$ is defined via plastic flow rules, cf. (2.10). Therefore, we use a yield function $\Phi = \Phi(S)$ for the stress-like variable $S = QP^T$, as discussed in the previous section. Recall that the current plastic strain $\mathbf{P} = \mathbf{P}(t)$ at time t acts as a constant problem parameter, whereas we consider only the evolution $\dot{\mathbf{P}}$ as the plastic quantity of interest, cf. [15]. This idea is used for the yield function Φ which may depend as well on the fixed $\mathbf{P}$. By Lemma A.3.5, we can express the evolution of the plastic strain $\mathbf{P}$ by a trace-free matrix $\mathbf{B}$, such that $\dot{\mathbf{P}} = \mathbf{B}\mathbf{P}$ holds. Hence, the evolution $\dot{\mathbf{P}}$ can described by $\mathbf{B}$ only, since $\mathbf{P}$ itself is known. From now on we switch to local quantities $B = \mathbf{B}(x)$, etc. Following [21, Section 2] and [15, Section 2.2], the flow rule in the finite-strain setting is given by

$$B = \dot{P}P^{-1} = \lambda \partial \Phi(S)$$

where plastic evolution occurs if and only if a critical stress is reached, i.e.,

$$\lambda \geq 0, \quad \Phi(S) \leq 0, \quad \lambda \Phi(S) = 0$$

holds simultaneously. Therefore, we can define the *elastic region* $\mathbb{E}$ from the yield function Φ by

$$\mathbb{E} := \{S \colon \Phi(S) \leq 0\}.$$

The following ideas are taken from [21, Section 2]. Similar to Section 3.2, we define a convex polar function – or *dissipation* function – corresponding to $\mathbb{E}$ by

$$\Delta : \mathrm{sl}(d) \to \mathbb{R},\ \Delta(B) := \sup_{S\in\mathbb{E}} \langle B, S\rangle, \tag{8.7}$$

on the tangential space $\mathrm{sl}(d) = T_I\,\mathrm{SL}(d)$. Note that $B = \dot{P}P^{-1} \in \mathrm{sl}(d)$. Applying the convex analysis Theorem 3.2.2 we have again that

$$B \in N_{\mathbb{E}}(S) \quad \Leftrightarrow \quad S \in \partial\Delta(B). \tag{8.8}$$

We want to extract Q from $S = QP^T$ in order to have derivative information with respect to P, cf. (8.3). To this end, we define

$$\hat{\Delta}(P, \dot{P}) := \Delta(B) = \Delta(\dot{P}P^{-1}) \tag{8.9}$$

to separate the components of B.

Remark 8.3.1. *Note that $\hat{\Delta}$ is rate-independent in the sense of Definition 2.2.3, i.e., for a time-scaled $Q(t) := P(\alpha t)$ we have $\dot{Q} = \alpha\dot{P}$, and therefore $\hat{\Delta}(Q, \dot{Q}) = \alpha\hat{\Delta}(P, \dot{P})$ for every $\alpha \in \mathbb{R}$. Hence, for a time-scaled input we have a scaled output without changing the solution path $t \to P(t)$.*

A computation similar to the one in Section 8.1 leads to

$$\frac{\partial}{\partial\dot{P}}\hat{\Delta}(P, \dot{P}) = \partial\Delta(\dot{P}P^{-1})P^{-T}.$$

Therefore, from $S = QP^T$ and (8.8) we conclude that

$$S \in \partial\Delta(\dot{P}P^{-1}) \quad \Leftrightarrow \quad Q = SP^{-T} \in \frac{\partial}{\partial\dot{P}}\hat{\Delta}(P, \dot{P}) \tag{8.10}$$

is an equivalent form of the flow rule. Together with the original definition of Q in (8.3), we conclude that

$$\boxed{0 \in \frac{\partial}{\partial P}\hat{W}_{\mathrm{pl}}(C, P) + \frac{\partial}{\partial\dot{P}}\hat{\Delta}(P, \dot{P})} \tag{8.11}$$

is the final form of the plastic flow in this section.

8.4. Global Formulation

To find a global form of the different components of the sections above a small review of the necessary ingredients is given. First of all, the unknown quantities over the domain Ω are given by the deformation field

$$\mathbf{y} : \Omega \to \mathbb{R}^d$$

which assigns each particle $x \in \Omega$ the new position $\mathbf{y}(x)$, and hence defines the deformed domain by $\Omega^{\mathbf{y}} := \mathbf{y}(\Omega)$. Note that $\mathbf{y}$ is also implicitly time-dependent. Like in the small-strain setting, we do not consider any higher derivatives above first-order information, such that internal inertial forces are not considered. Time is a model parameter defining the path from the initial state to the deformed domain. This path is also called *homotopy*.

The other main variable in our model is the field of plastic strains

$$\mathbf{P} : \Omega \to \mathrm{SL}(d)$$

that keeps track of the history of the plastic effects. An implicit time-dependence is assumed too, of course. From $\mathbf{y}$ and $\mathbf{P}$ the field of elastic strains $\mathbf{F}_{\mathrm{el}}$ can be recovered, and hence the Piola–Kirchhoff stresses $\mathbf{T}$ and $\mathbf{\Sigma}$. Additional internal variables of the plastic strain are not necessary since only kinematic hardening is considered in this part of the thesis. Note that the internal variable of the hardening effects are identified by the plastic strain itself, cf. (2.18). Therefore, the system of unknowns is complete.

So far the energy counting in all presented effects is defined as the plastic energy

$$(\mathbf{y}, \mathbf{P}) \mapsto \mathcal{W}_{\mathrm{pl}}(\nabla \mathbf{y}, \mathbf{P})$$

in (8.6). External load controlling the deformation process are not mentioned yet. Therefore, we introduce a volumetric force density $\mathbf{b}$ and a surface force density $\mathbf{s}$

$$\mathbf{b} : \Omega \to \mathbb{R}^d, \quad \mathbf{s} : \partial\Omega \to \mathbb{R}^d,$$

acting on the test body and its boundary, respectively. Similar to Section 2.2 in the first part, the second Piola–Kirchhoff stress $\mathbf{\Sigma}$ (8.1) fulfills

$$-\operatorname{div} \mathbf{\Sigma} = \mathbf{b} \tag{8.12}$$

in Ω. Note that the divergence is understood componentwise. To include the forces $\mathbf{b}$ and $\mathbf{s}$ formally into the energetic model, we define the load functional

$$\mathbf{l}(\mathbf{y}) := \int_\Omega \mathbf{b} : \mathbf{y}\, dx + \int_{\partial\Omega} \mathbf{s} : \mathbf{y}\, dS$$

acting on the deformation field. To underline the linear behavior and to highlight the time-dependence, it is sometimes written as a linear functional

$$\langle \mathbf{l}(t), \mathbf{y} \rangle := \int_\Omega \mathbf{b}(t) : \mathbf{y}\, dx + \int_{\partial\Omega} \mathbf{s}(t) : \mathbf{y}\, dS. \tag{8.13}$$

Now we have everything at hand to define the global energy

$$\boxed{\mathcal{E}(\mathbf{y}, \mathbf{P}) := \mathcal{W}_{\mathrm{el}}(\nabla \mathbf{y} \mathbf{P}^{-1}) + \mathcal{W}_{\mathrm{hd}}(\mathbf{P}) + \langle \mathbf{l}(t), \mathbf{y} \rangle.} \tag{8.14}$$

We can directly derive a first lemma from this energy statement. It arises from the independence of all plastic components on $\mathbf{y}$, and from the direct connection between the stress $\mathbf{\Sigma}$, the hyperelastic energy (7.11) for hyperelastic materials, and the force equilibrium (8.12).

Lemma 8.4.1 (Elastic Equilibrium [21, 22])**.** *Let the global energy $\mathcal{E}$ in* (8.14) *depend smoothly on* $\mathbf{y}$*. Then a plastic material with a hyperelastic elastic energy fulfills*

$$\nabla_{\mathbf{y}}\mathcal{E}(\mathbf{y},\mathbf{P}) = 0. \tag{8.15}$$

The second global statement is the integrated version of the plastic flow (8.11) from the previous section. The pointwise local inclusion of zero can be integrated, such that

$$\boxed{0 \in \partial_{\mathbf{P}}\mathcal{E}(\mathbf{y},\mathbf{P}) + \partial_{\dot{\mathbf{P}}}\hat{\mathbf{\Delta}}(\mathbf{P},\dot{\mathbf{P}})} \tag{8.16}$$

holds for the (time-dependent) plastic strain field $\mathbf{P}$. In [21, Section 2] it is shown that a time integration of the flow rule (8.16) leads to

$$\boxed{\mathcal{E}(\mathbf{y}(0),\mathbf{P}(0)) + \int_0^t \langle \dot{\mathbf{l}}(s), \mathbf{y}(s)\rangle\, ds = \mathcal{E}(\mathbf{y}(t),\mathbf{P}(t)) + \int_0^t \hat{\mathbf{\Delta}}(\mathbf{P}(s),\dot{\mathbf{P}}(s))\, ds,} \tag{E}$$

which is called *energy balance.*

Remark 8.4.2. *The energy balance highlights the following fact. The time-dependent derivative of the load term* $\dot{\mathbf{l}}(s)$ *is a physical work term if integrated in time. Nevertheless, there is a gap between the energy* $\mathcal{E}$ *at the initial state* $t = 0$ *to an arbitrary time point. This gap is given by the integral over* $\hat{\mathbf{\Delta}}$*. The portion of energy that is* lost *is called dissipated energy. If no plastic evolution occurs (*$\dot{\mathbf{P}} \equiv 0$*), and hence* $\hat{\mathbf{\Delta}} \equiv 0$*, then no energy is dissipated.*

We can derive another statement from (8.15)&(8.16). They are equivalent in terms of the directional derivative statements

$$\begin{aligned} \nabla_{\mathbf{y}}\mathcal{E}(\mathbf{y},\mathbf{P})[\mathbf{z}] &= 0 \\ \nabla_{\mathbf{P}}\mathcal{E}(\mathbf{y},\mathbf{P})[\mathbf{D}] + \hat{\mathbf{\Delta}}(\mathbf{P},\mathbf{D}) &\geq 0 \end{aligned} \tag{8.17}$$

for each compatible direction $(\mathbf{z},\mathbf{D})$ that does not violate possible boundary constraints on the deformation field. Whereas the first line is obvious, the second line is not. To verify the second line, we notice that $\hat{\mathbf{\Delta}}$ is 1-homogeneous in the second component (that is precisely the rate-independence from Remark 8.3.1). From [21, Section 2] we have

$$\langle \partial_{\dot{\mathbf{P}}}\hat{\mathbf{\Delta}}(\mathbf{P},\dot{\mathbf{P}}), \mathbf{D}\rangle = \hat{\mathbf{\Delta}}(\mathbf{P},\mathbf{D})$$

for every compatible $\mathbf{D}$.

A closer look at (8.17) reveals the following. For every fixed $\mathbf{P}$ the first condition states that $\mathbf{y}$ is a critical point. This is a necessary condition for a global minimizer of $\mathcal{E}$ by assuming that there is a global minimizer. On the other hand, for a fixed $\mathbf{y}$ the second condition of (8.17) states that for each direction $\mathbf{D}$ the decrease of the energy in this direction must not be larger than the "distance" of $\mathbf{P}$ and $\mathbf{D}$ given by $\hat{\mathbf{\Delta}}$. Note that the length of $\mathbf{D}$ does not matter, since both terms are 1-homogeneous in $\mathbf{D}$.

The form (8.17) holds only information for a fixed point in time. To transform this onto a solution notion along the whole time axis $[0, T]$, an integrated version of $\hat{\mathbf{\Delta}}$ needs to be introduced. As already mentioned in Remark 8.4.2, this is the total dissipation of the plastic strain solution path $t \to \mathbf{P}(t)$. Formally we define the distance between two plastic strain states $P_1, P_2 \in \mathrm{SL}(d)$ by

$$D(P_1,P_2) := \inf\left\{\int_0^1 \hat{\Delta}(P(s),\dot{P}(s)) :\ P \in \mathcal{C}^1([0,1];\mathrm{SL}(d)),\ P(0)=P_1,\ P(1)=P_2\right\} \tag{8.18}$$

as the length of the shortest differentiable path from P_1 to P_2 in $\mathrm{SL}(d)$ using the metric $\hat{\Delta}$. Since $\hat{\Delta} : \mathrm{SL}(d) \times T\,\mathrm{SL}(d) \to \mathbb{R}$ measures objects on $\mathrm{sl}(d)$ (the tangential space at the unit matrix), it is also called a Finsler metric. The global version of this distance between two plastic strain fields $\mathbf{P}_1$ and $\mathbf{P}_2$ is naturally given by

$$\mathcal{D}(\mathbf{P}_1,\mathbf{P}_2) := \int_\Omega D(\mathbf{P}_1(x),\mathbf{P}_2(x))\,dx. \tag{8.19}$$

Note that $\mathcal{D}$ defines a quasimetric in the space of plastic strain fields, since $\mathcal{D}(\mathbf{P}_1,\mathbf{P}_2) = 0 \Leftrightarrow \mathbf{P}_1 \equiv \mathbf{P}_2$, and the triangle inequality holds as an immediate consequence of the definition by shortest paths. It is only a quasimetric since we cannot assume symmetry, though. With a global distance function at hand, we can formulate a variant of (8.17) that detaches from the local point of view of the current $\mathbf{P}$. We formulate the global energy stability

$$\mathcal{E}(\mathbf{y}(t),\mathbf{P}(t)) \le \mathcal{E}(\tilde{\mathbf{y}},\tilde{\mathbf{P}}) + \mathcal{D}(\mathbf{P}(t),\tilde{\mathbf{P}}) \tag{S}$$

for all deformation fields $\tilde{\mathbf{y}}$ compatible to boundary values, and all plastic strain fields $\tilde{\mathbf{P}}$. This complies to the necessary minimization conditions of (8.17).

Remarks 8.4.3.

(i) If the process is purely elastic, we have $\mathcal{D} \equiv 0$, and then (S) *is a minimization problem in* $\mathbf{y}$.

(ii) The definition of (S) *can be relaxed by allowing the distance of each path from* $\mathbf{P}(t)$ *to* $\tilde{\mathbf{P}}$*, and not just the shortest one.*

At this point, we are able to define the concept of solutions of the finite-strain plasticity problem. Since the presented system is energy driven, the solutions are called energetic solutions.

Definition 8.4.4 (Energetic Solution). *Let* $\mathbf{y} : [0,T] \times \Omega \to \mathbb{R}^3$ *be a deformation field fulfilling possible boundary conditions on* $\partial\Omega$*, and* $\mathbf{P} : [0,T] \times \Omega \to \mathrm{SL}(d)$ *be a plastic strain field. Moreover, let* $\mathbf{y}$ *and* $\mathbf{P}$ *fulfill both the energy balance* (E) *and the global stability criterion* (S) *for all* $t \in [0,T]$*. Then we call the path* $t \mapsto (\mathbf{y}(t,\cdot),\mathbf{P}(t,\cdot))$ *an energetic solution.*

8.5. Von Mises Dissipation

So far the dissipation function (8.7) was never specified explicitly. A von Mises type yield function for a norm $\|\cdot\|$ on the stress space is proposed in [15, Section 4] in the form

$$\Phi(S) := \|\mathrm{dev}(S)\| - \sigma_0$$

for finite-strain mechanics, similar to the small-strain case in Section 2.4.1. The operator dev denotes the deviator of a matrix. In [21, page 65] the corresponding dissipation function (8.7) is explicitly given by

$$\Delta(\xi) := \left(\sigma_0^2 \|\mathrm{sym}(\xi)\|_F^2 + \sigma_1^2 \|\mathrm{anti}(\xi)\|_F^2\right)^{\frac{1}{2}}, \tag{8.20}$$

with the symmetric $\mathrm{sym}(\xi) := \frac{1}{2}\left(\xi + \xi^T\right)$ and the antisymmetric $\mathrm{anti}(\xi) := \frac{1}{2}\left(\xi - \xi^T\right)$ parts. This defines a metric on $\mathrm{sl}(d)$ if $\sigma_0, \sigma_1 > 0$.

However, the case of $\sigma_1 = 0$ is of special interest later on. Unfortunately, in this case

$$\Delta_{\mathrm{sym}}(\xi) := \sigma_0 \|\mathrm{sym}(\xi)\|_F \tag{8.21}$$

is not a metric on $\mathrm{sl}(d)$, since all antisymmetric, trace-free matrices ξ in $\mathrm{sl}(d)$ fulfill $\Delta(\xi) = 0$. In the symmetric case there is an explicit form of the dissipation distance (8.18).

Theorem 8.5.1 (cf. [21, Corollary 6.2]). *For $P_1, P_2 \in \mathrm{SL}(d)$, and $\Delta = \Delta_{sym}$, as in* (8.21)*, we have for the dissipation distance* (8.18)

$$D(P_1, P_2) = \sigma_0 \left\| \log\left(\delta P\, \delta P^T\right)^{\frac{1}{2}} \right\|_F, \qquad \text{where } \delta P := P_2 P_1^{-1}.$$

Another result of the von Mises dissipation is the invariance under matrix multiplication from the right. We show the special case of our interest (multiplication with P_1^{-1})

Lemma 8.5.2. *For $P_1, P_2 \in \mathrm{SL}(d)$ and Δ from* (8.20)*, we have*

$$D(P_1, P_2) = D(I, \delta P)$$

for the corresponding increment defined via $\delta P = P_2 P_1^{-1}$.

Proof. In order to prove the equality of the above dissipation distances, we consider the original definition of D in (8.18), and introduce the sets of connecting paths

$$\Pi := \left\{P \in \mathcal{C}^1([0,1], \mathrm{SL}(d)) \colon P(0) = P_1, P(1) = P_2\right\} \quad \text{and}$$
$$\delta\Pi := \left\{\delta P \in \mathcal{C}^1([0,1], \mathrm{SL}(d)) \colon P(0) = I, P(1) = \delta P = P_2 P_1^{-1}\right\}.$$

With these definitions at hand, the transformation $T \colon \Pi \to \delta\Pi$, $T(P) := P P_1^{-1}$ is a bijection between these sets. We observe for $P \in \Pi(P_1, P_2)$, the corresponding $T := T(P) \in \delta\Pi(P_1, P_2)$ and any $s \in [0,1]$ that

$$\dot{P}(s) P(s)^{-1} = \left(\dot{P}(s) P_1^{-1}\right)\left(P(s) P_1^{-1}\right)^{-1} = \dot{T}(s) T(s)^{-1}$$

which directly implies equal length of the transformed path induced by the Finsler metric. ☐

From the original definition of D in (8.18) it is clear that D is only a distance function of $\mathrm{SL}(d)$ if, and only if, the dissipation function Δ (8.7) is a metric on $\mathrm{sl}(d)$.

In order to keep the distance properties by using the general linear group $\mathrm{SL}(d)$ including antisymmetric components (cf. Section 7.2.2), we assume that $\sigma_1 > 0$, in the limit sense that $\sigma_1 \to 0$. Then we may still utilize Theorem 8.5.1 in this case and accept the resulting model error.

Another possibility is to consider no plastic spin, and thus relying on the symmetric plastic range $\mathrm{SL}(d)^+_{\mathrm{sym}}$ (7.8). As mentioned in [22, Remark 4.2.9], we have to redefine the split of the plastic dissipation functional (8.9) by

$$\hat{\Delta}(P, \dot{P}) \coloneqq \Delta\left(P^{-\frac{1}{2}}\dot{P}P^{-\frac{1}{2}}\right),$$

and arrive at another dissipation distance D^+_{sym} defined by the shortest path in $\mathrm{SL}(d)^+_{\mathrm{sym}}$. That special case is analyzed in depth in Section 9.2.1. Nevertheless, the computations of Section 8.3 yield equivalent results in that case.

8.6. Existence Theory

In this section we investigate sufficient analytical criteria for the existence of energetic solutions (in terms of Definition 8.4.4) for the global rate-independent finite strain problem (E)&(S). For convenience, we recall the problem here. We denote by $\mathcal{Y}$ the function space of deformation functions $\mathbf{y} : \Omega \times [0, T] \to \mathbb{R}^d$ with respect to possible boundary values, and by $\mathcal{P}$ the function space of plastic strains $\mathbf{P} : \Omega \times [0, T] \to \mathrm{SL}(d)$. The particular definition of the solution space $\mathcal{Q} \coloneqq \mathcal{Y} \times \mathcal{P}$ is given later. The task is to find a $\mathbf{q} = (\mathbf{y}, \mathbf{P}) : \Omega \times [0, T] \to \mathcal{Q}$, such that,

$$\begin{aligned} \mathcal{E}(t, \mathbf{y}(t), \mathbf{P}(t)) &\le \mathcal{E}(t, \tilde{\mathbf{y}}, \tilde{\mathbf{P}}) + \mathcal{D}(\mathbf{P}(t), \tilde{\mathbf{P}}) \qquad \forall(\tilde{\mathbf{y}}, \tilde{\mathbf{P}}) \in \mathcal{Q} \\ \mathcal{E}(0, \mathbf{y}(0), \mathbf{P}(0)) + \int_0^t \langle \dot{\mathbf{l}}(s), \mathbf{y}(s)\rangle \, ds &= \mathcal{E}(t, \mathbf{y}(t), \mathbf{P}(t)) + \int_0^t \hat{\mathbf{\Delta}}(\mathbf{P}(s), \dot{\mathbf{P}}(s)) \, ds \end{aligned} \tag{S\&E}$$

for each point in time $t \in [0, T]$. Note that we highlight the explicit time-dependency of the energy $\mathcal{E} : [0, T] \times \mathcal{Q} \to \mathbb{R} \cup \{\infty\}$, which is most of the times implicitly assumed.

The main existence result for en energetic solution of the above stated problem is based on Theorem 4.2.1 in [22]. A list of conditions is given there, which is summarized below.

First of all, we need to specify the solution space $\mathcal{Q} = \mathcal{Y} \times \mathcal{P}$. For the space of deformation yields $\mathcal{Y}$ we simply assume the first-order Sobolev space

$$\mathbf{y}(t) \in \mathcal{Y} \coloneqq W^{1,p_{df}}_{\Gamma}(\Omega, \mathbb{R}^d) \tag{8.22}$$

with $d < p_{\mathrm{df}} < \infty$, and a Dirichlet boundary $\Gamma \subset \partial\Omega$ with positive measure. Furthermore, there are some prescribed boundary values

$$\mathbf{y}|_\Gamma \equiv \mathbf{g}.$$

Of course, we need weak differentiability to be able to compute the deformation gradient $\mathbf{F} = \nabla \mathbf{y}$. The high regularity p_{df} simply arises from polyconvex materials (cf. Section 7.4.2) to be able to integrate the determinant on $\mathbf{F}$ which is a polynomial of degree d. The plastic strain functions are given by

$$\mathbf{P}(t) \in \mathcal{P} := W^{1,p_{gr}}(\Omega, \mathrm{SL}(d)) \cap L^{p_{pl}}(\Omega, \mathrm{SL}(d)) \tag{8.23}$$

with regularity parameters $1 < p_{\mathrm{gr}}, p_{\mathrm{pl}} < \infty$. Explicit conditions on these parameters are to be given later.

The conditions on the energies begin with the hyperelastic elastic energy W_{el} from Sections 7.4.1 and 7.4.2. Moreover, W_{el} is elliptic in the sense that

$$W_{el}(F_{\mathrm{el}}) \geq a + c \, \|F_{el}\|_F^{p_{el}} \tag{8.24}$$

for a growth rate $p_{\mathrm{el}} > 1$, and constants $a \in \mathbb{R}, c > 0$. The additional kinematic hardening energy W_{hd} from Section 8.2 takes the form

$$W_{\mathrm{hd}}(P, Z) := k_1 \, \|P\|_F^{p_{\mathrm{pl}}} + k_2 \, \|Z\|_F^{p_{\mathrm{gr}}} \tag{8.25}$$

with $k_1, k_2 > 0$, and $p_{\mathrm{pl}}, p_{\mathrm{gr}} > 1$ This completes the assumption on the plastic energy parts, as given in (8.6).

The parameters p_{el} and p_{pl} of the energetic quantities need to fulfill the inequality

$$\frac{1}{p_{\mathrm{el}}} + \frac{1}{p_{\mathrm{pl}}} < \frac{1}{d} \tag{8.26}$$

in order to prepare an application of Hölder techniques in the proof of the theorem below.

The other main component is dissipation. In Section 8.4 an abstract dissipation distance $\mathcal{D}$ is introduced, and a specific choice of the von Mises dissipation is given in Section 8.5. We assume that $\mathcal{D}$ is given by a von Mises distance, and that it is a quasidistance in the classical sense, i.e., $\mathcal{D}$ fulfills the triangle inequality, and, furthermore, $\mathcal{D}(\mathbf{P}_1, \mathbf{P}_2) = 0$ if and only if $\mathbf{P}_1 \equiv \mathbf{P}_2$. Symmetry of $\mathcal{D}$ is not needed.

As far as the external influences are concerned, we rely on the load functional $\mathbf{l}$, as given in (8.13). We assume that both appearing force fields $\mathbf{s}$ and $\mathbf{b}$ are sufficiently smooth, cf. [22, Remark 4.2.5].

Theorem 8.6.1 ([22, Theorem 4.2.1])**.** *Under the assumptions stated above in this Section there exists an energetic solution* $(\mathbf{y}, \mathbf{P})\colon [0, T] \to \mathcal{Y} \times \mathcal{P}$ *to the rate-independent system* (S&E)

The existence theorem above requires a long list of conditions. As discussed in the following remarks, some of them may not match practical needs.

Remarks 8.6.2.

(i) Existence of an energetic solution is already ensured under weaker, but more technical conditions. We avoid these by restricting to certain polyconvex material models and von Mises dissipation in this thesis. See the original source [22] for more details.

(ii) *The existence Theorem 8.6.1 requires pretty strong regularity and growth assumptions on the material, especially the inequality* (8.26)*. Classical kinematic hardening (see Part I, Chapter 3) yields a hardening energy with* $p_{pl} = 2$*, that already contradicts* (8.26)*. Say, we use twice the regularity* $p_{pl} = 4$ *for the hardening energy, then we need* $p_{el} > 4$ *in two dimensions (*$d = 2$*), and even* $p_{el} > 12$ *for* $d = 3$*.*

(iii) *The Mooney-Rivlin energy* (7.13) *has* $p_{el} = 2$ *and is therefore not applicable for Theorem 8.6.1.*

(iv) *For the Ogden energy* (7.15)*,* p_{el} *is given by the lowest exponent* p_i *with* $\alpha_i \neq 0$*. Hence, by (ii) only high exponents are possible in order to comply with Theorem 8.6.1.*

(v) *Whether or not the von Mises dissipation* $\mathcal{D}$ *defines a qausimetric depends on the particular choice of the range of the plastic strain and on the inclusion or exclusion of plastic spin. See Section 8.5 for more details.*

Another question that arises in the context of existence is the uniqueness of the solution. Unfortunately, there are very simple counterexamples indicating that there no uniqueness results.

Example 8.6.3 (Ambiguity of the Solution in two Dimensions)**.** Here is a sketch of a counterexample without any actual numerical treatment involved. It is an observation that probably everyone knows from the real world. Consider a two-dimensional rod $\Omega = [0, 10] \times [0, 1]$. We assume that the material is isotropic, so it has no directional preference. We also assume for simplicity that the dissipation distance satisfies $\mathcal{D}(\mathbf{I}, \mathbf{P}) = \mathcal{D}(\mathbf{I}, -\mathbf{P})$ (this is actually true for the von Mises dissipation and no plastic spin, as seen in Theorem 8.5.1). Then, we apply an external force to both ends of the rod such that we compress the rod. Now there are basically two natural outcomes for the resulting deformed domain $\Omega^{\mathbf{y}}$. Either the rod stays straight and reduces it's length, or the rod bends. The situation is visualized in Figure 8.1. In the case that the rod bends, both the upward and downward bending will result in the same energy, and in the same dissipation distance to the initial configuration. Therefore, the solution cannot be unique.

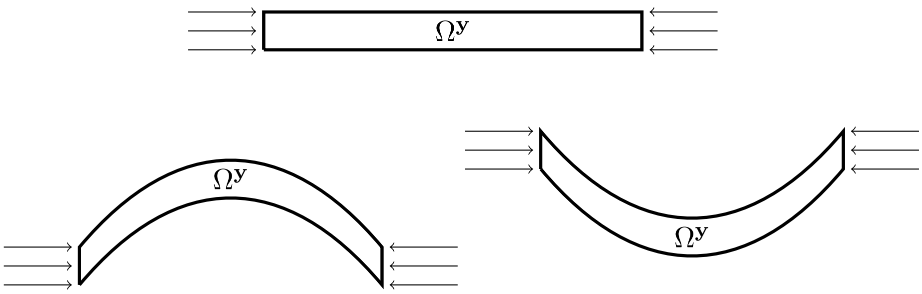


Figure 8.1.: Compression test of a rod with multiple possible natural outcomes.

9. Discrete Problem

9.1. Time Discretization

The definition of the global energy stability (S) leads to a natural scheme for a time discretization. Consider a partition of the time interval $[0,T]$ into the discrete time points $0 = t_0 < t_1 < \ldots < t_N = T$ for $N \in \mathbb{N}$. Of course, we assume that a suitable $(\mathbf{y}^0, \mathbf{P}^0) \in \mathcal{Y} \times \mathcal{P}$ is given at the initial time step t_0. The idea of a natural discretization scheme is that we assume $\mathbf{y}^{k-1}, \mathbf{P}^{k-1}$ to be known for the current time step t_k. Then, regarding the inequality (S), for the current time step t_k we want to minimize the energy $\mathcal{E}$ while having to keep the dissipation distance $\mathcal{D}$ to the previous plastic strain $\mathbf{P}^{k-1}$ small. Therefore, we arrive at the time-discrete minimization problems below.

$$\begin{aligned}&\text{Find } (\mathbf{y}^1, \mathbf{P}^1), \ldots, (\mathbf{y}^N, \mathbf{P}^N) \in \mathcal{Q} \;\text{ such that for } k \in \{1, \ldots, N\} : \\ &(\mathbf{y}^k, \mathbf{P}^k) \text{ minimizes } (\mathbf{y}, \mathbf{P}) \mapsto \mathcal{E}(t_k, \mathbf{y}, \mathbf{P}) + \mathcal{D}(\mathbf{P}^{k-1}, \mathbf{P})\;.\end{aligned} \tag{9.1}$$

The existence of minimizers $(\mathbf{y}^1, \mathbf{P}^1), \ldots, (\mathbf{y}^N, \mathbf{P}^N)$ can be shown by weaker assumptions than the ones we need for Theorem 8.6.1 for the time-continuous case. But, however, the stated assumptions in Section 8.6 imply the necessary conditions for solution of the time-increment problem above. Details on the particular conditions are found in [22, Chapter 2.1.2].

Theorem 9.1.1 ([22, Proposition 2.1.4])**.** *The time-incremental minimization problems from* (9.1) *have a solution for any partition of the time interval* $[0,T]$ *if the assumptions of Theorem 8.6.1 are fulfilled.*

Similar to the time-continuous problem (S&E) we cannot hope for uniqueness of the solution, since we can simply repeat Example 8.6.3 from the preceding section.

9.2. Reformulation of the Minimization Problem

In order to attack the minimization problems (9.1) numerically, we will employ several transformation steps. The goal will be to reach an equivalent formulation without having to deal with the manifold range of the solution space $\mathcal{Q}$, especially the nonlinear plastic strain range $\mathrm{SL}(d)$. Space discretization and finding proper finite element spaces for the solution approximations is analytically and numerically easier if we have to deal with a problem set in a vector space. The canonical vector space corresponding to the manifold $\mathrm{SL}(d)$ is the tangential space at the unit matrix, which is called the Lie algebra and is usually denoted by $\mathrm{sl}(d)$. Therefore, we want to construct a mapping from $\mathrm{sl}(d)$ to $\mathrm{SL}(d)$. This requires slightly different techniques depending on the presence of plastic spin. Thus, we study both cases separately.

9.2.1. Without Plastic Spin

In the case of no plastic spin, or *spinless* plasticity, we restrict the plastic range to the positive definite, symmetric matrices, as already discussed in Section 7.2.2. Recall the spinless plastic range

$$\mathbf{P}(x) \in \mathrm{SL}(d)^+_{\mathrm{sym}} := \Big\{A \in \mathrm{SL}(d) \colon A \text{ is symmetric and positive definite}\Big\}.$$

Inspired by [22, Remark 4.2.9] we introduce the mapping

$$\mathrm{SL}(d)^+_{\mathrm{sym}} \times T\,\mathrm{SL}(d)^+_{\mathrm{sym}} \to \mathrm{sl}(d)^+_{\mathrm{sym}} : \ (P, \xi) \mapsto P^{-\frac{1}{2}} \xi P^{-\frac{1}{2}}$$

into the tangential space at the unit matrix. See Appendix A.3.2 for more details. Note that $\mathrm{sl}(d)^+_{\mathrm{sym}} = \mathbb{S}^d_0$ is the vector space of trace-free symmetric matrices. Using the mapping above for the von Mises dissipation (cf. Section 8.5) we note that we can assume $\sigma_1 = 0$ since $\mathrm{anti}(\xi) = 0$ for all $\xi \in \mathrm{sl}(d)^+_{\mathrm{sym}}$. Hence,

$$\hat{\Delta}(P, \dot{P}) = \Delta(P^{-\frac{1}{2}} \dot{P} P^{-\frac{1}{2}}) = \sigma_0 \big\| \mathrm{sym}(P^{-\frac{1}{2}} \dot{P} P^{-\frac{1}{2}}) \big\|_F = \sigma_0 \big\| P^{-\frac{1}{2}} \dot{P} P^{-\frac{1}{2}} \big\|_F$$

defines a metric, since on $\mathrm{SL}(d)^+_{\mathrm{sym}}$

$$\sigma_0 \big\| P^{-\frac{1}{2}} \dot{P} P^{-\frac{1}{2}} \big\|_F = 0 \quad \Leftrightarrow \quad \dot{P} = 0.$$

Remark 9.2.1. *Although it seems more elegant to use the symmetric mapping into* $\mathrm{sl}(d)^+_{\mathrm{sym}}$ *above, we can still keep using the original dissipation metric* $\|\,\mathrm{sym}\left(\dot{P}(s) P(s)^{-1}\right)\|_F$ *on the tangential bundle* $\mathrm{SL}(d)^+_{\mathrm{sym}} \times T\,\mathrm{SL}(d)^+_{\mathrm{sym}}$*, as introduced in* (8.9)*. This way we avoid interfering with the derivation of the energetic system. From the invariance of the Frobenius norm of the symmetric part under symmetric similarity transformations (Lemma A.4.1) we can conclude that both choices induce the same metric, i.e.,*

$$\begin{aligned} \left\| \mathrm{sym}\left(P(s)^{-\frac{1}{2}} \dot{P}(s) P(s)^{-\frac{1}{2}}\right) \right\|_F &= \left\| \mathrm{sym}\left(P(s)^{\frac{1}{2}} P(s)^{-\frac{1}{2}} \dot{P}(s) P(s)^{-\frac{1}{2}} P(s)^{-\frac{1}{2}}\right) \right\|_F \\ &= \left\| \mathrm{sym}\left(\dot{P}(s) P(s)^{-1}\right) \right\|_F. \end{aligned}$$

In order to use Theorem 8.5.1 note that $P_2 P_1^{-1}$ is in general not in $\mathrm{SL}(d)^+_{\mathrm{sym}}$, even if $P_1, P_2 \in \mathrm{SL}(d)^+_{\mathrm{sym}}$. Therefore, we introduce the *symmetric increment* operator

$$\delta P^+_{\mathrm{sym}} = \delta P^+_{\mathrm{sym}}(P_1, P_2) := P_1^{-\frac{1}{2}} P_2 P_1^{-\frac{1}{2}} \in \mathrm{SL}(d)^+_{\mathrm{sym}} \tag{9.2}$$

for two matrices $P_1, P_2 \in \mathrm{SL}(d)^+_{\mathrm{sym}}$. We will neglect the arguments since they are usually clear from the context.

In order to show that Theorem 8.5.1 still holds in $\mathrm{SL}(d)^+_{\mathrm{sym}}$ we show a small auxiliary result first that corresponds to Lemma 8.5.2.

Lemma 9.2.2. *For* $P_1, P_2 \in \mathrm{SL}(d)^+_{\mathrm{sym}}$ *we have for the*

$$D(P_1, P_2) = D(I, \delta P^+_{\mathrm{sym}}).$$

Proof. Similar to the proof of Lemma 8.5.2, we introduce the sets of connecting paths

$$\begin{aligned}\Pi(P_1, P_2) &:= \Big\{P \in \mathcal{C}^1([0,1], \mathrm{SL}(d))\colon P(0) = P_1, P(1) = P_2\Big\} \qquad \text{and} \\ \delta\Pi(P_1, P_2) &:= \Big\{\delta P \in \mathcal{C}^1([0,1], \mathrm{SL}(d))\colon P(0) = I, P(1) = \delta P^+_{\mathrm{sym}}\Big\}\ .\end{aligned}$$

Note that we intentionally define the paths in $\mathrm{SL}(d)$ since for now we cannot assume that a minimizing path takes values only in $\mathrm{SL}(d)^+_{\mathrm{sym}}$. The corresponding transformation mapping $T\colon \Pi(P_1, P_2) \to \delta\Pi(P_1, P_2)$, $T(P) = P_1^{-\frac{1}{2}} P P_1^{-\frac{1}{2}}$ is again bijective. We observe for $P \in \Pi(P_1, P_2)$, the corresponding $T := T(P) \in \delta\Pi(P_1, P_2)$ and any $s \in [0,1]$ that

$$\dot{T}(s)T(s)^{-1} = P_1^{-\frac{1}{2}}\dot{P}(s)P_1^{-\frac{1}{2}}(P_1^{-\frac{1}{2}}P(s)P_1^{-\frac{1}{2}})^{-1} = P_1^{-\frac{1}{2}}\dot{P}(s)P(s)^{-1}P_1^{\frac{1}{2}}.$$

Hence the norm of the symmetric part fulfills

$$\left\|\mathrm{sym}\left(\dot{T}(s)T(s)^{-1}\right)\right\|_F = \left\|\mathrm{sym}\left(\dot{P}(s)P(s)^{-1}\right)\right\|_F$$

by invariance under symmetric similarity transformations (see Lemma A.4.1). Therefore, the dissipation distance is invariant under the transformation T. ☐

Symmetric Dissipation Distance With the properties above at hand we conclude from Theorem 8.5.1 that

$$D(I, \delta P^+_{\mathrm{sym}}) = \sigma_0 \left\|\log\left(\delta P^+_{\mathrm{sym}}\right)\right\|_F$$

holds for the shortest path from I to $\delta P^+_{\mathrm{sym}}$ defined on the general group $\mathrm{SL}(d)$! We define the dissipation distance D^+_{sym} as in (8.18) but only for differentiable paths in $\mathrm{SL}(d)^+_{\mathrm{sym}}$. Since the shortest path is from a subset of the admissible path from D, we get

$$D^+_{\mathrm{sym}}(I, \delta P^+_{\mathrm{sym}}) \geq \sigma_0 \left\|\log\left(\delta P^+_{\mathrm{sym}}\right)\right\|_F.$$

Nevertheless, we can show the equality.

Lemma 9.2.3. *Let* $\delta P \in \mathrm{SL}(d)^+_{\mathrm{sym}}$. *The von Mises dissipation distance from* I *to* $\delta P^+_{\mathrm{sym}}$ *in* $\mathrm{SL}(d)^+_{\mathrm{sym}}$ *is explicitly given by*

$$D^+_{\mathrm{sym}}(I, \delta P^+_{\mathrm{sym}}) = \sigma_0 \left\|\log\left(\delta P^+_{\mathrm{sym}}\right)\right\|_F. \tag{9.3}$$

Proof. Consider the path

$$P : [0,1] \to \mathrm{SL}(d)^+_{\mathrm{sym}}, \quad P(t) := \exp(t\ \log(\delta P^+_{\mathrm{sym}})),$$

with connects I and $\delta P^+_{\mathrm{sym}}$ within $\mathrm{SL}(d)^+_{\mathrm{sym}}$. From

$$\dot{P}(t) = \log(\delta P^+_{\mathrm{sym}})P(t),$$

the dissipation distance of this path is given by

$$\int_0^1 \sigma_0 \left\|\dot{P}(t)P(t)^{-1}\right\|_F dt = \int_0^1 \sigma_0 \left\|\log\left(\delta P^+_{\mathrm{sym}}\right)\right\|_F dt = \sigma_0 \left\|\log\left(\delta P^+_{\mathrm{sym}}\right)\right\|_F.$$

Hence, we also have $D^+_{\mathrm{sym}}(\delta P) \leq \sigma_0 \left\|\log\left(\delta P^+_{\mathrm{sym}}\right)\right\|_F$ and thus the desired equality. □

Global Reformulation With the above reformulations in mind, we can write the global time-incremental minimization problems (9.1) in the displacement $\mathbf{y}$ and in the plastic increment $\delta\mathbf{P}^+_{\mathrm{sym}}$. To express the next unknown state $\mathbf{P}^k$, we define the inverse increment operator

$$I^+_{\mathrm{sym}}(P, \delta P^+_{\mathrm{sym}}) := P^{\frac{1}{2}}\, \delta P^+_{\mathrm{sym}}\, P^{\frac{1}{2}}, \tag{9.4}$$

such that

$$\delta P^+_{\mathrm{sym}} = P_1^{-\frac{1}{2}} P_2 P_1^{-\frac{1}{2}} \quad \Leftrightarrow \quad P_2 = I^+_{\mathrm{sym}}(P_1, \delta P^+_{\mathrm{sym}}).$$

As a consequence, each of the minimization problems from (9.1) can be considered in the equivalent form

$$\begin{aligned}&\text{Find } (\mathbf{y}^k, \delta\mathbf{P}^k) \text{ that minimizes } \mathcal{E}\left(\mathbf{y}^k, I^+_{\mathrm{sym}}\left(\mathbf{P}^{k-1}, \delta\mathbf{P}^k\right)\right) + \mathcal{D}^+_{\mathrm{sym}}(\delta\mathbf{P}^k)\\ &\text{and set } \mathbf{P}^k := I^+_{\mathrm{sym}}\left(\mathbf{P}^{k-1}, \delta\mathbf{P^k}\right),\end{aligned} \tag{9.5}$$

where we abbreviate $\delta\mathbf{P} = \delta\mathbf{P}^+_{\mathrm{sym}}$ for better readability. Note that from Lemma 9.2.3 the dissipation distance depends only on $\delta\mathbf{P}^+_{\mathrm{sym}}$, and hence the unit matrix as an argument is dropped in the formulation above.

Tangential Space In the spinless case of plasticity, the lemma below is essential.

Lemma 9.2.4. *Let* $\mathbb{S}^d_0 := \left\{A \in \mathbb{R}^{d\times d} :\ A^T = A,\ \mathrm{trace}(A) = 0\right\}$ *denote the trace-free symmetric matrices. Then, the canonical matrix exponential*

$$\exp : \mathbb{S}^d_0 \to \mathrm{SL}(d)^+_{\mathrm{sym}},\ \exp(A) := \sum_{i=0}^{\infty} \frac{A^i}{i!}$$

is bijective.

Proof. First of all, the eigenvalue decomposition of symmetric matrices $A = VJV^T$ into an orthogonal V (containing the eigenvectors) and a diagonal J (containing the eigenvalues) has the property

$$\exp(A) = V \exp(J) V^T.$$

An immediate consequence is that the determinant of the range of the matrix exponential is always one if the domain is trace-free. Moreover the scalar exponential function is a bijective mapping from $\mathbb{R} \to (0, \infty)$. Hence, we can prescribe both eigenvalues and eigenvectors in $\mathbb{S}^d_0$ and $\mathrm{SL}(d)^+_{\mathrm{sym}}$ at the same time which determines matrices uniquely. □

This allows to rewrite the global problem (9.5) again in terms of $\delta\mathbf{P}^+_{\mathrm{sym}} = \exp(\delta\mathbf{B})$ for a uniquely determined $\delta\mathbf{B}$ mapping $x \in \Omega$ to matrices from $\mathbb{S}^3_0$:

$$\begin{aligned}&\text{Find } (\mathbf{y}^k, \delta\mathbf{B}^k) \text{ that minimizes } \mathcal{E}\Big(\mathbf{y}^k, I^+_{\mathrm{sym}}(\mathbf{P}^{k-1}, \exp(\delta\mathbf{B}^k))\Big) + \mathcal{D}^+_{\mathrm{sym}}(\exp(\delta\mathbf{B}^k))\\ &\text{and set } \mathbf{P}^k := I^+_{\mathrm{sym}}\left(\mathbf{P}^{k-1}, \exp(\delta\mathbf{B}^k)\right),\end{aligned} \tag{9.6}$$

where the term $\exp(\delta\mathbf{B})$ is understood pointwise. Now the log term in the dissipation functional from Lemma 9.2.3 cancels out. In summary, the objective functional of the time-discrete increment problem at time step t_k is given by

$$\boxed{\mathcal{L}^+_{\mathrm{sym}}(\mathbf{y}^k, \delta\mathbf{B}^k) := \mathcal{E}\Big(\mathbf{y}, I^+_{\mathrm{sym}}(\mathbf{P}^{k-1}, \exp(\delta\mathbf{B}^k))\Big) + \sigma_0 \int_\Omega \left\|\delta\mathbf{B}^k(x)\right\|_F dx\,.} \tag{9.7}$$

Remarks 9.2.5.

(i) *The objective functional $\mathcal{L}^+_{\mathrm{sym}}$ is defined over the vector space $\mathcal{Y} \times \mathcal{B}$ where the deformation function space $\mathcal{Y}$ is already given in* (8.22)*, and $\mathcal{B}$ is a suitable Sobolev space over Ω with values in $\mathbb{S}^d_0$. Hence, no manifold-valued spaces appear in this equivalent formulation.*

(ii) *The reformulated dissipation term of $\mathcal{L}^+_{\mathrm{sym}}$ corresponds to the von Mises dissipation in the small-strain case in the first part of this thesis, cf.* (3.5).

9.2.2. With Plastic Spin

The case of included plastic spin needs a careful treatment in order to arrive at similar minimization functionals as seen above in the spinless case. Now we consider the general matrix group SL(d) for the plastic range, but we stick to the case of von Mises dissipation (cf. Section 8.5). Therefore, we consider the dissipation function Δ from (8.20), given by

$$\Delta(\xi) := \left(\sigma_0^2 \left\|\mathrm{sym}(\xi)\right\|_F^2 + \sigma_1^2 \left\|\mathrm{anti}(\xi)\right\|_F^2\right)^{\frac{1}{2}},$$

and the resulting dissipation distance $\mathcal{D}$ (8.18). As already discussed before, D is only a quasi-metric on SL(d) (and therefore sufficient for the existence theory) if and only if $\sigma_0, \sigma_1 > 0$. On the other hand, unfortunately, we have only an explicit form of the dissipation distance D in the case of $\sigma_1 = 0$. This explicit form is given in Theorem 8.5.1 by

$$D(\delta P) = D(I, \delta P) = \sigma_0 \left\|\log\left(\delta P\, \delta P^T\right)^{\frac{1}{2}}\right\|_F$$

for the increment $\delta P = P_2 P_1^{-1}$ between the plastic states P_1 and P_2 in SL(d). Recall from Lemma 8.5.2 that we can neglect the dependency on the unit matrix and hence express D only in the increment δP. For the general case σ_1 no explicit expression is known to the author. However, an implicit form of the dissipation function is given in in [21, page 86]. If we take the full Frobenius norm ($\sigma_0 = \sigma_1$) in (8.20), then we have

$$\begin{aligned} &D_{\sigma_1=\sigma_0}(\delta P) \\ &= \min\left\{\sigma_0\,(\|S\|_F^2 + \|A\|_F^2)^{\frac{1}{2}} \ :\ S^T = S,\, A^T = -A,\, \delta P = \exp(S-A)\exp(2A)\right\}. \end{aligned} \tag{9.8}$$

Lemma 9.2.6. *If we assume that the plastic increment δP is symmetric, then we have*

$$D(\delta P) = \sigma_0 \left\|\log\left(\delta P\right)\right\|_F$$

independent of $\sigma_1 > 0$.

Proof. From the implicit formula (9.8), we see that $S = \log(\delta P)$ and $A = 0$ is an admissible choice and yields

$$D_{\sigma_1=\sigma_0}(\delta P) \leq \sigma_0 \left\|\log\left(\delta P\right)\right\|_F.$$

The equality is established by considering the case $\sigma_1 = 0$. In this case the equality follows from Theorem 8.5.1. However, for a fixed σ_0 and an arbitrary $\sigma_1 > 0$ the dissipation function fulfills

$$\Delta_{\sigma_1>0}(\xi) \geq \Delta_{\sigma_1=0}(\xi),$$

and therefore,

$$D_{\sigma_1>0}(\delta P) \geq D_{\sigma_1=0}(\delta P).$$

This shows that $S = \log(\delta P)$ and $A = 0$ is indeed the optimal choice. ☐

Symmetric Increments Lemma 9.2.6 states that symmetric plastic increments are favorable, since we have an explicit expression of the dissipation distance in this case. In the following we want to motivate why *assuming* that δP is symmetric can be justified in practice.

Let the two plastic strain states P_1 and P_2 in $\mathrm{SL}(d)$ be close to each other. Then, the increment $\delta P = P_2 P_1^{-1}$ is close the unit matrix. In other words, we can assume that

$$\|\delta P - I\|_F = \varepsilon$$

for a small $\varepsilon > 0$. From orthogonality of the symmetric part $\mathrm{sym}(\delta P)$ and the antisymmetric part $\mathrm{anti}(\delta P)$ of the plastic increment we have

$$\|\delta P\|_F^2 = \|\mathrm{sym}(\delta P)\|_F^2 + \|\mathrm{anti}(\delta P)\|_F^2.$$

The norm is the antisymmetric part of δP is small since

$$\begin{aligned} \|\mathrm{anti}(\delta P)\|_F &= \frac{1}{2}\left\|\delta P - \delta P^T\right\|_F = \frac{1}{2}\left\|(\delta P - I) - (\delta P - I)^T\right\|_F \\ &\leq \frac{1}{2}\left(\|\delta P - I\|_F + \left\|(\delta P - I)^T\right\|_F\right) = \|\delta P - I\|_F = \varepsilon. \end{aligned}$$

On the other hand, the norm of the symmetric part fulfills

$$\|\mathrm{sym}(\delta P)\|_F^2 = \|\delta P\|_F^2 - \|\mathrm{anti}(\delta P)\|_F^2 \geq 1 - \varepsilon^2,$$

since $\delta P \in \mathrm{SL}(d)$ has at least one singular value greater or equal to one. Therefore, we can assume that δP is symmetric if we consider small increment steps in the time-discrete problem. We formulate the main assumption of the case of included plastic spin:

$$\boxed{\delta P \text{ is symmetric in the time-discrete problem formulation.}} \tag{9.9}$$

Since we assume that $\delta P \approx I$, we can also conclude that we have positive definiteness, i.e.,

$$\delta P \in \mathrm{SL}(d)^+_{\mathrm{sym}}.$$

In order to investigate the consequences of this assumption, we introduce the inverse increment operator

$$I(P, \delta P) := \delta P\, P, \tag{9.10}$$

which is used to recover the current plastic strain from the increment, i.e,

$$P_2 = I(P_1, \delta P) \quad \Leftrightarrow \quad \delta P = P_2 P_1^{-1}.$$

Note that symmetry of the plastic increment δP does not in imply symmetry of the plastic strain itself since products of symmetric matrices are generally not symmetric. Therefore, we are *not* in the case of spinless plasticity.

With the above stated assumptions at hand, we can reformulate the time-incremental minimization problem in terms of the plastic increments.

$$\begin{aligned} &\text{Find } (\mathbf{y}^k, \delta\mathbf{P}^k) \text{ that minimizes } \mathcal{E}\left(\mathbf{y}^k, \delta\mathbf{P}^{\mathbf{k}}\,\mathbf{P}^{k-1}\right) + \mathcal{D}(\delta\mathbf{P}^{\mathbf{k}}) \\ &\text{and set } \mathbf{P}^k := \delta\mathbf{P}^{\mathbf{k}}\,\mathbf{P}^{k-1}. \end{aligned} \tag{9.11}$$

Since the space of unknowns of the minimization problems is now the same as in the spinless case, we can utilize Lemma 9.2.4 to express δP bijectively by a field of symmetric, trace-free matrices δB. Then, the minimization takes place again in a vector space setting, and we arrive at the global problem similar to (9.6)

$$\begin{aligned} &\text{Find } (\mathbf{y}^k, \delta\mathbf{B}^{\mathbf{k}}) \text{ that minimizes } \mathcal{E}\Big(\mathbf{y}, \exp(\delta\mathbf{B}^k)\,\mathbf{P}^{k-1}\Big) + \mathcal{D}(\exp(\delta\mathbf{B}^k)) \\ &\text{and set } \mathbf{P}^k := \exp(\delta\mathbf{B}^k)\mathbf{P}^{k-1} \end{aligned} \tag{9.12}$$

By Lemma 9.2.6 the exp and the log terms cancel out in the explicit dissipation form. This leads to the objective functional at time step t_k

$$\boxed{\mathcal{L}(\mathbf{y}^k, \delta\mathbf{B}^k) := \mathcal{E}\Big(\mathbf{y}^k, \exp(\delta\mathbf{B}^k)\mathbf{P}^{k-1}\Big) + \sigma_0 \int_\Omega \left\|\delta\mathbf{B}^k(x)\right\|_F dx\,.} \tag{9.13}$$

Note that the Remarks 9.2.5 also apply for $\mathcal{L}$.

Nonsymmetric Plastic Increments A nonsymmetric plastic increment δP does not allow the reformulation in the tangential space of symmetric, trace-free matrices as described above. Nevertheless, we can still take advantage of the cancellation of exp and log in the dissipation distance by assuming that $\sigma_1 = 0$ in the von Mises dissipation (8.20). By doing so we lose the quasidistance properties of the dissipation distance D, and thus sufficient existence conditions for the energetic solution. In certain applications, however, this may not have a negative effect. Instead of considering the tuple $(\mathbf{y}, \delta\mathbf{P})$ in the increment problems (9.11), we consider the triple $(\mathbf{y}, \delta\mathbf{P}, \delta\mathbf{L})$, where δP is a matrix field of symmetric, positive definite matrices, and $\delta\mathbf{L}$ is a matrix field of rotations in $\mathrm{SO}(d)$, keeping track of the orthogonal parts of the plastic increment. We use that polar decomposition

$$\delta\mathbf{L}\,\delta\mathbf{P}$$

is always well-defined and we can represent all matrix fields in $\mathrm{SL}(d)$ by this split.

We can again go back to the space of trace-free, symmetric matrices $\mathbb{S}^d_0$ and arrive at the objective functionals

$$\tilde{\mathcal{L}}(\mathbf{y}^k, \delta\mathbf{B}^k, \delta\mathbf{L}^k) \coloneqq \mathcal{E}\Big(\mathbf{y}^k, \delta\mathbf{L}^k \exp(\delta\mathbf{B}^k)\mathbf{P}^{k-1}\Big) + \sigma_0 \int_\Omega \Big\|\delta\mathbf{B}^k(x)\Big\|_F \,dx\,.$$

The main difficulty of this form is an efficient space discretization of the manifold-valued $\delta\mathbf{L}^k$ function. In this thesis we concentrate in the following on the first two presented objective functionals $\mathcal{L}^+_{\mathrm{sym}}$ (9.7) and $\mathcal{L}$ (9.13) only.

9.3. Space Discretization

The global increment problems (9.6) and (9.12) will be discretized in space by finite element methods. The resulting objective functionals (9.7) and (9.13) depend on three functions over Ω that need to be discretized. There is the deformation field $\mathbf{y}$, the tangential plastic increment $\delta\mathbf{B}$ with values in $\mathbb{S}^d_0$, and the previous plastic strain $\mathbf{P}^{k-1}$ with values in $\mathrm{SL}(d)$. We consider a conforming triangulation $\mathcal{T}$ of the domain Ω by a set of m geometric elements $T_1, \dots, T_m \in \mathcal{T}$. We denote by n the number of vertices $v_1, \dots, v_n$ of the resulting grid.

9.3.1. Discrete Function Spaces

Deformation Field The first function to be discretized is the displacement field $\mathbf{y} : \Omega \to \mathbb{R}^d$ that is given as a first-order Sobolev function from the function space $\mathcal{Y}$ (8.22). Hence, it is canonical to use first-order Lagrange finite elements for an approximation $\mathbf{y}_h$ of $\mathbf{y}$. To this end, consider the nodal basis $\{\boldsymbol{\phi}_i\}_{i=1,\dots,n}$ in d dimensions with the Lagrange property $\boldsymbol{\phi}_i(v_j) = \delta_{ij}$ for the vertices v_j of the grid. Then the discrete approximation of the deformation field is given by

$$\mathbf{y}_h(x) \coloneqq \sum_{i=1}^{n} \sum_{j=1}^{d} \bar{y}_{ij}\, \boldsymbol{\phi}_i(x)\, e_j \tag{9.14}$$

with canonical unit vectors $e_j \in \mathbb{R}^d$ and real-valued coefficients $\bar{y}_{ij}$.

Tangential Plastic Increment The tangential plastic increment $\delta\mathbf{B} : \Omega \to \mathbb{S}_0^d$ is also a first-order Sobolev function since we need derivative information for the chain rule in the gradient term in the hardening functional (cf. Section 8.2). Therefore, first-order Lagrange finite elements are again suitable. A basis of $\mathbb{S}_0^d$ is given in [30, Section 3.2] by

$$B_1 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad B_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

of the two dimensional space S_0^2 if $d = 2$. For the case $d = 3$, the basis matrices for the five-dimensional space $\mathbb{S}_0^3$ are

$$B_1 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B_2 := \frac{1}{\sqrt{6}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$

$$B_3 := \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B_4 := \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad B_5 := \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

In general, we have the dimension

$$d_p := \frac{d^2 + d}{2} - 1$$

for the space S_0^d, cf. Remark 4.3.1 in the first part of this thesis. The choice of the bases above forms an orthonormal bases of $\mathbb{S}_0^d$ corresponding to the Frobenius inner product. Therefore, it defines an isometry between $\mathbb{R}^{d_p}$ equipped with the Euclidean norm $\|\cdot\| 2$ and $\mathbb{S}_0^d$ with the Frobenius norm, as

$$\Big\| \sum_{j=1}^{d_p} a_j B_j \Big\|_F = \|a\|_2 \tag{9.15}$$

holds for all $a \in \mathbb{R}^d$. In summary, the discrete plastic strain field is represented by

$$\delta\mathbf{B}_h(x) := \sum_{i=1}^{n} \sum_{j=1}^{d_p} \bar{b}_{ij}\, \boldsymbol{\phi}_i(x)\, B_j$$

with scalar coefficients $\bar{b}_{ij}$.

Some numerical tricks can be applied to evaluate the objective functionals $\mathcal{L}$ (9.13) and $\mathcal{L}_{\text{sym}}^+$ (9.7). Since first-order Lagrangian shape functions are non-negative, the dissipation term can be approximated by a lumped sum

$$\begin{aligned} \int_\Omega \|\delta\mathbf{B}_h(x)\|_F \, dx &\approx \sum_{i=1}^{n} \int_\Omega \Big\| \sum_{j=1}^{d_p} \bar{b}_{ij}\, \boldsymbol{\phi}_i(x)\, B_j \Big\|_F \, dx \\ &= \sum_{i=1}^{n} \int_\Omega \boldsymbol{\phi}_i(x)\, dx \, \|\bar{b}_i\|_2 =: \sum_{i=1}^{n} \gamma_i \, \|\bar{b}_i\|_2. \end{aligned} \tag{9.16}$$

In this form we separate the block components of the plastic strain. This is a structural necessity for the TNNMG solver from Section 5.2.

We will also consider the case $k_2 = 0$ in the hardening functional (8.25). In this case no gradient information is needed. Thus, we can choose zero-order (piecewise constant) Lagrange finite elements to decrease the number of degrees of freedom of the fully discrete problem. Hence, consider a zero order basis $\{\boldsymbol{\theta}_i\}_{i=1,\dots,m}$ with the Lagrange property $\boldsymbol{\theta}_i|_{T_j} \equiv \delta_{ij}$ for the grid elements $T_j \in \mathcal{T}$. By reusing the bases of $\mathbb{S}_0^d$ from above, the zero-order discrete approximation of $\delta\mathbf{B}$ is given by

$$\delta\mathbf{B}_h(x) \coloneqq \sum_{i=1}^{m}\sum_{j=1}^{d_p} \bar{b}_{ij}\,\boldsymbol{\theta}_i(x)\,B_j.$$

Since zero-order approximations are piecewise constant, and each base function $\boldsymbol{\theta}_i$ has support only on one grid element, so no lumping of the dissipation as in (9.16) is necessary. We have the exact sum

$$\begin{aligned}\int_\Omega \|\delta\mathbf{B}_h(x)\|_F \; dx &= \sum_{i=1}^{m}\int_\Omega \left\| \sum_{j=1}^{d_p} \bar{b}_{ij}\,\boldsymbol{\theta}_i(x)\,B_j \right\|_F \; dx \\ &= \sum_{i=1}^{m}\int_\Omega \boldsymbol{\phi}_i(x)\,dx\;\left\|\bar{b}_i\right\|_2 \eqqcolon \sum_{i=1}^{m}\gamma_i\;\left\|\bar{b}_i\right\|_2,\end{aligned}$$

where $\gamma_i = |T_i|$ is the volume of the grid elements in this case.

Previous Plastic Strain The discretization of the previous plastic strain $\mathbf{P}^{k-1}$ may need more advanced methods since it takes values in the nonlinear matrix manifold $\mathrm{SL}(d)$. In the case of zero-order finite elements for $\delta\mathbf{B}_h$, we can simply choose the same zero-order Lagrange basis $\{\boldsymbol{\theta}_i\}$ and approximate $\mathbf{P}^{k-1}$ by

$$\mathbf{P}_h^{k-1}(x) \coloneqq \sum_{i=1}^{m} \bar{P}_i\,\boldsymbol{\theta}_i(x)$$

for $\mathrm{SL}(d)$-valued coefficients $\bar{P}_i$.

If gradients of the plastic strain are present in the hardening model ($k_2 > 0$), then a different approach is required. We need to construct a finite element space of weakly differentiable functions with values in $\mathrm{SL}(d)$. Various such spaces have been constructed in [17] under the name of Geometric Finite Elements. However, the construction and evaluation of these special finite elements is expensive and complicated in implementation. Since this is beyond the scope of this thesis, a simpler approach is used.

We use the extrinsic vector space $\mathbb{R}^{d\times d}$ of $\mathrm{SL}(d)$ for the range of the previous plastic strain and discretize this vector space by first-order Lagrange finite elements. Hence, the coefficients correspond to the vertices v_j of the grid. On each grid vertex v_j we demand $\mathbf{P}_h^{k-1}(v_j) \in \mathrm{SL}(d)$. From the nonlinear nature of $\mathrm{SL}(d)$ an evaluation of $\mathbf{P}_h^{k-1}$ at an arbitrary point $x \in \Omega$ returns most likely a value not in $\mathrm{SL}(d)$. Hence, we compute the orthogonal projection of $\mathbf{P}_h^{k-1}(x)$ onto $\mathrm{SL}(d)$ for each evaluation. Therefore, this method is called projection-based finite element method. Details about how to compute those projections are found in the Appendix A.3.

A missing numerical details is that the first-order finite element spaces of $\delta\mathbf{B}_h$ and $\mathbf{P}_h^{k-1}$ are generally not compatible. We combine explicit computations of values in $\mathrm{SL}(d)$ (by the matrix exponential) with a projection-based method. In order to avoid complicated analytical obstacles and implementation issues, we neglect this problem in the hope that the overall discretization error dominates error quantities arising form this incompatibility. The update product (9.6) for the next approximated plastic strain $\mathbf{P}_h^k$ is then simply computed at the grid vertices to determine new coefficients.

Gradients of $\mathbf{P}_h^{k-1}$ are not projected for reasons of simplicity. Since these gradients are only used to regularize the model, we assume that a precise evaluation of the gradients is not necessary and decide to accept the additional discretization error.

9.3.2. Discrete Plastic Strain Gradients

We still need a proper implementation of the plastic strain gradient for the hardening energy (8.25). Depending on the underlying plastic strain range this procedure differs and is formulated for both $\mathrm{SL}(d)^+_{\mathrm{sym}}$ and $\mathrm{SL}(d)$ separately.

Symmetric range $\mathrm{SL}(d)^+_{\mathrm{sym}}$ The plastic increment between two plastic states P_1 and P_2 is given in the form

$$\delta P^+_{\mathrm{sym}} = \delta P^+_{\mathrm{sym}}(P_1, P_2) := P_1^{-\frac{1}{2}} P_2 P_1^{-\frac{1}{2}} \in \mathrm{SL}(d)^+_{\mathrm{sym}}$$

in (9.2). In order to invert the increment and compute the current plastic strain we need the operator

$$I^+_{\mathrm{sym}}(P, \delta P^+_{\mathrm{sym}}) := P^{\frac{1}{2}}\, \delta P^+_{\mathrm{sym}}\, P^{\frac{1}{2}}$$

from (9.4). We lift this to the discrete and global level, and replace the plastic increment by the exponential mapping from the tangential space. The resulting current plastic strain at time step t_k is therefore given by

$$\mathbf{P}_h(x) = \left(\mathbf{P}_h^{k-1}(x)\right)^{\frac{1}{2}} \exp\left(\delta\mathbf{B}_h^k(x)\right)\left(\mathbf{P}_h^{k-1}(x)\right)^{\frac{1}{2}}$$

with the discretized global functions described in the previous section.

The gradient $\nabla\mathbf{P}_h(x)$ can be expressed by combining the product rule and the chain rule. It takes the form

$$\begin{aligned}\nabla\mathbf{P}_h(x) =& \nabla\left(\left(\mathbf{P}_h^{k-1}(x)\right)^{\frac{1}{2}}\right) \exp\left(\delta\mathbf{B}_h^k(x)\right)\left(\mathbf{P}_h^{k-1}(x)\right)^{\frac{1}{2}} \\ &+ \left(\mathbf{P}_h^{k-1}(x)\right)^{\frac{1}{2}} \nabla\exp\left(\delta\mathbf{B}_h^k(x)\right)\left[\nabla\delta\mathbf{B}_h^k(x)\right]\left(\mathbf{P}_h^{k-1}(x)\right)^{\frac{1}{2}} \\ &+ \left(\mathbf{P}_h^{k-1}(x)\right)^{\frac{1}{2}} \exp\left(\delta\mathbf{B}_h^k(x)\right)\nabla\left(\left(\mathbf{P}_h^{k-1}(x)\right)^{\frac{1}{2}}\right).\end{aligned}$$

To compute this numerically, we need the Fréchet-derivative of the matrix square root and of the matrix exponential.

The Fréchet-derivative of the matrix square root of a matrix A in direction E can be obtained by setting $s(A) := A^{\frac{1}{2}}$ and deriving $s(A)^2 - A = 0$ in direction E. This leads to the Sylvester-like equation

$$s(A)X + Xs(A) = E, \tag{9.17}$$

which has the unique solution $X = \nabla s(A)[E] \in \mathbb{R}^{d\times d}$ if all eigenvalues of $s(A)$ are positive [18, Sec. B.14.]. This is the case for $A \in \mathrm{SL}(d)^+_{\mathrm{sym}}$. The square root $A^{\frac{1}{2}}$ is also well-defined. For symmetric A and small $d \in \{2, 3\}$ the problem (9.17) is a linear equation with only few unknowns, and it can be solved directly. Furthermore, $A^{\frac{1}{2}}$ can be computed by the eigenvalue decomposition of A, which can be computed directly for $d \in \{2, 3\}$.

The Fréchet-derivative $\nabla \exp(A)[E]$ of the matrix exponential of A in direction E can be computed by evaluating

$$\exp \begin{pmatrix} A & E \\ 0 & A \end{pmatrix} = \begin{pmatrix} \exp(A) & \nabla \exp(A)[E] \\ 0 & \exp(A) \end{pmatrix}$$

and extracting the upper right block [18, Sec. 10.6]. The matrix exponential itself is computed by the *Scaling and Squaring Method* [18, Sec. 10.3.], which is presented in Appendix A.5. With these tools at hand, we can numerically compute the gradient $\nabla \mathbf{P}_h(x)$ at any point $x \in \Omega$.

General range $\mathrm{SL}(d)$ In the general case, the plastic increment between two plastic states P_1 and P_2 is given by

$$\delta P = \delta P(P_1, P_2) := P_2 P_1^{-1} \in \mathrm{SL}(d)$$

in (9.10). The resulting current plastic strain at time step t_k is here given by

$$\mathbf{P}_h(x) = \exp\Big(\delta \mathbf{B}_h^k(x)\Big) \mathbf{P}_h^{k-1}(x).$$

The gradient $\nabla \mathbf{P}_h(x)$ is therefore again given by combining the product rule and the chain rule. The result is simpler than in the other case above. We get

$$\begin{aligned} \nabla \mathbf{P}_h(x) = & \nabla \exp\Big(\delta \mathbf{B}_h^k(x)\Big) \Big[\nabla \delta \mathbf{B}_h^k(x)\Big] \mathbf{P}_h^{k-1}(x) \\ & + \exp\Big(\delta \mathbf{B}_h^k(x)\Big) \nabla \Big(\mathbf{P}_h^{k-1}(x)\Big). \end{aligned}$$

How we construct the gradient of the exponential function is already given in the other case above.

9.3.3. Discrete Minimization Functional

The space discretization of the deformation field $\mathbf{y}^k$ and the tangential plastic increment $\delta\mathbf{B}^k$ is now complete with the finite element spaces presented above. We arrive at an algebraic minimization functional

$$\boxed{L(\bar{w}) = L(\bar{y}, \bar{b}) \coloneqq \mathcal{E}\Big(\mathbf{y}_h^k, I^*(\mathbf{P}_h^{k-1}, \exp(\delta\mathbf{B}_h^k))\Big) + \sigma_0 \sum_{i=1}^{r} \gamma_i \left\|\bar{b}_i\right\|_2} \tag{9.18}$$

in the coefficients $\bar{w} = (\bar{y}, \bar{p})$, and the inverse increment operator $I^* = I$ (9.10) for the case of plastic spin, and $I^* = I^+_{\text{sym}}$ (9.4) without plastic spin. The sum term is defined with $r \in \{n, m\}$, i.e., either the number of grid vertices n if we approximate the plastic increment by first-order elements, or the number of grid elements m in the zero-order case, respectively. The space of the coefficients is simply given by

$$\bar{w} = (\bar{y}, \bar{p}) \in \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} = \mathbb{R}^N$$

for the total number of deformation coefficients $N_1 = dn$, and the number of plastic increment coefficients $N_2 = rd_p$. Since the finite element spaces of the appearing unknowns are linear, the coefficient space is too.

Remarks 9.3.1.

(i) *The energy term $\mathcal{E}$ is smooth since all internal components are differentiable in the coefficients.*

(ii) *Nevertheless, we cannot expect convexity of $\mathcal{E}$ in the coefficients $\bar{w}$.*

(iii) *The dissipation term is precisely the same as in the small-strain fully discrete case* (4.17) *without isotropic hardening.*

10. The Inexact Proximal Newton Method

The discrete objective functional L (9.18) is defined in the space of the coefficients $\bar{w} \in \mathbb{R}^N$, where N is the total number of unknowns. By the reformulations performed in Section 9.2, we are able to remove all nonlinearities from the domain space. As a consequence, we receive a rather complicated structure within L itself. We notice that L has two properties, both of which complicate the numerical treatment for the minimization. First, L is nonconvex, which prevents us from using standard Newton-like minimization techniques. And second, there are nonsmooth components in the dissipation term, which even blocks the application of first-order gradient methods. To the author's knowledge, there is no stable and efficient algorithm for minimizing such functionals with a convergence guarantee to local minimizers for an arbitrary initial value.

To fill this gap, a Proximal Newton method was developed in collaboration with Anton Schiela and Bastian Pötzl. In the following, the solver, its main properties and implementation aspects are presented. Furthermore, the application to the discrete minimization problems of finite-strain plasticity is given. A great overview of the development of Proximal Newton methods in recent years, combined with a deep analysis of the properties given in this thesis, can be found in the article [26]. In a follow-up article [25] we improve the Proximal Newton method by introducing inexactness criteria, which are briefly given in Section 10.4. The main statements of this chapter, as well as a majority of the notation, are taken from these two articles.

10.1. General Setting

The Proximal Newton method is designed to solve composite minimization problems of the form

$$\min_{x \in X} F(x) := f(x) + g(x). \tag{10.1}$$

The minimization problem is defined on a real Hilbert space $(X, \langle \cdot, \cdot \rangle_X)$ with the corresponding scalar product. We have the induced norm $\|v\|_X = \sqrt{\langle v, v \rangle_X}$ and the dual space X^*. In the minimization problem above the functions $f, g : X \to \mathbb{R} \cup \{\infty\}$ take different roles.

The first component f is called the smooth part of the objective functional F. To ensure that the steps in the following sections are well-defined we demand at least that $f \in \mathcal{C}^1(X)$, and that the first derivative is Lipschitz-continuous. Moreover, we assume that for each $x \in X$ there is a linear mapping $H_x : X \to X^*$ with an upper bound

$$\|H_x\|_{X \to X^*} \leq M$$

of the operator norm for some $M > 0$. Furthermore, we demand an ellipticity constraint

$$(H_x v)(v) \geq \kappa_1(x) \|v\|_X^2 \quad \forall v \in X$$

with $\kappa_1(x) \in \mathbb{R}$ depending on $x \in X$. We want to emphasize that $\kappa_1(x)$ can take negative values. Additionally, we demand that the mapping H_x is connected to f by the *semi-smoothness* condition

$$\|f'(y) - f'(x) - H_x(y-x)\|_{X^*} = o(\|y-x\|_X) \quad \forall x, y \in X. \tag{10.2}$$

Remark 10.1.1. *Consider a twice continuously differentiable $f \in \mathcal{C}^2(x)$ and choose $H_x := f''(x)$. If there is a lower bound $m \in \mathbb{R}$ independent of $x \in \mathbb{R}$, s.t.,*

$$(H_x v)(v) \geq m$$

for every $x, v, \in X$, then f satisfies the properties stated above.

In contrast to f the second component g is not assumed to be differentiable. Therefore, we call g the nonsmooth part of F. We only demand that g is lower semi-continuous and satisfies a convexity bound of the form

$$g(sx + (1-s)y) \leq sg(x) + (1-s)g(y) - \frac{\kappa_2}{2} s(1-s) \|x-y\|_X^2$$

with a $\kappa_2 \in \mathbb{R}$. Note that negative values of κ_2 extent the classical notion of convexity, whereas g is called strong convex for a positive κ_2.

However, for a fast local convergence result, we assume that $\kappa_1(x) + \kappa_2 > 0$ near optimal solutions of (10.1), cf. [26, Theorem 1].

10.2. Proximal Newton Steps

The Proximal Newton method is an iterative method for solving the minimization problem (10.1). Consider an arbitrary initial iterate $x^0 \in X$. Then one Proximal Newton step consists of two main parts, as described below.

10.2.1. Second-Order Model Problems

In order to find the next iterate, we reduce the complexity of the original minimization problem. Therefore we replace the smooth part f around x in (10.1) by a second-order model. Since we have no convexity properties of f at hand, we regularize the second-order model by an additional quadratic regularization term $\frac{\omega}{2} \|\delta x\|_X^2$ with a given weight $\omega \geq 0$. A similar replacement of g is not possible due to the nondifferentiability. Hence, the descent of the value in g is given directly. The resulting second-order minimization functional $\lambda_{x,\omega} : X \to \mathbb{R}$ is given by

$$\lambda_{x,\omega}(\delta x) := f'(x)\delta x + \frac{1}{2}(H_x \delta x)(\delta x) + \frac{\omega}{2} \|\delta x\|_X^2 + g(x + \delta x) - g(x). \tag{10.3}$$

We denote the minimizer, providing it exists, by

$$\Delta x(\omega) := \arg\min_{\delta x \in X} \lambda_{x,\omega}(\delta x)\,. \tag{10.4}$$

From the convexity and ellipticity bounds on f and g in Section 10.1 we get immediately the following result:

Lemma 10.2.1 ([26, Proposition 1])**.** *If* $\omega > -(\kappa_1(x) + \kappa_2)$ *a unique solution* $\Delta x(\omega)$ *always exists.*

Of course, usually we do not know $\kappa_1(x)$ and κ_2 beforehand. Thus, the minimization (10.4) may fail in practice. In this case, we abort the computation of the Proximal Newton step and retry with a larger regularization weight. How to choose the regularization weights is discussed in Section 10.3.

10.2.2. Sufficient Decrease Criterion

In the case that the minimization (10.4) succeeded, we do not update the current iterate x right away. The increment step $\Delta x(\omega)$ may not be a descent step, i.e, $F(x+\Delta x(\omega)) \geq F(x)$ is possible. Furthermore, even if $\Delta x(\omega)$ is a decrease step, the decrease may not be sufficiently large for a global convergence of the Proximal Newton method.

Therefore, an easy to check *sufficient decrease criterion* was developed in [26]. After the increment step $\Delta x(\omega)$ is computed, we check whether

$$F(x + \Delta x(\omega)) - F(x) \leq \gamma \lambda_{x,\omega}(\Delta x(\omega)) \tag{10.5}$$

holds for the second-order model $\lambda_{x,\omega}$ (10.4) and a fixed decrease parameter $\gamma \in (0,1)$. Although it is not at all obvious that this inequality is holds for a sufficiently large regularization weight ω, it can be shown that it does [26, Lemma 3]. Details are intentionally omitted because this result requires lengthy and technical calculations.

In summary, if the sufficient decrease criterion is not satisfied, we reject the current Proximal Newton step and retry with a larger regularization weight ω.

If the sufficient decrease criterion is fulfilled, then the iterate is updated via

$$x^{k+1} := x^k + \Delta x(\omega).$$

Moreover, we update the regularization weight according to the strategies described in Section 10.3. We stop the iteration, once a termination criterion is satisfied. Possible such criteria are given in Section 10.5.

10.3. Choice of Regularization Weights

The regularization parameter $\omega \geq 0$ plays a major role in the formulation of the Proximal Newton steps. Both the minimization of the second-order model (10.3), and the acceptance of the sufficient decrease criterion (10.5) are only successful for sufficiently large values of ω.

On the other hand, large values of ω lead to small increments $\Delta x(\omega)$ of the second-order model (10.3). This is due to the penalization of large increments by the regularization term $\frac{\omega}{2}\left\|\delta x\right\|_X^2$. So we have to find strategies to keep ω as large as possible to compute admissible increments $\Delta x(\omega)$, and at the same time keep ω as small as possible to avoid too small increments.

10.3.1. A Simple Strategy

A very simply approach is taken in [26] and [25]. In this approach, we start with an arbitrary initial weight $\omega > 0$. Each time a Proximal Newton step fails (due to the failing minimization or the descent criterion) we double the value of ω for the next try. If a step was successful, we want to make ω to decay quickly to zero for a fast local convergence. Hence, we multiply ω by 2^{-n}, where n is the number of consecutively successful Proximal Newton steps. This first strategy will be denoted by $[\omega]_1$ in the numerical tests in Chapter 11.

10.3.2. Adaptive Strategies

More sophisticated strategies are currently being investigated in [19] and in Bastian Pötzl's dissertation [24]. The main idea is to compare a predicted decrease with the actual decrease in the values of F from (10.1). For the minimizer $\Delta x(\omega)$ of the second-order model (10.3), the actual decrease is given by

$$a_{\text{red}}(x,\omega) := F(x+\Delta x(\omega)) - F(x),$$

while the predicted decrease of the second-order model is

$$p_{\text{red}}(x,\omega) := \lambda_{x,\omega}(\Delta x(\omega)).$$

In the case $a_{\text{red}} \approx p_{\text{red}}$, the prediction of the second order model matches the actual measurements on the original functional F. Therefore, we can assume that a small weight ω is sufficient for the next step. In the other cases, a penalization routine is established based on control strategies in other research areas.

One strategy is taken from the step size control in the numerics of ordinary differential equations [6, Chapter 5.2]. From there, we can define a damping prefactor function that provides appropriate regularization weights for the next step. In the numerical results, we will refer to this strategy by the notation $[\omega]_2$.

Another approach originates in numerical schemes of semi-smooth Newton methods [33]. In this approach, we want to reach $a_{\text{red}} \approx p_{\text{red}}$ in the next step, so we assume equality in the sufficient decrease criterion (10.5) and rearrange it for ω. This third strategy is denoted by $[\omega]_3$ in the numerical results.

Another advantage of these more sophisticated strategies is that we do not need to distinguish whether the sufficient decrease criterion (10.5) is satisfied or not. We can can prove that they increase or decrease the regularization weight appropriately. Unfortunately, introducing more details about these strategies opens a new topic unrelated to the rest of this thesis and is therefore intentionally omitted. For a comprehensive introduction and numerical realization of $[\omega]_2$ and $[\omega]_3$, we encourage the reader to take a closer look at [24].

10.4. Inexactness Criteria

Solving the second-order model problem (10.3) is generally nontrivial. Although the smooth part f is replaced by a quadratic approximation, it is still a nonsmooth optimization problem due to the nonsmooth part g. Special solvers are necessary to treat those kind of problems, as already seen in the first part of this thesis in Chapter 5.

A current drawback of the current form of the Proximal Newton method is that it states that $\Delta x(\omega)$ is a minimizer, so it is computed with numerical exactness. To overcome the problem of time-consuming computation of numerically exact solutions, we introduce certain inexactness criteria for the solver of the second-order model problems (10.3). These inexactness criteria reduce the wall time of solving the second-order model problems while preserving convergence results.

A detailed overview of the derivation of these criteria can be found in [25], resulting from a collaboration with Bastian Pötzl and Anton Schiela. In general, we need to preserve two properties: Local convergence near the minimizer $\Delta x(\omega)$, and global convergence far from the minimizer.

10.4.1. Local Criterion

Local convergence is ensured by a relative error criterion. Therefore, let $\Delta s(\omega)$ denote the current iterate of the inner solver of the second-order model problems (10.3). In the case of a converging solver we have $\Delta s(\omega) \to \Delta x(\omega)$. The local convergence criterion is said to be satisfied when

$$\frac{\|\Delta x(\omega) - \Delta s(\omega)\|_X}{\|\Delta x(\omega)\|_X} \leq \eta \tag{10.6}$$

applies to a parameter $\eta \in (0, 1)$. This criterion is generally of an analytical nature, since we do not know the exact minimizer $\Delta x(\omega)$. Nevertheless, the application of certain algorithms, such as TNNMG (Section 5.2), allows us to evaluate (10.6) in practice, cf. Section 10.6.4. A necessary condition of the parameter η is that we need $\eta \to 0$ during the progressing computation of Proximal Newton steps. In [25], a simple choice is

$$\eta = q^k,$$

is considered, where k is the number of the current Proximal Newton iteration and $q \in (0, 1)$ can be chosen as a problem parameter.

10.4.2. Global Criterion

Global convergence can be preserved if the decrease $F(x + \Delta s(\omega)) - F(x)$ of the inexact Proximal Newton step is below the corresponding decrease of another converging method.

Of course, this only makes sense if each iteration step of the other method is easier to compute than a Proximal Newton step. A natural choice is a Proximal Gradient method, where we use first-order model problems, such as (10.3), just without the second-order term $(H_x\delta x)(\delta x)$. Unfortunately, even such first-order problems are as difficult to solve as the original second-order problem.

One solution of this was to introduce a subgradient model of F given by

$$\lambda^{\mu}_{x,\hat{\omega}}(\delta x) := f'(x)\delta x + \mu\,\delta x + \frac{\hat{\omega}}{2}\,\|\delta x\|_X^2\,,$$

where $\mu \in \partial g(x)$ is a (theoretically arbitrary) Fréchet-subderivative of g at x. For a sufficiently high regularization weight $\hat{\omega}$ this has a minimizer $\Delta x(\hat{\omega})^{\mu}$. The subgradient model can be minimized efficiently since it is a quadratic problem with a scaled fixed quadratic part. The resulting global convergence criterion is then given by the inequality

$$\lambda_{x,\omega}(\Delta s(\omega)) \leq \lambda^{\mu}_{x,\hat{\omega}}(\Delta x^{\mu}(\hat{\omega})), \tag{10.7}$$

where we compare the current decrease in the second-order model problem (10.3) with the decrease in the subgradient model problem. In [25] a necessary value of the corresponding regularization weight of the subgradient method is derived by

$$\hat{\omega} = -\frac{\|f'(x)+\mu\|^2_{X^*}}{2\lambda_{x,\omega}(\Delta s(\omega))}, \tag{10.8}$$

where numerical tricks can be used to efficiently evaluate this dual norm term. Another crucial property is that

$$\hat{\omega} < \hat{\omega}_{\max} \tag{10.9}$$

for some large algorithmic bound $\hat{\omega}_{\max} > 0$. To keep $\hat{\omega}$ small in practice, we use a subgradient μ that makes $\|f'(x)+\mu\|_{X^*}$ small. If the subdifferential of the nonsmooth part g has a certain structure that we can use, then an optimal μ can be found.

10.5. Convergence Properties

The convergence results presented here cover both the case of the numerically exact minimization of the second-order model problems (10.3), as well as the inexact solution presented in Section 10.4. If we consider inexact solutions and both the local condition (10.6) and the global conditions (10.7)&(10.9) are satisfied, then we set

$$\Delta x(\omega) := \Delta s(\omega)$$

to be the inexact minimizer of the second-order model (10.3).

To investigate analytical and numerical convergence, a closer look is taken at the sufficient decrease criterion (10.5). This criterion is used to decide whether or not to accept a proximal Newton step. On the left side we have a direct difference of the function values

$$F(x+\Delta x(\omega)) - F(x).$$

In practice, this suffers from numerical cancellations, especially as the increments $\Delta x(\omega)$ become smaller. On the right hand side of (10.5) we have

$$\gamma \lambda_{x,\omega}(\Delta x(\omega)),$$

which also suffers from numerical cancellations in the second-order model (10.3). Therefore, this criterion is not suitable near a minimizer.

To improve this situation, we are currently experimenting with the criterion

$$[f'(x+\Delta x(\omega)) - f'(x) - H_x(\Delta x(\omega))]\Delta x(\omega) \leq \frac{1-\gamma}{2}\omega \left\|\Delta x(\omega)\right\|_X^2$$

as a more stable alternative to (10.5) when we are close to convergence of the proximal Newton method. We are able to show that this criterion is eventually satisfied for a sufficiently large ω.

The Proximal Newton iteration stops when the increment $\Delta x(\omega)$ and the regularization weight become small. Regarding inexactness, the local convergence parameter η (10.6) should also be small to guarantee accurate solutions at the local minima of F. To combine these considerations, we use the criterion

$$\frac{1+\omega}{1-\eta}\left\|\Delta x(\omega)\right\|_X \leq \varepsilon \tag{10.10}$$

for some $\varepsilon > 0$. We set $\eta = 0$ in the case of numerically exact minimization of the second-order model problems (10.3).

We are finally able to establish the global convergence result of the Proximal Newton method:

Theorem 10.5.1 ([26, Theorem 2])**.** *Let all assumptions in Section 10.1 be satisfied. Moreover, the second-order problems in Section 10.2 are solved either exactly or inexactly, so that in the latter case the local and global criteria of Section 10.4 are satisfied. If each Proximal Newton step satisfies the sufficient decrease criterion* (10.5)*, then all accumulation points of the sequence of iterates* x^k *generated by the Proximal Newton method are stationary points of the problem* (10.1)*.*

Remark 10.5.2. *The requirements of Section 10.1 are automatically met if we have the following setting:*

(i) The objective functional F *has a global minimum.*

(ii) The smooth part f *is twice continuously differentiable with a lower ellipticity bound given in Remark 10.1.1.*

(iii) We use the Hessian matrix $f''(x)$ *for the linear operator* H_x*.*

(iv) The nonsmooth part g *is convex with* $\kappa_2 = 0$*.*

10.6. Application to Finite-Strain Plasticity

Recall the fully discrete minimization problem (9.18)

$$L(\bar{w}) = L(\bar{y}, \bar{b}) \coloneqq \mathcal{E}\Big(\mathbf{y}_h^k, I^*(\mathbf{P}_h^{k-1}, \exp(\delta\mathbf{B}_h^k))\Big) + \sigma_0 \sum_{i=1}^{r} \gamma_i \left\|\bar{b}_i\right\|_2 .$$

Note that the domain of L is the space of coefficients in $\mathbb{R}^N$. The Proximal Newton method can be defined directly on the space $\mathbb{R}^N$, since it is a Hilbert space. However, this leads to grid-dependent results. For each refinement of the grid, the basis and the underlying Hilbert space of the method change. Therefore, we define our method directly on the Lagrange finite element space

$$\mathcal{Q}_h = \mathcal{Y}_h \times \mathcal{B}_h.$$

Here, $\mathcal{Y}_h$ is the first-order Lagrange space of the deformation field, and $\mathcal{B}_h$ the Lagrange space for the tangential plastic increment. The order of the latter is zero or one, depending on the gradient regularization of the plastic strain in the hardening energy. In any case, the product space $\mathcal{Q}_h$ is a subspace of an H^1 space and inherits a canonical scalar product. Therefore, it is a Hilbert space and suitable for the Proximal Newton method.

We can identify a function in $\mathcal{Q}$ by the coefficients $\bar{w} = (\bar{y}, \bar{b})$. Thus, we minimize (9.18) in the coefficients. Nevertheless, we use the norm $\|\cdot\|_X$ of the finite element space $X = \mathcal{Q}_h$ for the regularization term of the Proximal Newton steps.

The smooth part of (9.18) is given by

$$f(\bar{w}) \coloneqq \mathcal{E}\Big(\mathbf{y}_h^k, I^*(\mathbf{P}_h^{k-1}, \exp(\delta\mathbf{B}_h^k))\Big),$$

where we use the linear mappings

$$\bar{y} \mapsto \mathbf{y}_h^k(\bar{y}), \quad \bar{b} \mapsto \delta\mathbf{B}_h^k(\bar{b})$$

for the coefficients $\bar{w} = (\bar{y}, \bar{b})$ into the finite element spaces.

We conclude that f is at least twice continuously differentiable, since it is a combination of a number of smooth functions. First, there is the linear mapping from the coefficients $\bar{w}$ into Lagrange finite element spaces $\mathcal{Q}_h = \mathcal{Y}_h \times \mathcal{P}_h$. As an absolutely convergent power series, the matrix exponential is also a smooth function. Moreover, the computation of the inverse is well-defined in $\mathrm{SL}(d)$, and the derivatives of the inverse can be expressed explicitly, as already considered in (8.4). The same approach leads to the second derivative. Finally, the elastic energies are polyconvex (Section 7.4.2) with smooth functions W. Therefore, we can also assume differentiability.

The nonsmooth part is given by

$$g(\bar{w}) \coloneqq \sigma_0 \sum_{i=1}^{r} \gamma_i \left\|\bar{b}_i\right\|_2 , \tag{10.11}$$

which is obviously convex in $\bar{w}$, since the Euclidean norm is, and all appearing coefficients are positive. We can skip the formulation in the finite element space $\mathcal{B}_h$ here, since the basis was chosen to be isometric to the coefficient space, cf. (9.15).

In summary, by Remark 10.5.2, all requirements for the application of the Proximal Newton method are met.

10.6.1. Solving Second-Order Subproblems with TNNMG

From the special form of the nonsmooth part (10.11), we see similarities to the nonsmooth part of the small-strain plasticity objective functional (4.17) with kinematic hardening only. In both cases, the nonsmooth norm terms have the same block-separable structure. Moreover, the second-order problem (10.3) takes the form

$$\lambda(\delta\bar{w}) = \operatorname{quad}(\delta\bar{w}) + g(\bar{w} + \delta\bar{w}). \tag{10.12}$$

with a quadratic and a nonsmooth part. The quadratic part arises from the second-order approximation of f at $\bar{w}$ and the norm regularization term $\frac{\omega}{2}\left\|\mathbf{w}_h(\delta\bar{w})\right\|_X^2$. Note that this minimization functional is almost the same as in the small-strain case (4.17). However, there are two notably differences.

The first difference is an additional shift by $\bar{w}$ in the nonsmooth part, which can be circumvented by shifting the solution space.

The second difference is the non-explicit form of the quadratic part. In the small-strain case an explicit matrix A (4.12) is constructed directly form the finite element space. However, in (10.3) the quadratic part is given by the bilinear form $H + \frac{\omega}{2}R$, where H and R are the representations of $f''(\bar{w})$ and the regularization term $\left\|\mathbf{w}_h(\delta\bar{w})\right\|_X^2$, respectively, in the finite element coefficient basis.

We can assume that the second-order problems are strongly convex and coercive for a sufficiently large regularization weight ω. Therefore, as given in Section 5.2 in the first part of this thesis, all the requirements are met for the TNNMG algorithm.

The application of TNNMG is almost the same as given in the first part of this thesis. There is only one new task we need to take care of. It is the exact solution of the nonlinear local smoothing steps of TNNMG, as given below.

10.6.2. Local Problems

In the nonlinear local smoothing step of TNNMG we minimize the functional λ (10.12) consecutively in the block components of the solution vector $\delta\bar{w} = (\delta\bar{y}, \delta\bar{b})$. For a block component $\delta\bar{y}_i$ of the deformation part $\delta\bar{y}$, we consider a minimization problem in d dimensions. For a tangential plastic strain component $\delta\bar{b}_i$, we consider the corresponding trace-free symmetric matrix B_i which is given in $d_p = d(d+1)/2 - 1$ dimensions, cf. Remark 4.3.1.

Local deformation minimization The local minimization problem for a deformation block is quadratic, because the nonsmooth part g of the functional λ (10.12) is independent of the deformation. The problem can be expressed by a matrix $H_{ii} \in \mathbb{R}^{d\times d}$ and a vector $b_i \in \mathbb{R}^d$, s.t.,

$$\delta\bar{y}_i = \arg\min_{z\in\mathbb{R}^d} \left\{ \frac{1}{2} z^T H_{ii} z - b_i^T z \right\}.$$

The solution is given by $\delta\bar{y}_i = H_{ii}^{-1} b_i$ if H_{ii} is an s.p.d. matrix. The symmetry of H_{ii} is always given, since H is a restricted Hessian matrix plus the regularization matrix. There may be a lack of positive definiteness if the regularization weight was too small. In this case, the functional λ (10.12) is not strongly convex and violates the requirements of TNNMG. Therefore, the Proximal Newton step is rejected and a larger regularization weight is required. See Section 10.3.

Local plastic strain minimization The case of local minimization problems in a tangential plastic block component is more involved. The block components do not couple because the nonsmooth part of the minimization functional λ is block-separable. Therefore, we can express the local minimization problem by a matrix $H_{ii} \in \mathbb{R}^{d_p\times d_p}$, a vector $b_i \in \mathbb{R}^{d_p}$ and the internal shift by the previous Proximal Newton iterate $\bar{w}_i \in \mathbb{R}^{d_p}$, s.t.,

$$\delta\bar{b}_i = \arg\min_{z\in\mathbb{R}^{d_p}} \left\{ \frac{1}{2} z^T H_{ii} z - b_i^T z + \sigma_0\gamma_i \left\| \bar{w}_i + z \right\|_2 \right\}.$$

The prefactor σ_0 is the yield stress and γ_i is the volume of the Lagrange shape functions, as given in Section 9.3.1. We can transform the problem by shifting by $\bar{w}$ and setting $c_i := b_i + H_{ii}\bar{w}_i$ to obtain

$$\delta\bar{b}_i = \arg\min_{z} \left\{ \frac{1}{2} z^T H_{ii} z - c_i^T z + \sigma_0\gamma_i \left\| z \right\|_2 \right\}. \tag{10.13}$$

We solve the nonlinear problem (10.13) by assuming that H_{ii} is an s.p.d. matrix. If it is not, the requirements of TNNMG are not met. Therefore, we reject this Proximal Newton step and start over with a larger regularization weight ω.

Lemma 10.6.1. *Let $H_{ii} \in \mathbb{R}^{d_p\times d_p}$ an s.p.d. matrix, $c_i \in \mathbb{R}^{d_p}$, and $\sigma_0\gamma_i \geq 0$. Then the following statements are equivalent:*

- *(i)* $\left\| c_i \right\|_2 \leq \sigma_0\gamma_i$,
- *(ii)* $0 = \arg\min_z \left\{ \frac{1}{2} z^T H_{ii} z - c_i^T z + \sigma_0\gamma_i \left\| z \right\|_2 \right\}$.

Proof. From the assumptions on H_{ii} and $\sigma_0\gamma_i$ we conclude the existence of a unique minimizer.

"*(i)* ⇒ *(ii)*" Since the problem is convex, we show that the directional derivative from zero in an arbitrary direction d is non-negative which is a sufficient condition for the minimizer. Let $d \in \mathbb{R}^{d_p} \setminus \{0\}$. Then we have for

$$L_i(z) := \frac{1}{2} z^T H_{ii} z - c_i^T z + \sigma_0\gamma_i \left\| z \right\|_2 \tag{10.14}$$

that

$$\nabla L_i(0)[d] = -c_i^T d + \sigma_0\gamma_i \left\|d\right\|_2 .$$

From $c_i^T d \leq \left\|c_i\right\|_2 \left\|d\right\|_2$ and from $\sigma_0\gamma_i \geq \left\|c_i\right\|_2$, we conclude that

$$\nabla L_i(0)[d] \geq - \left\|c_i\right\|_2 \left\|d\right\|_2 + \sigma_0\gamma_i \left\|d\right\|_2 = (\sigma_0\gamma_i - \left\|c_i\right\|_2) \left\|d\right\|_2 \geq 0.$$

Hence, zero is the minimizer.

"$(ii) \Rightarrow (i)$" If $c_i = 0$ the statement (i) is true. Therefore, assume $c_i \neq 0$. From (ii), we have

$$0 \leq \nabla L_i(0)[d] = -c_i^T d + \sigma_0\gamma_i \left\|d\right\|_2$$

for all $d \in \mathbb{R}^{d_p}$. Choosing $d = c_i / \left\|c_i\right\|_2$ yields

$$0 \leq - \left\|c_i\right\|_2 + \sigma_0\gamma_i \quad \Leftrightarrow \quad \left\|c_i\right\|_2 \leq \sigma_0\gamma_i.$$

The corner case of $c_i = 0$ is handled by choosing an arbitrary $d \neq 0$. □

Now we dive deeper into the case $\left\|c_i\right\|_2 > \sigma_0\gamma_i$. From the lemma above, we know that in this case the unique minimizer z^* is not zero. Therefore, L_i (10.14) is differentiable around z^*, and

$$0 = \nabla L_i(z^*) = H_{ii} z^* - c_i + \sigma_0\gamma_i \frac{z^*}{\left\|z^*\right\|_2}$$

is a sufficient condition for the minimizer. Solving this low-dimensional nonlinear system is surprisingly difficult. To find z^*, we perform a damped Newton iteration. For a given initial value z^0 the iteration step is

$$z^{k+1} = z^k + \alpha^k d^k, \tag{10.15}$$

with

$$d^k := -\left(\nabla^2 L_i(z^k)\right)^{-1} \nabla L_i(z^k),$$

where

$$\nabla^2 L_i(z^k) = H_{ii} + \frac{\sigma_0\gamma_i}{\left\|z^k\right\|_2} - \frac{\sigma_0\gamma_i\, z^k (z^k)^T}{\left\|z^k\right\|_2^3}.$$

The Armijo damping parameter is chosen as $\alpha^k = \max\{2^{-l},\ l \in \mathbb{N}_0\}$, s.t.,

$$L_i(z^k + \alpha^k d^k) \leq L_i(z^k) + \rho\alpha^k \nabla L_i(z_k)^T d^k$$

for a safety parameter $\rho = 10^{-4}$.

Lemma 10.6.2. *The direction of steepest descent of L_i at zero is c_i, i.e.,*

$$c_i = \arg\min_{d\neq 0} \frac{\nabla L_i(0)[d]}{\left\|d\right\|_2}.$$

Furthermore, the minimum of L_i restricted to the line $\{\alpha c_i :\ \alpha \in \mathbb{R}\}$ is obtained at

$$\alpha := \frac{\left\|c_i\right\|_2^2 - \sigma_0\gamma_i \left\|c_i\right\|_2}{c_i^T H_{ii} c_i}.$$

Proof. Since the Euclidean norm is positive 1-homogeneous the directional derivative of L_i at zero in direction d is given by

$$\frac{\nabla L_i(0)[d]}{\|d\|_2} = -\frac{c_i^T d}{\|d\|} + \sigma_0\gamma_i,$$

which is minimized by any multiple of c_i. Since c_i is a descent direction, we know that $\alpha > 0$. The minimization on the line $\{\alpha c_i : \ \alpha \in \mathbb{R}\}$ reads

$$\min_{\alpha\in\mathbb{R}} L_i(\alpha c_i) = \min_{\alpha\in\mathbb{R}} \frac{1}{2}\alpha^2 c_i^T H_{ii} c_i - \alpha \|c_i\|_2^2 + \alpha\sigma_0\gamma_i \|c_i\|_2 ,$$

which is smooth in the variable α. So, for the minimizer α^* we have

$$\alpha^* \, c_i^T H_{ii} c_i = \|c_i\|_2^2 - \sigma_0\gamma_i \|c_i\|_2 .$$

☐

Remark 10.6.3. *In practice the block component* $z^0 := \frac{\|c_i\|_2^2 - \sigma_0\gamma_i\|c_i\|_2}{c_i^T H_{ii} c_i} c_i$ *is a good initial value for the damped Newton iterations.*

Lemma 10.6.4. *Let* $z^0 := \frac{\|c_i\|_2^2 - \sigma_0\gamma_i\|c_i\|_2}{c_i^T H_{ii} c_i} c_i$ *and* H_{ii} *be s.p.d.. Then the damped Newton-step* (10.15) *is always well-defined.*

Proof. First, we show that $\nabla^2 L_i(z^k)$ is s.p.d. Let $v \in \mathbb{R}^d \setminus \{0\}$. Then we have

$$v^T \nabla^2 L_i(z^k) v = v^T H_{ii} v + \frac{\sigma_0\gamma_i \|v\|_2^2}{\|z^k\|_2} - \frac{\sigma_0\gamma_i (v^T z^k)^2}{\|z^k\|_2^3}.$$

Since $(v^T z^k)^2 \leq \|v\|_2^2 \|z^k\|_2^2$ and $\sigma_0\gamma_i > 0$, we have

$$v^T \nabla^2 L_i(z^k) v = v^T H_{ii} v \geq 0.$$

Hence, we even have a uniform ellipticity bound by the smallest eigenvalue of H_{ii}. Second, we show that $z^k \neq 0$ for all k, i.e., $\nabla L_i(z^k)$ can be constructed. This follows directly from

$$L_i(z^0) = -\frac{\left(\|c_i\|_2^2 - \sigma_0\gamma_i \|c_i\|_2\right)^2}{2 c_i^T H_{ii} c_i} \overset{(\|c_i\|_2 > \sigma_0\gamma_i)}{<} 0 = L_i(0),$$

and that the sequence of damped Newton Steps with Armijo damping always produces a descending energy sequence. Therefore, $z^k = 0$ will never be reached. ☐

Convergence of the Newton method with Armijo damping is shown in Lemma A.6.3.

10.6.3. Automatic Differentiation

In the application of the Proximal Newton method, the practical problem arises to compute the explicit form of the second-order subproblem (10.3). For the smooth part f in the finite-strain plasticity model, we need the Hessian matrix $H_x = f''(x)$ and the gradient $f'(x)$ in the algebraic matrix or vector form for a given x. Early in the implementation phase, we decided to use an automatic differentiation tool for this task. Computing derivatives manually is error-prone and inflexible to implementation changes.

We use the free and open source tool ADOL-C[1] in our implementation. ADOL-C can be used directly with the existing code base and does not require any further adjustments to the implementation of the smooth energy part f. The only restriction is the use of generic programming, s.t., each method involved in the evaluation of f supports the data types of ADOL-C. Then we can directly compute $f'(x)$ and $f''(x)$ at a given point x.

Nevertheless, as will be seen later in the results, a large part of the total wall time of the Proximal Newton method is caused by the construction of the second-order problems (10.3). This is not a problem of ADOL-C per se, since even an evaluation of f at a point x is expensive, since we need projection techniques, cf. Section 9.3.1. Hence, it takes a multiple of that to compute the gradient, and even more to compute a Hessian matrix.

10.6.4. Evaluating Inexactness Criteria

The local and global inexactness criteria from Chapter 10 are sufficient conditions for preserving the convergence results of the Proximal Newton method while solving the second-order model problems (10.3) inexactly.

However, in practice these criteria are not trivial to check. Especially, the local convergence criterion (10.6)

$$\frac{\|\Delta x(\omega) - \Delta s(\omega)\|_X}{\|\Delta x(\omega)\|_X} \leq \eta$$

is difficult to verify in practice, since the exact solution $\Delta x(\omega)$ is not known. A study of the behavior of TNNMG applied to the second-order model problems (10.3) brought up some interesting insights. To be precise, TNNMG produces a series of corrections δ^j, such that the iterates are given by $\Delta s(\omega) = \Delta s^i(\omega) = \sum_{j=1}^{i} \delta^j$. From the convergence properties of TNNMG (see Theorem 5.2.1), we conclude that $\Delta s^i(\omega) \to \Delta x(\omega)$ as $i \to \infty$. We conjecture from the multigrid nature of TNNMG that there is a constant convergence rate $\theta < 1$ such that $\|\delta^{j+1}\|_X \approx \theta \|\delta^j\|_X$ holds after the first few iteration steps. Using the triangle inequality in a clever manner allows us to formulate a stronger local convergence criterion in the form

$$\frac{\|\Delta x(\omega) - \Delta s^i(\omega)\|_X}{\|\Delta x(\omega)\|_X} \leq \frac{\frac{\theta}{1-\theta} \|\delta^i\|_X}{\|\Delta s^i(\omega)\|_X - \frac{\theta}{1-\theta} \|\delta^i\|_X} \overset{!}{\leq} \eta \,.$$

---

[1] `https://github.com/coin-or/ADOL-C`

Note that this can be evaluated, since the rate θ can be computed after each TNNMG iteration. Although the estimate above is pretty generic, it is surprisingly sharp in practice by a relative factor of 0.8 to 0.9 [25].

In summary, the special structure of TNNMG allows to implement an efficient and explicit method to verify the local inexactness criterion (10.6). The global inexactness criteria from Section 10.4.2 are already in an explicit form, hence no modifications are necessary for the practical application.

11. Numerical Results

Various numerical tests are considered to test the application of the Proximal Newton solver to finite-strain plasticity problems. Some choices have to be made since the theory of this part of the thesis allows for a variety of different models. In detail, besides the obvious data such as the domain Ω, the boundary values and the external load $\mathbf{l}(t)$ (8.13), we have to choose a hyperelastic and polyconvex material from Sections 7.4.1 and 7.4.2, together with the corresponding material parameters. The plastic behavior is controlled by the hardening function from Section 8.2 and by the dissipation function, where we only consider the von Mises dissipation, as given in Section 8.5. Furthermore, the inclusion of plastic spin needs to be discussed. The case of spinless plasticity, as presented in Section 9.2.1, can be used numerically without further assumptions. In contrast, a model with plastic spin (Section 9.2.2) requires a symmetry assumption on the plastic increment.

On the algorithmic side, the Proximal Newton solver has multiple possible configurations. First, we have to decide whether to solve the second-order model problems (10.3) numerically exactly, or inexactly. In the latter case, we have to set the parameters $\eta \to 0$ and $\hat{\omega}_{\max}$ (Section 10.4). Regardless of the exactness decision, a regularization weight strategy (Section 10.3) and parameters for the sufficient decrease criterion (10.5) need to be specified.

Additional flexibility is given for the application of TNNMG as the preferred solver for the second-order minimization problems (10.3). We can generally solve the appearing linear correction steps with a direct solver or do that inexactly with single multigrid steps (cf. Section 5.2 in the first part of this thesis for details). We have seen in the application of TNNMG to small-strain plasticity problems in Chapter 6 that inexact multigrid steps are to prefer in terms of effectiveness of the solver. However, problems of the kind (10.3) may become numerically unstable if the regularization weight is just large enough, or even too small to make the problem strongly convex. Since we do not know a sufficient value of the regularization weight beforehand, it can be advantageous to use direct linear solvers in TNNMG.

In summary, we have a lot of degrees of freedom for testing the Proximal Newton solver on finite-strain plasticity problems. The goal is to consider as many settings as possible without losing track of the numerical tests. While preparing this thesis it was observed that the plastic spin has no effect on the minimization methods in terms of wall time, iteration numbers and convergence behavior. Therefore, we consider only spinless plastic strains with values in the symmetric set $\mathrm{SL}(d)^+_{\mathrm{sym}}$. As already discussed, the advantage of this choice is that we do not have to make compromises and further assumptions for the analytical existence of solutions.

11.1. Test Parameters

11.1.1. Elastic Energies

We start by defining the precise form of the elastic material laws which are used in the numerical tests. The first one is the isotropic linear St. Venant–Kirchhoff material (7.12)

$$W_{\mathrm{SVK}}(F) \coloneqq \frac{\lambda}{2}\operatorname{tr}(E)^2 + \mu \, \|E\|_F^2$$

with the Green–St. Venant strain tensor $E = \frac{1}{2}(F^T F - I)$. The Lamé parameters are chosen from [32, page 388] to model an existing steel material. In particular, we set $\mu = 1.107438 \cdot 10^5$ N/mm^2 and $\lambda = 8.01937 \cdot 10^4$ N/mm^2.

The second choice is the nonlinear Mooney–Rivlin model (7.13)

$$W_{\mathrm{MR}}(F) \coloneqq a \, \|F\|_F^2 + b \, \|\operatorname{cof}(F)\|_F^2 + c \det(F)^2 - d \ln(\det(F)) + e.$$

All appearing parameters are again measured in N/mm^2. To favor volume-preserving behavior, we set $d = 2a + b + c$, such that for uniform deformations $F = \alpha I$, the energy $W_{\mathrm{MR}}(\alpha I)$ is minimal for $\alpha = 1$. For reasons of comparability we want to have

$$W_{\mathrm{MR}}(F) = W_{\mathrm{SVK}}(F) + \mathcal{O}(\|E\|_F^3) \tag{11.1}$$

for the Mooney–Rivlin material. Recall the computation from Section 7.4.2, especially (7.14), which can be resolved for the Mooney–Rivlin parameters a, b and c. However, there is one degree of freedom left open. Therefore, we choose to penalize the determinant term with a large parameter $c = 10^4$ to further support volume-preserving material behavior. Note that the other parameters a, b and d are directly influenced by this choice. Increasing c far beyond 10^4 would cause a or b to become negative, and hence W_{MR} loses the polyconvexity. For $c = 10^4$ we get $a = 4.5323475 \cdot 10^4$ and $b = 1.0048425 \cdot 10^4$. The parameter e is set to $-3a - 3b - c$ by (7.14) only to enforce (11.1) in theory. Since it is constant it does not have an influence on the minimizing algorithms.

Other elastic materials are not considered in the numerical tests. Recall from Remarks 8.6.2 that we do not have a guaranteed existence of an energetic solution of the finite-strain plasticity problem for the energies above. A material compatible with the existence theory (Theorem 8.6.1) is the Ogden material (7.15). However, this requires impractical large exponent parameters of the material. Nevertheless, both the St. Venant–Kirchhoff and the Mooney–Rivlin material, as described above, yield satisfying results in the numerical tests, despite the lack of a proper existence result.

11.1.2. Kinematic Hardening

The kinematic strain hardening from Section 8.2

$$W_{\mathrm{hd}}(P, Z) \coloneqq k_1 \, \|P\|_F^{p_1} + k_2 \, \|Z\|_F^{p_2}$$

is chosen to be a classical linear hardening. Hence, we have for both exponents $p_{\mathrm{hd}} = p_{\mathrm{gr}} = 2$. This is similar to the small-strain results, where we have a corresponding exponent of two in (4.17). We will not consider other exponents in the tests of this part of the thesis. However, note that the energetic existence theory (Theorem 8.6.1) requires $p_{\mathrm{hd}} \gg 2$ to be impractical large.

Of special interest is the case $k_2 = 0$, i.e., no gradient regularization. Although a positive k_2 is also a requirement for the existence theory, no negative consequences in the practical test are found for small values of k_2, or even for $k_2 = 0$. The case of no gradient regularization has a positive effect on the wall time since the construction of the discrete plastic strain gradient is expensive (Section 9.3.2).

11.1.3. Other Parameters

All other parameters regarding the domain, the boundary conditions and the loads are specified in the following sections. Also the yield stress σ_0 of the von Mises dissipation, as given in Section 8.5 is different along the numerical tests. Since we restrict the tests to the symmetric subset $\mathrm{SL}(d)^+_{\mathrm{sym}}$ for the plastic strain range, no antisymmetric parts of the dissipation need to be considered. Hence, we can neglect the antisymmetric parameter σ_1.

11.1.4. Solver Settings

Most of the different parameters of the Proximal Newton solver are part of the consideration of the tests in the sections below. Nevertheless, the convergence criteria for each homotopy step (where a plastic increment is computed) are the same for all tests. We use the criterion

$$\frac{1+\omega}{1-\eta} \left\| \Delta x(\omega) \right\|_X \leq 10^{-10}$$

to terminate the Proximal Newton iteration, cf. (10.10).

The inner TNNMG solver (recall Section 5.2) for the second-order model problems (10.3) is fixed to a single exact nonlinear block Gauss–Seidel step for the nonlinear presmoothing phase. The other parameters are reused from the small-strain von Mises tests in Chapter 6, since we have to deal with equivalent problems here. For the linear correction phase we will use either CHOLMOD [3] or a single inexact multigrid step. This depends on the specific test case. TNNMG can report errors to the Proximal Newton solver in case that an abnormal behavior is detected and reject the current search for the minimizer $\Delta x(\omega)$ of the second-order model problems. In particular, an error is reported, if

- a nonlinear block Gauss–Seidel step does not find a minimum,
- the linear correction matrix is not positive definite (CHOLMOD has the ability to report this),
- line search does not find a minimum or
- TNNMG does not converge after 100 iterations.

All of these cases indicate that the second-order model problem is not strictly convex, and thus a larger regularization weight ω is required.

11.1.5. Test Machine

All test are executed single-threaded on a Intel(R) Core(TM) i7-9750H CPU with clock frequency fixed to 2600 Mhz to avoid overheating and to ensure comparability of all test runs. The test machine runs the current snapshot of Debian 12, including updates as of January 30, 2023. The C++ Code (cf. Appendix B) is compiled with the flags`-O3 -DNDEBUG` using the `gcc`[1] compiler in version 12.2.0.

11.2. Tests on a Simple Geometry

11.2.1. Regularization Strategies

In this test we want to compare different regularization strategies of the Proximal Newton solver. In Section 10.3 three different strategies are considered. Namely, the simple scaling strategy $[\omega]_1$, and the two more sophisticated adaptive strategies $[\omega]_2$ and $[\omega]_3$. To test the numerical influence of these three strategies we use a test body which has already been studied in the small-strain theory in the first part of this thesis. The three dimensional test body described in Section 6.1.2. Boundary conditions are placed as follows: the bottom face $(0,7) \times \{0\} \times (0,1)$ is a Dirichlet boundary for the deformation field, i.e., $\mathbf{y}(x) = x$ for $x_2 = 0$. External surface loads are applied at the top face $\Gamma = (0,4) \times \{7\} \times (0,1)$. For the tensile tests, we will compute numerical solutions scaled by some $\alpha > 0$ in the load functional

$$\langle \mathbf{l}_{\mathrm{pull}}(t), \mathbf{y} \rangle \coloneqq 10^3 \alpha \cdot \int_\Gamma \mathbf{y}_2(x)\, dS. \tag{11.2}$$

Note that the load only acts on the second (upward pointing) component $\mathbf{y}_2$ of the deformation field.

The geometry is coarsely discretized as seen in Section 6.1.2 by a grid of 25 cubical grid elements. Numerical tests on this are computed on a grid hierarchy consisting of two further uniform grid refinements resulting in 1600 cubical grid elements on the finest level. We use a St. Venant–Kirchhoff material with the parameters of the section above. Gradient regularization is not considered for the hardening term as we set $k_2 = 0$. the other linear kinematic hardening parameter k_1 is set to $2 \cdot 10^3$ N/mm^2. The yield stress is chosen to be $\sigma_0 = 5 \cdot 10^3$ N/mm^2.

In order not to interfere with inexactness effects on the Proximal Newton solver, we decided to use compute numerically exact solution of the second-order model problems (10.3) with TNNMG up to a relative accuracy of 10^{-8}. Since wall time considerations are not of interest in this test we use a direct solver CHOLMOD for the linear Newton problems of TNNMG. We do this to avoid an influence of possible numerically unstable multigrid steps on the number of Proximal Newton iterations.

The tensile tests are computed independent of each other with two homotopy steps for each test. In the first step the solution corresponding to the load function (11.2) is computed. In the second step no load function is considered such that the material unloads.

---

[1] `https://gcc.gnu.org/`

The results of the tests in terms of the deformation field and the plastic strain for different load parameters $\alpha \in \{1, 3, 5\}$ are displayed in Figure 11.1. The deformation field is shown directly by the representation of the test body. Plastic deformation is indicated by the color map. In detail, at each point $x \in \Omega$ we measure $\|\mathbf{P}_h^k(x) - I\|_F$, s.t., no plastic strain $\mathbf{P}_h^k(x) = I$ results in a zero in the color map.

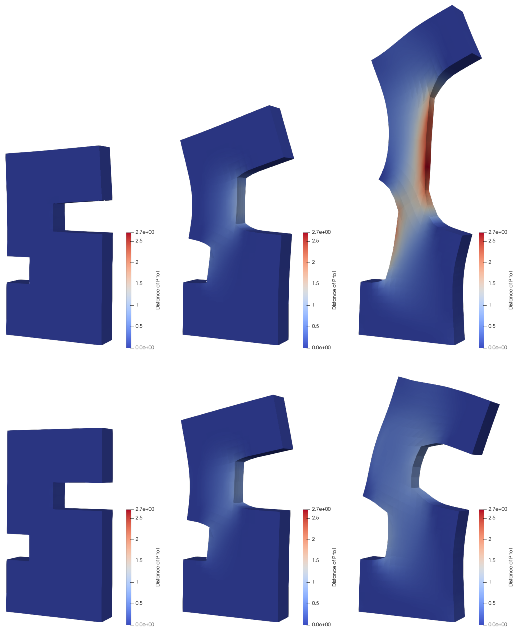

2.7e+00
2.5
2
1.5
1
0.5
0.0e+00
Distance of P to I
2.7e+00
2.5
2
1.5
1
0.5
0.0e+00
Distance of P to I
2.7e+00
2.5
2
1.5
1
0.5
0.0e+00
Distance of P to I
2.7e+00
2.5
2
1.5
1
0.5
0.0e+00
Distance of P to I
2.7e+00
2.5
2
1.5
1
0.5
0.0e+00
Distance of P to I
2.7e+00
2.5
2
1.5
1
0.5
0.0e+00
Distance of P to I

Figure 11.1.: Results of the pull test for $\alpha \in \{1, 3, 5\}$ in ascending order. Top row shows the result of the loading test. Bottom row shows the remaining strain after unloading.

For small loads ($\alpha < 1$), this results in a purely elastic deformation without any occurrence of a notable plastic strain. Therefore, those tests are not displayed as the result of the second homotopy step is equal to the initial state. From $\alpha = 1$ on, a plastic strain starts to evolve at the inner edges of the notches. In the last loading test for $\alpha = 5$) about half of the domain has a plastic strain different from the identity. While unloading a part of the remaining plastic strain vanishes. This is due to the kinematic hardening model. Although the yield surface is large it requires plastic unloading to reach a stress-free state in the second homotopy step.

As shown in Table 11.1, the number of accepted Proximal Newton steps is generally comparable among the three methods with a small advantage of the second strategy $[\omega]_2$ in terms ot both total and accepted Proximal Newton steps. Especially for the large load parameter $\alpha = 5$ one can see that overall the adaptive strategies $[\omega]_2$ and $[\omega]_3$ produce the same number of accepted steps as the simple strategy $[\omega]_1$, but the number of total Proximal Newton steps is larger for $[\omega]_1$. This indicates that this strategy aims for too low regularization weights ω causing a large number of solutions $\Delta x(\omega)$ of the second-order model problems (10.3) to be rejected. The adaptive strategies compare the

Strategy	$\alpha = 1$	unloading	$\alpha = 3$	unloading	$\alpha = 5$	unloading
$[\omega]_1$	6 (6)	6 (6)	10 (7)	13 (11)	27 (15)	45 (20)
$[\omega]_2$	5 (4)	6 (5)	9 (8)	6 (6)	18 (13)	25 (19)
$[\omega]_3$	8 (8)	8 (8)	13 (11)	10 (10)	25 (15)	26 (22)

Table 11.1.: Number of Proximal Newton iterations for the regularization update methods. The number in parentheses is the number of accepted steps, indicating that the subproblem was well-posed.

result of the second-order model problem with the actual decrease of the values of the objective function. There, a low regularization weight is only selected if the second-order model problem is a *good* approximation of the actual objective function. Details are widely discussed in the thesis of Bastian Pötzl [24].

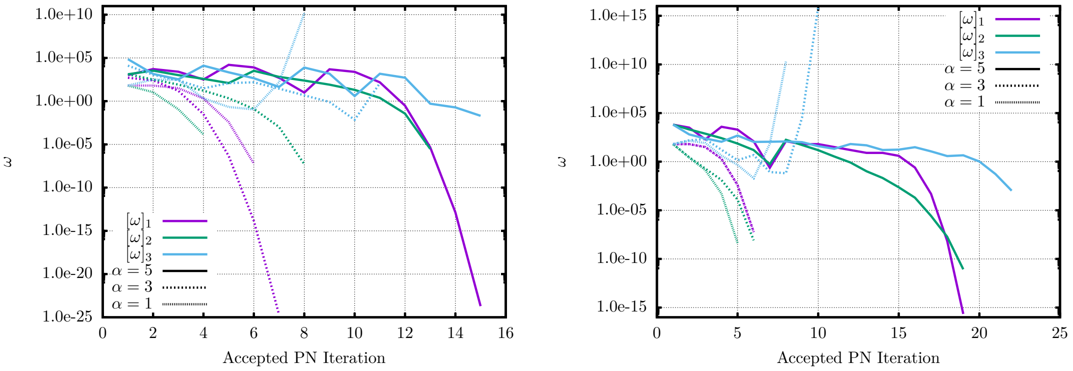


Figure 11.2.: Regularization weights of the Proximal Newton steps for $\alpha \in \{1, 3, 5\}$. Left: loading step. Right: Unloading step. Solid lines represent the values for $\alpha = 5$, medium dashed lines $\alpha = 3$, and the short dashed lines $\alpha = 1$.

For a deeper analysis of the results we take a look at the selected regularization weight ω of the tensile tests for the different strategies. In Figure 11.2 for both homotopy steps (loading and unloading) all relevant data are shown. We notice that the number of iterations is generally larger for a larger value of α. Therefore, we will analyze only the tensile tests with $\alpha = 5$, given by solid lines in the plots. The chosen accepted weights are about the same magnitude for the first 12 steps of loading, and the first 10 steps of unloading. The value of ω is not constant but it does not leave the interval $[1, 10^5]$. These are the phases of *globalization*, where the objective functional is nonconvex and the iterates are far from a local minimizer. Hence, large weights ω are needed to compute well-defined solutions of the second-order model problems. After the globalization phase the Proximal Newton solver eventually switches to the *local convergence* phase near a local minimum. By assuming that the objective functional is locally convex at the (local) minimizers, no regularization is needed in this phase of the Proximal Newton solver.

Both the simple strategy $[\omega]_1$ and the controller strategy $[\omega]_2$ seem to work well in the local convergence phase, and the regularization weight quickly drops to zero. Thus, both strategies seem to work well for the tensile tests with a slight preference for $[\omega]_2$ since it detects the local convergence phase faster.

The third remainder term based strategy $[\omega]_3$ shows a different behavior compared to the other two strategies. It does not detect the local convergence phase well, and, even worse, seems to encounter numerical instabilities at the end of the Proximal Newton iteration. Nevertheless, it finds the same local minimizers. This indicates that the instabilities are not a general problem of this method. However, a discussion and deep analysis of this strategy is not given in this thesis since it is beyond the scope. A detailed overview is given in [24].

In summary, for the other numerical tests we can choose either the simple non-adaptive strategy $[\omega]_1$ or the more sophisticated controller strategy $[\omega]_2$. Although no detailed explanation of $[\omega]_2$ is given in Section 10.3, we will stick to this strategy, since it gives slightly better results than $[\omega]_1$ in the other tests. However, all other numerical tests in this part of the thesis can be computed using $[\omega]_1$ if optimal performance is not required, or if parts of this method are to be placed in a custom implementation.

11.2.2. Gradient Regularization

In this section we will study the influence of the scalar k_2 in the hardening term (8.25). Although it is a necessary term for the existence result of Theorem 8.6.1, it is not included in the classical continuum mechanical model of hardening, which is the total free energy of the internal variables [15]. To investigate the influence of the parameter k_2 on the numerical tests, a small series of tests is computed to show both the effect on the result and on the performance of our Proximal Newton Method. The test object from Section 11.2.1 is therefore reused. To keep the number of tests small, only the case $\alpha = 5$ is considered. Both the grid and the boundary conditions remain unchanged. All material parameters are copied. We still consider an exact Proximal Newton step computation with TNNMG using the direct solver CHOLMOD for the linear corrections. In this case we consider the strain gradient coefficients $k_2 = 0$ and $k_2 = 10^i$ for $i = 1, 2, 3, 4$. The case of no gradient regularization, i.e. $k_2 = 0$, is of particular interest since there is no need to consider first-order finite elements for the tangential plastic increments $\delta\mathbf{B}_h$ and the previous plastic strain state $\mathbf{P}_h^{k-1}$, as discussed in Section 9.3.1. Also, a smaller finite element order significantly reduces the number of degrees of freedom of the discrete problem. Therefore, an additional test is performed with zero order finite elements for $\delta\mathbf{B}_h$ and $\mathbf{P}_h^{k-1}$.

The results for the different settings are displayed in Figure 11.3. It turns out that a small gradient regularization up to $k_2 = 10^2$ does not create relative differences in both the plastic strain and the deformation field beyond 1%. For moderate values $10^2 < k_2 < 10^4$ a significant difference is evolving in the plastic strain variable, resulting in nearly no plastic strain in the case of $k_2 = 10^4$. These properties of the results are as expected. Higher gradient regularization favors a flattening of the plastic strain. For low or no gradient regularization the distribution of the plastic strain varies between no plastic

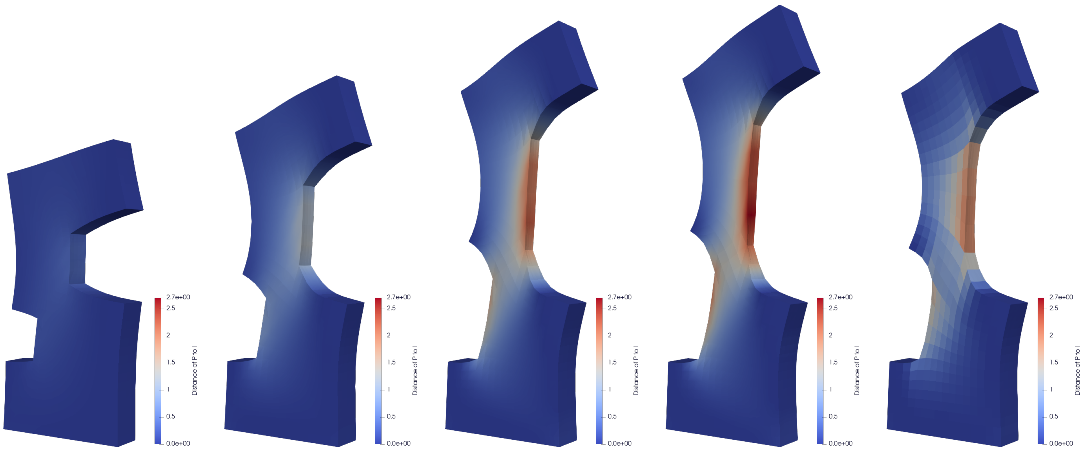


Figure 11.3.: Results of the tensile test for different values of k_2. Left to right: $k_2 = 10^4, 10^3, 10^2, 0$, and the case or zero order finite elements for the plastic strain spaces.

strain at the top and bottom face and the maximum at the inner faces of the notches. For a zero order finite element space for the plastic components no gradient regularization can be considered since the finite element functions are constant on each grid element. Therefore, we can only compare this result to the case of $k_2 = 0$ with order one finite elements. Nevertheless, the results are similar, but not equal. Although deformation is the same the value of the plastic strain at the inner faces of the notches is about 30% smaller for zero order elements. This is expected since the piecewise constant finite elements average over the cubical grid elements and cannot obtain local maxima at faces.

Besides the differences in the result the behavior of the Proximal Newton method using different gradient regularization parameters is analyzed. Therefore, we measure

- the total number of Proximal Newton steps (including the rejected steps)
- the number of accepted steps,
- overall wall time of the inner TNNMG iterations (and the average time per Proximal Newton step),
- overall wall time of the assembly of f' and f'' using ADOL-C (and the average time per accepted Proximal Newton step) and
- the energy at the minimum.

The numbers of the Proximal Newton steps indicate the condition of the global minimization problem, as well as the quality of the second-order model problems. The wall time of TNNMG is an indicator of the total number of TNNMG iterations along the Proximal Newton process and allows insight into the condition of the second-order model problems. On the other hand, the wall time of ADOL-C is generally independent of the problem itself and should scale with the number of accepted Proximal Newton steps, since only after an accepted step a new Hessian matrix and gradient need to be computed. The energy at the minimum can be used to validate the influence of the gradient regularization, since it adds an additional nonnegative term to the objective functional. Hence, we check whether a larger value of k_2 leads to a larger value at the minimizer.

k_2	PN steps	accepted	TNNMG [s]	ADOL-C [s]	energy at minimum
10^4	122	87	2041 (16.78)	4326 (49.72)	-7156.67
10^3	13	10	95 (7.00)	459 (45.90)	-14929.17
10^2	17	13	93 (5.47)	590 (45.38)	-21531.63
10	20	15	136 (6.80)	734 (48.93)	-23167.35
1	17	13	93 (5.47)	610 (46.92)	-23363.95
0	18	14	91 (5.06)	461 (32.92)	-23386.29
order 0	17	13	9 (0.59)	119 (9.16)	-21317.59

Table 11.2.: Proximal Newton measurements for different gradient regularization parameters. The test with order 0 is run on a finite element space with zero order of the plastic strain. The numbers in parenthesis is the average time per PN step (TNNMG), or per accepted PN step (ADOL-C).

The results of this measurements are displayed in Table 11.2. We start with the analysis of the results with first-order finite elements. For the gradient regularization parameters in the range $0 \leq k_2 \leq 10^3$ the number of Proximal Newton steps is comparable with 13 to 18 steps in total and 10 to 15 accepted steps. There is a slight tendency that a higher regularization requires fewer Proximal Newton steps until convergence. Therefore, we conclude that the gradient regularization is not a necessary part of the convergence of our method. Looking more closely at the wall times of TNNMG, we notice that the wall time per Proximal Newton step varies between 5 and 7 seconds. There is no pattern visible and we conjecture that the number of inner TNNMG iterations is independent of the parameter k_2. Also, for all parameters $k_2 > 0$, the average wall time per accepted step of the assembler ADOL-C is nearly constant around 46 to 47 seconds. This is expected since the construction of the Hessian matrix and the gradient is generally independent of the parameters of the objective functional. For no gradient regularization ($k_2 = 0$), the expensive evaluation of the plastic strain gradient can be skipped, resulting in only about 33 seconds per accepted Proximal Newton step for the ADOL-C routines.

A notable behavior of the Proximal Newton solver occurred for the high gradient regularization value $k_2 = 10^4$. The Proximal Newton solver encountered heavy convergence problems of the inner TNNMG solver leading to overall high regularization weights $\omega > 10^3$ throughout the whole solution process. This causes the computed increments $\Delta x(\omega)$ to be small and thus we need more Proximal Newton steps than for the other tests with lower values of k_2. The reason of this behavior is that the objective function $\mathcal{L}$ (9.18) is dominated by the gradient hardening term $k_2 \left\|\nabla P_h\right\|^2$ in the hardening functional (8.25). The elastic energy and the dissipation functional play only a minor role. Hence, we have an ill-conditioned composite minimization problem with differently scaled components. This causes the poor convergence properties of our method. As a conclusion we can say that we should avoid over-regularization in the plastic strain gradient terms, as they lead to worse solver behavior.

The case of zero-order finite elements of the plastic strain components results in almost the same number of Proximal Newton steps as in the first-order case with $k_2 = 0$. However, the wall times of TNNMG and ADOL-C are significantly lower by a factor of 4 for ADOL-C and by a factor of 9 for TNNMG. This is expected since the number of degrees of freedom in the plastic components is about 8 times lower for zero-order elements than for first-order elements.

The energy at the minimum increases with increasing gradient regularization, as postulated before. The zero-order result has a higher energy than the first-order result with $k_2 = 0$. This is not surprising since there are fewer degrees of freedom in the zero-order case.

11.2.3. Inexact Proximal Newton

In this test we investigate the inexactness features of the Proximal Newton method (Section 10.4) and their impact on the quality of the solution process, as well as on the wall time. In the previous tests it is obvious that our inner TNNMG solver is much faster than the time needed to assemble the second-order model problems (10.3) with ADOL-C. Therefore, if there is no better alternative for setting up the second-order model problem, then it seems that no improvement in the speed of the TNNMG solver is necessary. However, from the tests of small-strain plasticity with TNNMG combined with a direct solver (Section 6.1.2) it appears that the wall time does not scale linearly with the number of unknowns. On the other hand, the number of entries of the Hessian matrix f'' is linear in the number of unknowns. These entries originate from the finite element approximation and each row of the hessian stores only entries of neighboring degrees of freedom. This number is invariant under grid refinements. Therefore, we expect the wall time of ADOL-C to be linear in the number of unknowns. As a consequence, for very fine grids, the wall time of ADOL-C will be below the wall time of TNNMG.

Unfortunately, since the resources of the test machine are limited, we never reach the case where TNNMG with a direct linear solver becomes the bottleneck of the total wall time. However, from the above considerations, we still see the need to reduce the wall time of TNNMG.

For our test, we choose the already known tensile test from Section 11.2.1. Geometry and boundary conditions are copied. For a better overview we only consider one loading step with a tensile force corresponding to $\alpha = 5$ and we do not consider the second unloading step. However, we invest in a further grid refinement with 4 levels in total to reduce the gap in the wall time of TNNMG and ADOL-C. This leads to 12 800 cubical grid elements at the finest level.

There are four scenarios. For the linear correction steps of TNNMG, we consider either the direct solver CHOLMOD, as in the previous tests, or a single imprecise multigrid step. This is combined with either solving the second-order model problems (10.3) numerically exact, or, as presented in Section 10.4, with inexactness criteria. In the case of an inexact Proximal Newton step computation we use the parameter $\hat{\omega}_{\max} = 10^{10}$ for the global criterion (10.9) and for the local criterion (10.6) we use

$$\eta = 0.8^i,$$

where i is the number of the so far accepted Proximal Newton steps. This guarantees $\eta \to 0$.

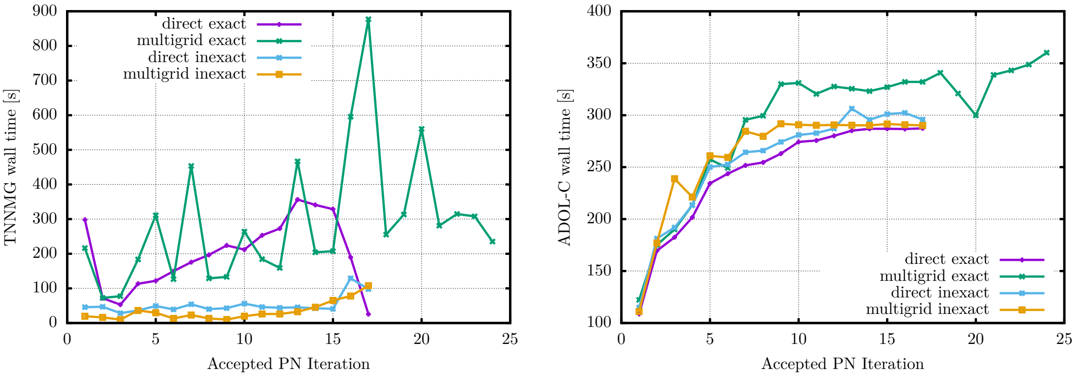


Figure 11.4.: Wall times per accepted Proximal Newton step. Left: TNNMG. Right: ADOL-C.

In Figure 11.4 the measurements for the four scenarios are displayed. In detail, we measure the wall time of both TNNMG and ADOL-C for each accepted Proximal Newton step. This may include possible wall times of rejected steps before a step is accepted.

First of all, the inexact Proximal Newton method combined with multigrid TNNMG requires 24 accepted Proximal Newton steps for convergence, while the other three scenarios converge after 17 accepted Proximal Newton steps. We expect that the number of iterations in the inexact case may be higher than in the exact case, since some quality of the second-order updates $\Delta x(\omega)$ is lost.

There are significant differences in the wall times of the TNNMG parts. The first steps take always longer because several Proximal Newton steps are rejected until the optimal initial weight ω is found. The other aspects are discussed separately for each of the four scenarios.

Exact Multigrid: Exact Proximal Newton with multigrid TNNMG requires both the largest total wall time and the most Proximal Newton steps. For each Proximal Newton step, at least 50 TNNMG iterations are necessary for numerically exact solutions. Since we set an iteration limit of 100, six Proximal Newton steps are rejected because TNNMG did not converge. These cases are visible as peaks in the wall time plot. The convergence rate of TNNMG was generally poor, indicating that problems with low regularization weights ω are ill-conditioned and not suitable for an iterative multigrid solver. Unfortunately, the regularization strategies of the Proximal Newton solver aim at keeping ω as low as possible in order to generate large increment steps.

Exact Direct: Using the direct solver for exact Proximal Newton steps avoids the problem of the ill-conditioned second-order model problems. Hence, on average, there are 15 TNNMG iterations per Proximal Newton step necessary. However, the size of the problem is large on the fine grid. Therefore, CHOLMOD becomes expensive. Up to the last three Proximal Newton iterations the wall time increases. This is due to the increasing number of *activated* nonzero plastic components that contribute to the linear correction problems of TNNMG. In the last three steps, TNNMG stops after a small number of iterations as the absolute corrections become numerically zero.

Inexact Direct: Inexact Proximal Newton with direct TNNMG is more efficient than the exact variants. The efficient local criterion (10.6) is satisfied after at most 3 TNNMG iterations. Therefore, each Proximal Newton steps is computed in a fraction of the time of its exact counterpart. This behavior continues until step 15. Then the parameter η becomes so small that more TNNMG iterations are necessary. Therefore, the wall time increases for the end, where 7 to 8 TNNMG iterations are performed.

Inexact Multigrid: The last scenario, inexact Proximal Newton with multigrid TNNMG, seems to be the most efficient. Slightly more TNNMG iterations are necessary for the Proximal Newton steps than in the direct case. Nevertheless, the local convergence criterion is satisfied by about 4 to 5 iterations for the first 13 steps. In the end, the local convergence phase with very low values of η require much more TNNMG iterations with a maximum of 33 in the last step. Therefore, the wall time increases towards the end.

As a summary for the four scenarios, TNNMG steps are much faster with a single multigrid step for the linear correction problems than using a direct solver. However, for the generally ill-conditioned second-order model problems (10.3), multigrid steps are not suitable if numerical exactness is required. The inexactness criteria seems to work well. Since there are no other disadvantages of the inexact Proximal Newton method, it seems favorable in general, as long as the local convergence criterion (10.6) can be evaluated in practice.

The wall times of ADOL-C of the accepted Proximal Newton steps are constantly increasing. The reason is to be found in the details of how ADOL-C works internally as described in [13]. In the initial state the tangential plastic increment $\delta\mathbf{B}_h^0$ is zero and during the Proximal Newton iterations the number of nonzeros decreases. To compute the entries of the Hessian matrix f'', ADOL-C stores the chain rule expression of the second-order derivatives. The evaluation of these chain rule expressions can be shortened if a multiplication with zero takes place. ADOL-C is able to recognize multiplications with zero. Therefore, computing the Hessian matrix with ADOL-C becomes more expensive as the number of nonzeros in the tangential plastic increment decreases.

11.3. Different Material Models on a Complex Geometry

In this test we consider another, more advanced geometry. We model a realistic paper clip with a total height of 30 mm and a total width of 10 mm. The wire diameter is 1 mm. The geometry is given in Figure 11.5 and is discretized by 980 prism-like elements. One uniform grid refinement is computed resulting in a total of 7840 grid elements at the finest level. Along the cross sections of the wire there are always 7 triangular elements to model the cylindrical geometry of the wire. The two circular ends of the wire are denoted by Γ_{in} for the inner end and by Γ_{out} for the outer end. Dirichlet boundary values for the deformation field are placed at the outer end Γ_{out}, s.t.,

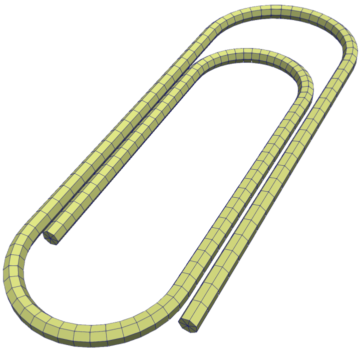


Figure 11.5.: Geometry and initial grid of the 3D paperclip.

$$\mathbf{y}_h|_{\Gamma_{\text{out}}} \equiv 0.$$

The external load is given by a Neumann function at the inner end Γ_{in} with the load functional

$$\langle \mathbf{l}_{\text{pull}}(t), \mathbf{y} \rangle := 200 \int_{\Gamma_{\text{in}}} \mathbf{y}_2(x)\, dS \tag{11.3}$$

for the first homotopy step. As before, we also consider a second homotopy step without load to observe the unloading of the material.

In this test we compare the two material models presented in Section 10.1. For both the St. Venant–Kirchhoff model and the Mooney–Rivlin model we use a high value for the yield stress by setting $\sigma_0 = 2 \cdot 10^3$ N/mm^2. This increases the elastic region and reduces the effect of plastic unloading in the second homotopy step. Moreover, we choose a rather low kinematic hardening parameter $k_1 = 2 \cdot 10^3$ N/mm^2. This choice increases the deformation caused by the plastic strain. It is only an academic choice to improve the visibility of plastic effects in the results. One can safely use other values of k_1 in custom tests. Since gradient regularization does not have a significant impact on the solution quality (Section 11.2.2), we do not consider gradient regularization and set $k_2 = 0$. This reduces the computation time. For a further reduction of the computation time, the order of the plastic finite elements is set to zero.

The solver parameters are given as follows. We use an inexact Proximal Newton method because it is superior to the exact counterpart (cf. Section 11.2.3). For the second-order model problems, we use TNNMG with a direct linear solver. Although the multigrid variant was slightly faster in the test above, the multigrid scheme is not suitable here since we only consider two grid levels. All other solver parameters are unchanged from the previous tests.

We start by discussing the results for the St. Venant–Kirchhoff material. The result of the loading step is given in Figure 11.6. The deformation is again given directly by the representation of the test object. The plastic strain is given by the color map. For a better orientation, the initial undeformed state is also included. A large deformation of the paper clip is visible. The external load almost completely unwound the wire, creating an almost straight rod. High values of plastic strain are found in the curvatures of the paper clip, while no plastic strain appears in the formally straight parts of the paper clip. This is what we expect since the straightening on the wire causes higher strains at the curvatures.

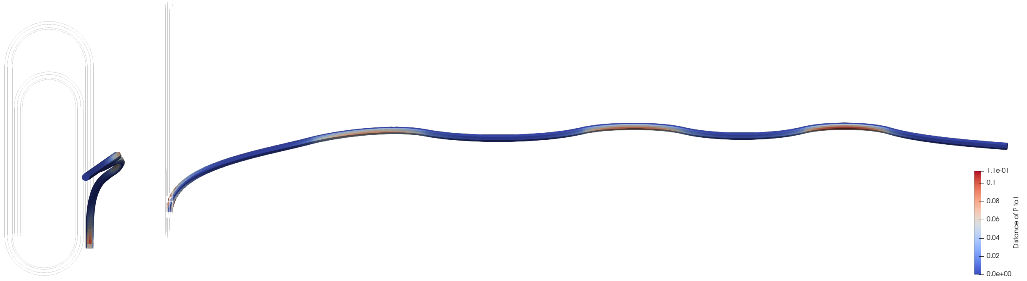


Figure 11.6.: Deformation of the paperclip including reference state for orientation after loading. Left: Front view. Right: Side view.

The result of the following unloading test is shown in Figure 11.7. We can see the elastic unloading of the completely plastic strain free straight parts of the paperclip. The only remaining plastic strain is found in the curvatures, resulting in larger remaining angles after the second homotopy step.

To our surprise, the results of the same test with the nonlinear Mooney–Rivlin material are almost the same. We expected the resulting deformation and plastic strain to be comparable because we chose the parameters of the Mooney–Rivlin material in Section 10.1 so that they overlap for moderate strains. However, in practice we have visually the same images as given in Figures 11.6 and 11.7. Therefore, we computed the differences of the results of the loading step.

In Figure 11.8 we see the absolute difference of the values of the plastic strain. As indicated in the color map, the maximal difference is about $4.4 \cdot 10^{-4}$, which is negligible for the total plastic strain with values up to $1.1 \cdot 10^{-1}$. The other image in the figure is the relative difference in the deformation field. The only notable difference is found at the outer end of the paper clip where the Dirichlet values are set. There we find a difference of 2.1%. Therefore, we can conclude that both material models give the almost

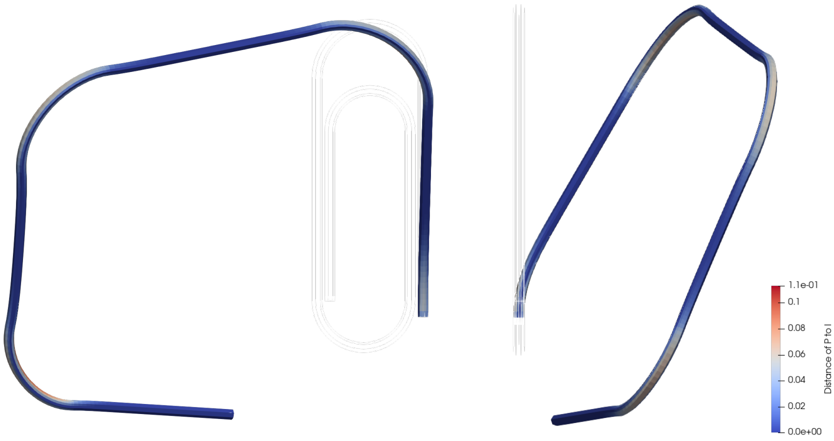


Figure 11.7.: Deformation of the paperclip including reference state for orientation after unloading. Left: Front view. Right: Side view.

the same results.

However, the influence of the two materials on the Proximal Newton solver is different. The evaluation of the Mooney–Rivlin material is slightly more expensive than the simple St. Venant–Kirchhoff material. Of course, computation of the gradient and the Hessian matrix with ADOL-C is then also more expensive. In total, the Proximal Newton method took 40493 seconds for the Mooney–Rivlin material and only 36395 seconds for the St. Venant–Kirchhoff material.

In the following we take a closer look at the iteration process of the Proximal Newton method. In Figure 11.9 the chosen regularization weights are given. First, we notice that a large number of accepted steps are necessary for convergence. We have 616 and 677 steps for the loading step. For the unloading step, we have 210 and 212 Proximal Newton steps. The values of the regularization weights are comparable for both materials. For the loading test we have a long phase of a the regularization weight between 10^{-2} and 10^{2} for about 600 Proximal Newton steps. Hence, we have a long globalization phase. From the convergence results of the Proximal Newton method we suppose that eventually the local convergence phase starts. This happens after about 650 steps where the regularization weight quickly drops to zero. For the unloading test similar results are observed. There, the local convergence phase appears to start already after 200 Proximal Newton steps.

A possible explanation for the high number of Proximal Newton steps in the globalization phase is found in the complexity of the geometry. For an energetic decent direction $\Delta x(\omega)$ of the second-order model problems (10.3), no large step size is possible. Large step sizes lead to unfavorable configurations, especially in the deformation field. To find a minimizing solution path in the given complicated geometry of the paperclip, only small linear steps $\Delta x(\omega)$ are allowed. These small steps can be seen in Figure 11.10.

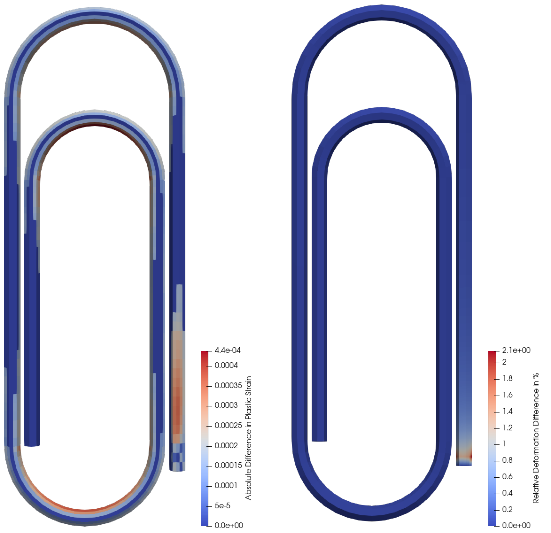

4.4e-04
0.0004
0.00035
0.0003
0.00025
0.0002
0.00015
0.0001
5e-5
0.0e+00
Absolute Difference in Plastic Strain
2.1e+00
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0.0e+00
Relative Deformation Difference in %

Figure 11.8.: Absolute difference in the plastic strain field (left) and relative difference in the deformation field (right) between the St. Venant–Kirchhoff and Mooney–Rivlin materials.

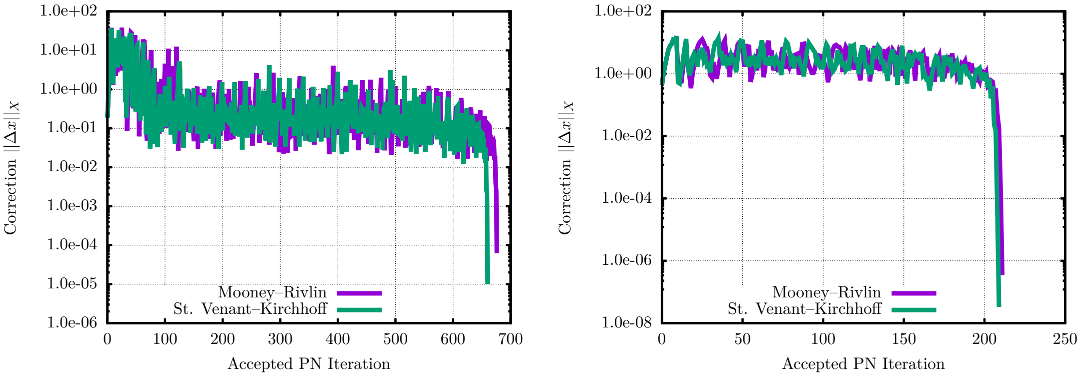


Figure 11.10.: Computed corrections of the Proximal Newton steps for loading (left) and unloading (right).

For the loading test, larger increments with a norm of up to 50 are admissible in the first accepted Proximal Newton steps. In this phase, a rough approximation of the result is already formed. Then, starting from step 25, the long globalization phase begins, in which the norm of the corrections is consistently varying around the value 0.3. In this phase the local optimizations in the deformation field are performed. Finally, in the local convergence phase, the corrections drop to zero.

Similar results are observed for the unloading test. In this case, only the initial phase with large corrections is missing. This is due to the fact that we have the large deformations left in the result as a consequence of the plastic strain.

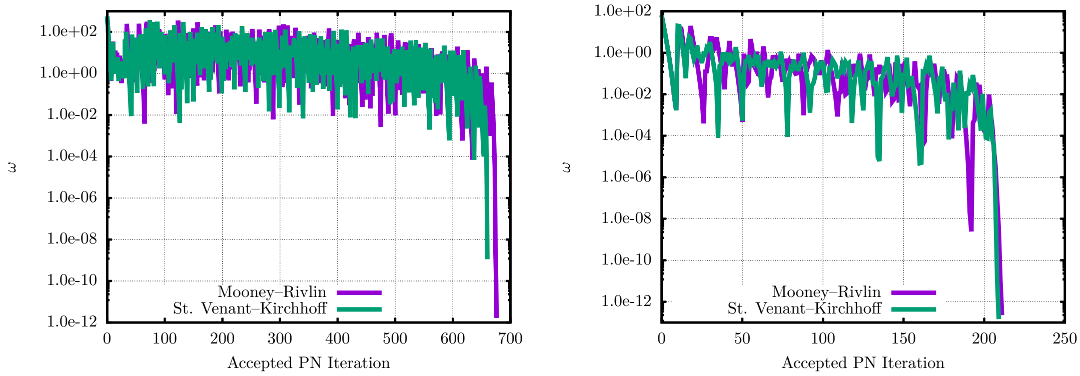


Figure 11.9.: Regularization weights ω for the loading step (left) and the unloading step (right).

To highlight the first phase of large corrections, the resulting energies (value of the objective functional) after each accepted Proximal Newton step are displayed in Figure 11.11. We can see that for the loading test most of the energy reduction is already done after the first few steps. The rest is just local optimization. For the unloading test, this first phase is not visible. Here we have a decaying decrease of the energy throughout all iterations.

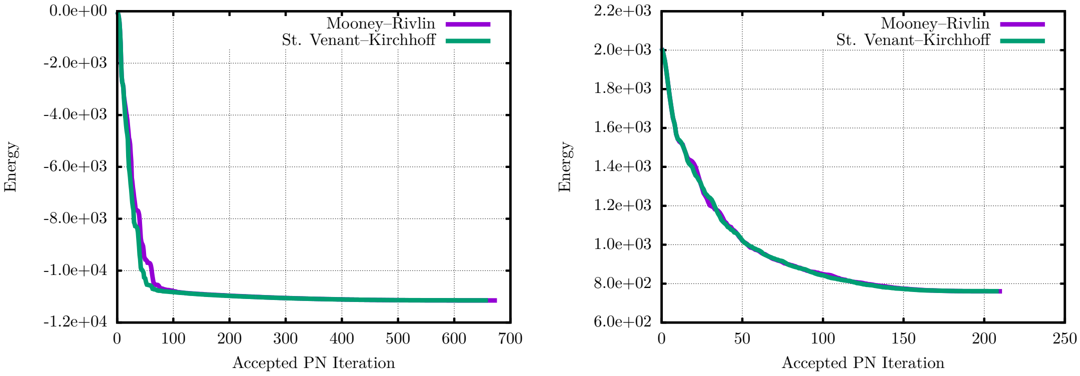


Figure 11.11.: Energies at the Proximal Newton steps for loading (left) and unloading (right).

11.4. Comparison to other Methods

The search for suitable comparison methods turns out to be more complicated than expected. Few algorithms exists that can handle minimization problems which are nonsmooth and nonconvex at the same time. In general two families of methods can be found: Proximal Gradient methods and subgradient methods. Both methods are briefly introduced in the following sections. However, we are not able to give suitable numerical results since the efficiency of these methods is pretty bad compared to the Proximal Newton method. Convergence of these methods in a suitable time is only given for very coarse grids and moderate loads. In contrast, the Proximal Newton method is able to solve those easy problem by a few iteration steps in under one second.

Nevertheless, the bottleneck of our Proximal Newton method in terms of wall time is the computation of the Hessian matrix f'' of the smooth part of the functional. If we have to rely on automatic differentiation tools, such as ADOL-C, we either have to deal with it or consider one of the alternative methods below. The advantage of both the Proximal Gradient method and the subgradient method is that we do not need a second-order derivative information, hence no Hessian matrix. Even if the convergence is slow and very limited, these methods could for instance be used as a preconditioner of the minimization problem to find a suitable initial value for the Proximal Newton method.

11.4.1. Proximal Gradient Method

The family of methods is closely related to the Proximal Newton method. It is applicable to composite minimization problems

$$\text{Minimize}_{x\in X}\ F(x) := f(x) + g(x) \tag{11.4}$$

on a space X with a differentiable part f and a convex part g, such that the objective functional F may be nonsmooth and nonconvex. The precise meaning of *differentiable* and *convex* may vary in literature. For the sake of comparability, we consider the classical definitions of these properties. This forms a similar setting to the Proximal Newton method, as given in Section 10.1.

The Proximal Gradient method is an iterative method that reduces the global minimization problem (11.4) into a series of first-order model problems

$$\lambda_{x,\omega}(\delta x) := f'(x)\delta x + \frac{\omega}{2}\left\|\delta x\right\|_X^2 + g(x+\delta x) - g(x), \tag{11.5}$$

closely related to the second-order model problems (10.3) of the Proximal Newton method without the Hessian matrix term. For a sufficiently large value of the regularization weight ω there is a solution

$$\Delta x(\omega) := \arg\min_{\delta x\in X} \lambda_{x,\omega}(\delta x).$$

A large disadvantage of the first-order model problems above is that the minimization of these problems is as complex as minimizing a second-order model problem of the Proximal Newton method. This makes each Proximal Gradient steps as slow as a Proximal Newton step. For this method, alternative sufficient decrease criteria and convergence criteria have to be defined.

However, as already discussed above, we are not able to find a suitable combination of these criteria to create an efficient Proximal Gradient method which compares to the Proximal Newton method. A more in-depth introduction and discussion on this topic is found in the thesis of Bastian Pötzl [24]. Basically, the motivation for this method is to spare the time needed for the computation of the Hessian matrix f'' in the Proximal Newton method. However, the number of iteration steps of the Proximal Gradient method appears to be unbounded for large problems and high loads. Thus, convergence below 10^5 iteration steps is only possible for unrefined grids. But then the computation of the Hessian matrix is again fast and there is no need to avoid using the Proximal Newton method, which solves those small problems in only 5 iteration steps. In conclusion, the Proximal Gradient method seems not suitable for finite-strain plasticity problems.

11.4.2. Subgradient Method

Another, even simpler method to solve composite minimization problems of the form (11.4) is given by the subgradient method. As the name already indicates, we compute a series of subgradients of F in this method. A subgradient of F is always possible to find since f is differentiable, and g (as a convex function) has a subdifferential ∂g.

In each subgradient step, at the current iterate x, we compute a subgradient

$$\mu \in f'(x) + \partial g(x).$$

As long as the x is not a local minimizer of F, we are able to find a subgradient μ, which is a descend direction. Even the steepest descend direction can be found efficiently if the nonsmooth part g contains some structure. In the case of small-strain plasticity, we have

$$g(x) = \sum_{i=1}^{n} \gamma_i \left\| x_i \right\|_2 ,$$

cf (9.18). The subdifferential is directly given by

$$\partial g(x) = \sum_{i=1}^{n} \gamma_i \partial \left\| \cdot \right\|_2 (x_i),$$

with

$$\partial \left\| \cdot \right\|_2 (x_i) = \begin{cases} \left\{ \frac{x_i}{\|x_i\|_2} \right\} & x_i \neq 0, \\ \{\mu : \ \|\mu\|_2 \leq 1\} & x_i = 0. \end{cases}$$

This allows to find the steepest descend direction μ. When μ is found, we set

$$x \leftarrow x + \alpha^k \mu$$

with a step size $\alpha^k \in (0, 1]$. There are different step size strategies in literature. Basically, one approach is to have a constant $\alpha^k := \alpha$ for all steps, or to consider null sequences $\alpha^k \to 0$. A great overview is given in [12].

A beauty of this method is the simplicity. The only complicated part is to find the optimal descend direction in the subdifferential. There is no need of solving the model problems (11.5)&(10.3).

A convergence criterion for subgradient methods is given by a sufficiently small distance of the subdifferential to zero

$$\operatorname{dist}(\partial F(x), 0) \leq \varepsilon$$

for some $\varepsilon > 0$. However, in [12] it is shown that at least $\mathcal{O}(\varepsilon^{-2})$ subgradient steps are necessary for convergence. In practice, we were not able to find a simple enough finite-strain plasticity problem to get a convergence result.

In summary, the subgradient method is also not suitable to solve finite-strain plasticity problems. However, since this method is very easy to implement and each step is computed fast, it may be used to find initial values for the Proximal Newton method.

11.5. Conclusion

We are able to present a minimization problem for finite-strain plasticity problem with a convergence guarantee for all stated problems, as shown in Theorem 10.5.1. The general iteration scheme is simple since we can we reduce the finite-strain minimization problem into a series of quadratic minimization problems (10.3) which take the form of small-strain minimization problems (4.17). These problems are already discussed in detail in the first part of this thesis and solved with the efficient TNNMG method in Section 5.2. Therefore, no further solver structure is necessary to implement the Proximal Newton method.

Even though computing the second-order model problems is expensive, the competitor algorithms which can be found in literature (Section 11.4) are not suitable for our problems and result in very inefficient convergence behavior, or no convergence at all.

In order to speed up the solution of the second-order model problems, an inexact variant of the Proximal Newton method is presented in Chapter 10. The numerical test indicates that this improves the efficiency of the Proximal Newton method a lot, as seen in Section 11.2.3. Moreover, our method is able to solve simple problems as given in Section 11.2.1, and eventually converges even for complex geometries, as seen in Section 11.3.

Therefore, in conclusion, the Proximal Newton method is suitable for all kinds of finite-strain increment problems. The only wall time bottleneck is the computation of the Hessian matrices of the smooth parts of the objective functional. However, this is a general problem of all second-order solvers, and therefore not a particular problem of the Proximal Newton method.

A. Appendix

A.1. The Dissipation Function for the Tresca Yield Criterion

We want to show the Tresca case of Theorem 3.2.4 from Section 3.2. First, we consider only kinematic hardening and define the set

$$K := \{\sigma \in \mathbb{S}^d \ : \ \max_{i,j=1,\dots,d} |\sigma_i - \sigma_j| \leq \sigma_0\}$$

for the admissible stresses. Before we show Theorem 3.2.4, we show an intermediate result.

Lemma A.1.1. *For $p \in \mathbb{S}^d$ we have*

$$\sup_{\sigma \in K} \{p : \sigma\} = \sup_{\sigma_1,\dots,\sigma_d \in \mathbb{R}} \Big\{ \sum_{i=1}^{d} p_i \sigma_i \mid \max_{i,j=1,\dots,d} |\sigma_i - \sigma_j| \leq \sigma_0 \Big\},$$

where $p_1, \dots, p_d$ and $\sigma_1, \dots, \sigma_d$ denote the principal values of p and σ, respectively.

Proof. Let Q_1 and Q_2 be any two orthogonal matrices. Then

$$\sup_{\sigma \in K} \{Q_1^T \sigma Q_1 : Q_2^T p Q_2\} = \sup_{\sigma \in K} \{Q_2 Q_1^T \sigma Q_1 Q_2^T : p\} = \sup_{\sigma \in K} \{\widetilde{Q}^T \sigma \widetilde{Q} : p\}, \tag{A.1}$$

where we have written $\widetilde{Q} = Q_1 Q_2^T$. The definition of K depends only on the singular values of σ. Hence σ is in K if and only if $\widetilde{Q}^T \sigma \widetilde{Q}$ is. Therefore, the last expression is equal to

$$\sup_{\sigma \in K} \{\widetilde{Q}^T \sigma \widetilde{Q} : p\} = \sup_{\sigma \in K} \{\sigma : p\}. \tag{A.2}$$

Since this holds for all orthogonal matrices Q_1 and Q_2, we can pick Q_1 such as to diagonalize σ, and Q_2 such as to diagonalize p. Then (A.1) together with (A.2) is the assertion. □

Proof of Theorem 3.2.4. We start by assuming kinematic hardening only. Hence, we have no internal variable η. Moreover, the only relevant case is $d = 3$, since for $d = 2$ the Tresca yield criterion is equivalent to the von Mises yield criterion (Lemma 2.4.3). By Lemma A.1.1 we have

$$D(p) = \sup_{\sigma_1,\sigma_2,\sigma_3 \in \mathbb{R}} \Big\{ p_1\sigma_1 + p_2\sigma_2 + p_3\sigma_3 \ : \ \max_{i,j=1,\dots,3} |\sigma_i - \sigma_j| \leq \sigma_0 \Big\}.$$

We express σ in new variables

$$(\sigma_1, \sigma_2, \sigma_3) \mapsto \big(\tilde{\sigma}_1 = \sigma_1 - \sigma_3,\ \tilde{\sigma}_2 = \sigma_2 - \sigma_3,\ \tilde{\sigma}_3 = \sigma_3 - \sigma_3 = 0\big).$$

This does not change the scalar product, since

$$\begin{aligned}
\sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3 &= (\tilde{\sigma}_1 + \sigma_3) p_1 + (\tilde{\sigma}_2 + \sigma_3) p_2 + \sigma_3 p_3 \\
&= \tilde{\sigma}_1 p_1 + \tilde{\sigma}_2 p_2 + \sigma_3 (p_1 + p_2 + p_3) \\
&= \tilde{\sigma}_1 p_1 + \tilde{\sigma}_2 p_2 \qquad \text{because } p \text{ is trace-free,} \\
&= \tilde{\sigma}_1 p_1 + \tilde{\sigma}_2 p_2 + \tilde{\sigma}_3 p_3 \qquad \text{because } \tilde{\sigma}_3 = 0.
\end{aligned}$$

Therefore, we can write

$$D(p) = \sup_{\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3 \in \mathbb{R}} \Big\{ \tilde{\sigma}_1 p_1 + \tilde{\sigma}_2 p_2 + \tilde{\sigma}_3 p_3 \ :\ \max\{|\tilde{\sigma}_1|, |\tilde{\sigma}_2|, |\tilde{\sigma}_1 - \tilde{\sigma}_2|\} \le \sigma_0,\ \tilde{\sigma}_3 = 0 \Big\}.$$

The maximization problem is now two-dimensional, and the admissible set is the convex hull of the six points

$$\begin{aligned}
a_1 &= (\sigma_0, 0), & a_2 &= (\sigma_0, \sigma_0), & a_3 &= (0, \sigma_0), \\
a_4 &= (-\sigma_0, 0), & a_5 &= (-\sigma_0, -\sigma_0), & a_6 &= (0, -\sigma_0).
\end{aligned}$$

The supremum of a linear objective functional is obtained at one of the corners of this hexagon. Hence, we have

$$D(p) = \max \left\{ \langle a_1, (p_1, p_2) \rangle, \ldots \langle a_6, (p_1, p_2) \rangle \right\}.$$

Computing the six scalar products leads to

$$D(p) = \sigma_0 \max \left\{ p_1, p_1 + p_2, p_2, -p_1, -p_1 - p_2, -p_2 \right\}.$$

Now we revert the shift from before by using $p_1 + p_2 + p_3 = 0$, s.t.,

$$D(p) = \sigma_0 \max \left\{ p_1, p_2, p_3, -p_1, -p_2, -p_3 \right\} = \sigma_0 \left\| p \right\|_2 .$$

This is the claim for kinematic hardening only.

The case of additional isotropic hardening is covered similar to the proof of the von Mises yield criterion in Theorem 3.2.4.

□

A.2. Geometric Multigrid Method

For completeness, we give a quick overview of the geometric multigrid method we use for the linear correction steps of the TNNMG method from Section 5.2. Our intention of the multigrid method is to have a cheap way to get an inexact, but still qualitatively good, solution $\tilde{x}$ of the linear equation

$$Ax = b, \quad \text{such that,} \quad \tilde{x} \approx x = A^{-1} b. \tag{A.3}$$

In this context, the matrix $A \in \mathbb{R}^{n \times n}$ is the truncated Newton matrix of the linearization step of TNNMG. We assume that A is symmetric and positive definite, and therefore a unique solution x exists.

A.2.1. Gauss–Seidel Steps

Gauss–Seidel steps are another simple and cheap way to create approximations of solution of (A.3). However, we will use them in the multigrid method. Therefore, we give a short sketch of how they are performed.

Definition A.2.1 (Gauss–Seidel Step). *Let* $x^i \in \mathbb{R}^n$ *be an approximation of the solution* $x \in \mathbb{R}^n$. *A Gauss-Seidel step for* x^i *is given by the algorithm:*

- *Set* $x^{i+1} \leftarrow x^i$
- *For* $k = 1, \dots, n$ *update all components of* x^{i+1}*:*

$$x_k^{i+1} \leftarrow \frac{1}{A_{kk}} \left(b_k - \sum_{j=1, j\neq k}^{n} A_{kj} x_j^{i+1} \right)$$

Although the method is simple, there are strong properties that can be shown.

Theorem A.2.2 ([14, Theorem 4.4.18.]). *Let* $A \in \mathbb{R}^{n\times n}$ *symmetric and positive definite. Then a Gauss-Seidel step is well-defined, since* $A_{kk} > 0$. *Furthermore, a repeated application of Gauss–Seidel steps converges to the unique solution*

$$\lim_{i\to\infty} x^i = x = A^{-1}b$$

for every initial iterate $x^0 \in \mathbb{R}^n$.

However, the downside is a slow convergence rate. Therefore, we will only consider single Gauss–Seidel steps in the multigrid method. A deeper analysis of multigrid methods, e.g. [14], indicates that Gauss-Seidel methods only eliminate local errors, or, in a different wording, high-frequent error terms. In order to improve the convergence, and therefore the quality of each step, low frequent error quantities need proper handling, too. There the idea of multigrid methods originates. Error frequencies correspond to the grid fineness. Therefore, we will apply Gauss–Seidel steps to a family of grids with different mesh sizes.

A.2.2. Prolongation and Restriction Operators

Let the domain Ω be covered by a hierarchy of grids $\mathcal{T}_0, \mathcal{T}_1, \dots, \mathcal{T}_N$, such that $\mathcal{T}_i$ is a subgrid of $\mathcal{T}_{i+1}$ for all $i = 0, 1, \dots, N-1$. This grid *hierarchy* is usually given by grid refinements. Therefore, we call grids with a higher index *fine* and grid with a lower index *coarse*. On each grid grid $\mathcal{T}_i$ we have a finite element space V_i for the global functions of unknowns. Of course, we have

$$n_i \coloneqq \dim V_i \leq \dim V_{i+1} \eqqcolon n_{i+1}.$$

By interpolating the basis of V_i in the basis of V_{i+1} we can define a linear prolongation operator $\mathcal{P}_i : V_i \to V_{i+1}$. Functions in V_i are represented by a coefficient vector in $\mathbb{R}^{n_i}$.

Hence, we can describe $\mathcal{P}_i$ by a matrix $P_i \in \mathbb{R}^{n_{i+1} \times n_i}$ that maps the coefficients in V_i into the corresponding interpolated coefficients in V_{i+1}. On the other hand, the transpose of P_i is automatically a restriction operator from a fine grid to a coarse grid. For convenience, we denote the restriction matrices by $R_i := P_i^T$.

Note that the linear system (A.3) is given on the finest grid $\mathcal{T}_N$, hence x is a coefficient vector in the finite element space V_N. For a consistent notation, we denote $A_N := A$ and $b_N := b$. With these definitions at hand, we can transfer the linear problem (A.3) to the next, coarser, grid $\mathcal{T}_{N-1}$ by setting

$$A_{N-1} := R_N^T A_N R_N, \quad b_{N-1} := R_N^T b_N.$$

Note that information is lost on the coarse grid.

A.2.3. Multigrid Steps

To overcome the issue of information loss we actually never restrict the solution iterate x^i, but rather the residual $r^i := b - Ax^i$. The residual is an error-like quantity and we lose only the high frequent error terms. The key idea is to apply the Gauss–Seidel steps to all grids levels of the grid hierarchy and to restrict the remaining residuals until the coarsest grid $\mathcal{T}_0$ is finally reached. We can assume that $A_0 \in \mathbb{R}^{n_0 \times n_0}$ is a small matrix, and hence a linear system with A_0 can be efficiently solved directly. A complete multigrid step used for the computations of this thesis is given as follows.

Definition A.2.3 (Multigrid Step)**.** *Consider the linear system $Ax = b$ in the coefficient space of a finite element space on a grid hierarchy. Moreover, let $\tilde{x} \approx x$ be a known approximation of the solution of the linear system. If there is a coarser grid, let R, P be the restriction/prolongation matrices as described above. A multigrid step $\tilde{y} = \mathrm{MG}(A, b\tilde{x})$ is given by the recursive algorithm*

- *If A is the linear system on the coarsest grid level:*
  - ***Return** the direct solution $\tilde{y} = A^{-1}b$*
- *Else:*
  - *Perform a Gauss-Seidel step on $\tilde{x}$*
  - *Restrict the residual $r := R(b - A\tilde{x})$*
  - *Set the vector $e = 0$ as approximation of the residual $e \approx r$.*
  - *Call recursively a multigrid step $y := \mathrm{MG}(R^T AR, e, r)$*
  - ***Return** the prolonged correction $\tilde{y} := \tilde{x} + Py$*

Remarks A.2.4.

(i) The variant of the multigrid step is called V-cycle, *since it goes recursively "down" to the coarsest grid and comes back "up" to the finest grid.*

(ii) Convergence theory is complicated. A comprehensive investigation can be found in, e.g., [14].

A.3. The Matrix Group SL(d)

In this section a quick overview about the most important results of the matrix group SL(d) is given. The definition is simple with the canonical determinant $\det\colon \mathbb{R}^{d\times d} \to \mathbb{R}$ at hand.

Definition A.3.1. *Let $d \in \mathbb{N}$. Then the set*

$$\mathrm{SL}(d) := \{A \in \mathbb{R}^{d\times d} :\ \det(A) = 1\}$$

is called the special linear group.

At first, we verify that this set forms a group.

Lemma A.3.2. SL(d) *is a group with the canonical matrix product as the group operation.*

Proof. This follows directly from the property that each matrix with determinant one is regular. A neutral element is given by the unit matrix. ☐

Remark A.3.3. *Matrix groups are Lie groups.*

A.3.1. Geometry

Investigation of Lie groups is done in the research field of differential geometry and is completely unrelated to the works of this thesis. Nevertheless, in contrast to other Lie groups such as GL(d) (matrices with non-zero determinant) and SO(d) (orthogonal matrices), we can view SL(d) as the level set of the determinant function. This is a polynomial restriction on the matrix space. This is done by recognizing that the determinant of a matrix $A \in \mathbb{R}^{d\times d}$ is a polynomial of degree d in the components of A. Hence, the level $\det(A) = 1$ is of a smooth kind and creates a smooth manifold $\mathcal{M} = \mathrm{SL}(d)$ with $d^2 - 1$ dimensions.

Lemma A.3.4. *For each $A \in \mathrm{SL}(d)$ the normal of $\mathcal{M}$ at A is given by $\nabla\det(A)$. Moreover, we have $\nabla\det(A) = \mathrm{cof}(A) = \det(A)A^{-T} = A^{-T}$.*

Proof. This is a classical result from smooth level set based sets. The second statement is true since $\det(A) = 1$. ☐

This direct geometric approach allows us to define tangent space at $A \in \mathrm{SL}(d)$ by using the orthogonality to the normal. Therefore, we define

$$T_P\,\mathrm{SL}(d) := \{A \in \mathbb{R}^{d\times d},\ A : P^{-T} = 0\}. \tag{A.4}$$

This tangential space can be given even more directly by using the equality

$$A : (BC) = (B^T A) : C = (AC^T) : B$$

for arbitrary $d \times d$ matrices A, B, C. This inserted into (A.4) yields

$$T_P\,\mathrm{SL}(d) = \{A \in \mathbb{R}^{d\times d},\ (P^{-1}A) : I = 0\ \Leftrightarrow\ \mathrm{tr}(P^{-1}A) = 0\},$$

or, equally,

$$T_P \operatorname{SL}(d) = \{A \in \mathbb{R}^{d\times d},\ (AP^{-1}) : I = 0 \ \Leftrightarrow\ \operatorname{tr}(AP^{-1}) = 0\}.$$

This leads to

$$T_P \operatorname{SL}(d) = \{PA :\ \operatorname{tr}(A) = 0\} = \{AP :\ \operatorname{tr}(A) = 0\}$$

There are three direct consequences.

Lemma A.3.5.

(i) The tangential space of $\operatorname{SL}(d)$ *at the unit matrix* I *is given by the vector space of trace-free matrices. This is denoted by* $\operatorname{sl}(d) := T_I \operatorname{SL}(d)$.

(ii) For each $P \in \operatorname{SL}(d)$ *the tangential space* $T_P \operatorname{SL}(d)$ *is given by* $P \operatorname{sl}(d)$*, i.e., each tangential space can be expressed by* $\operatorname{sl}(d)$ *multiplied from the left by* P.

(iii) For each $P \in \operatorname{SL}(d)$ *the tangential space* $T_P \operatorname{SL}(d)$ *is given by* $\operatorname{sl}(d)P$*, i.e., each tangential space can be expressed by* $\operatorname{sl}(d)$ *multiplied from the right by* P.

A.3.2. Positive Definite, Symmetric Subset $\operatorname{SL}(d)^+_{\mathrm{sym}}$

In this section some properties of the symmetric subset

$$\operatorname{SL}(d)^+_{\mathrm{sym}} := \Big\{A \in \operatorname{SL}(d) \colon A \text{ is symmetric and positive definite}\Big\}$$

are presented. This does *not* form a matrix group, since multiplication does generally not preserve symmetry. Nevertheless, this set has some properties that we need in this thesis.

Lemma A.3.6. *Consider a differentiable path* $P \in \mathcal{C}^\infty([0,1], \operatorname{SL}(d)^+_{\mathrm{sym}})$*. For each* $s \in [0,1]$*, there is a symmetric, trace-free* $B(s) \in \mathbb{S}^d_0$*, such that*

$$\dot{P}(s) = P(s)^{\frac{1}{2}} B(s) P(s)^{\frac{1}{2}}.$$

Proof. We show the equivalent result that $P(s)^{-\frac{1}{2}} \dot{P}(s) P(s)^{-\frac{1}{2}}$ is symmetric and trace-free. Symmetry follows directly from the symmetry of $P(s)$, which induces symmetry on $\dot{P}(s)$. For the trace, we consider an eigenvalue decomposition

$$P(s) = U(s)\Sigma(s)U(s)^T$$

with an orthogonal curve $U(s)$ and a diagonal curve $\Sigma(s)$. It follows from [27] that around each s the curves U and Σ can be chosen such that they are differentiable in s. Since U takes values in the Lie-group SO(3), we have $\dot{U}(s) = U(s)A(s)$ for an anti-symmetric matrix $A(s)$. Furthermore, $\Sigma(s)$ takes values in SL(3), and hence $\dot{\Sigma}(s) = \Sigma(s)D(s)$ for a trace-free diagonal $D(s)$. This yields

$$\begin{aligned} \dot{P}(s) &= \dot{U}(s)\Sigma(s)U(s)^T + U(s)\dot{\Sigma}(s)U(s)^T + U(s)\Sigma(s)\dot{U}(s)^T \\ &= U(s)\left(A(s)\Sigma(s) + \Sigma(s)D(s) + \Sigma(s)A(s)^T\right)U(s)^T. \end{aligned}$$

Combined with $P(s)^{-\frac{1}{2}} = U(s)\Sigma(s)^{-\frac{1}{2}}U(s)^T$, we have

$$P(s)^{-\frac{1}{2}}\dot{P}P(s)^{-\frac{1}{2}} = U(s)\left(\Sigma(s)^{-\frac{1}{2}}A(s)\Sigma(s)^{\frac{1}{2}} + D(s) + \Sigma(s)^{\frac{1}{2}}A(s)^T\Sigma(s)^{-\frac{1}{2}}\right)U(s)^T.$$

Since $D(s)$ is trace-free, $A(s)$ has a zero diagonal, the trace of the term in brackets is zero. Moreover, the trace is invariant under orthogonal transformations, thus the claim is shown. ☐

A.3.3. Projection onto $\mathrm{SL}(d)$

We quickly demonstrate how we can project a matrix $A \in \mathbb{R}^{d\times d}$ orthogonally (in the Frobenius inner product) onto the group $\mathrm{SL}(d)$. This is surprisingly difficult. We use an iterative method that yields suitable results in practice.

The algorithm is inspired by [9] and starts with a quasiprojection step in the direction $\operatorname{cof}(A)$. Note that the extrinsic derivative $\nabla\det(\cdot) = \operatorname{cof}(\cdot)$ is Frobenius-orthogonal to the polynomial restriction $\det(\cdot) = 1$ of the manifold $\mathrm{SL}(d)$. Therefore, the initial iterate is given by

$$\tilde{P}^0 := A + \gamma_0 \operatorname{cof}(A) = A + \gamma_0 \det(A)A^{-T}\,,$$

with $\gamma_0 \in \mathbb{R}$ such that $\det(\tilde{P}^0) = 1$, and hence $\tilde{P}^0 \in \mathrm{SL}(d)$. From there, we compute a sequence $(\tilde{P}^i)_{i\in\mathbb{N}}$ in $\mathrm{SL}(d)$ with the following procedure:

For each $\tilde{P}^i \in \mathrm{SL}(d)$ let $\tilde{Q}^i$ the closest point to A in the tangential space $T_{\tilde{P}^i}\,\mathrm{SL}(d)$. Then, the next quasiprojection is done in direction $\tilde{Q}^i - A$. For the next iterate $\tilde{P}^{i+1}$ find a $\gamma^{i+1} \in \mathbb{R}$, such that

$$\tilde{P}^{i+1} := A + \gamma^{i+1}(\tilde{Q}^i - A)$$

is again in $\mathrm{SL}(d)$. Although we have no convergence proof at hand, we observe similar convergence properties as those in [9]. In practice, the algorithm stopped after at most four iterations in our simulations with the criterion $\left\|\tilde{P}^{i+1} - \tilde{P}^i\right\|_F \le 10^{-8}$. For an efficient implementation, note that it is not necessary to compute the points $\tilde{Q}^i$ explicitly. Since the direction $\tilde{Q}^i - A$ is orthogonal to $T_{\tilde{P}^i}\,\mathrm{SL}(d)$, we can directly use the direction $\operatorname{cof}(\tilde{P}^i)$ and compute the next iterate by

$$\tilde{P}^{i+1} := A + \gamma^{i+1}\operatorname{cof}(\tilde{P}^i)\,.$$

The γ^i can be computed directly by solving a cubic polynomial equation, and choosing the solution γ^i with the smallest absolute value.

A.4. A Property of Symmetric Matrices

Lemma A.4.1. *Let $A, B \in \mathbb{R}^{d\times d}$ symmetric matrices with a regular B. Then the following holds for similarity transformations with B:*

$$\|A\|_F = \|\operatorname{sym}(A)\|_F = \|\operatorname{sym}(BAB^{-1})\|_F.$$

Proof. For the term on the right, we have by definition

$$\begin{aligned}\|\mathrm{sym}(BAB^{-1})\|_F^2 &= \frac{1}{4}\|BAB^{-1} + B^{-1}AB\|_F^2 \\ &= \frac{1}{2}\|BAB^{-1}\|_F^2 + \frac{1}{2}(BAB^{-1}) : (B^{-1}AB).\end{aligned}$$

Since the Frobenius inner product is invariant under adjoint transformations, we have

$$(BAB^{-1}) : (B^{-1}AB) = (BAB^{-1}B) : (B^{-1}A) = (AB^{-1}B) : (BB^{-1}A) = A : A = \|A\|_F^2$$

for a symmetric B. Moreover, the Frobenius norm is invariant under similarity transformations with a symmetric matrix, since the absolute value of the eigenvalues is equal to the singular values in this case. Hence, we have shown the claim. ☐

A.5. Robust and Efficient Computation of Matrix Exponential

The practical computation of the matrix exponential is not trivial [23]. Therefore, we will demonstrate how we perform this in the implementation of the numerics of this thesis.

The canonical method to compute the matrix exponential is given by computing

$$\exp(M) = \sum_{k=0}^{\infty} \frac{M^k}{k!}$$

up to numerical "convergence". In general, this method is unstable for matrices with large norm $\|M\| \gg 1$, since a large number of summands are needed until $\frac{\|M^k\|}{k!}$ falls below a certain tolerance. This causes cancellation errors. We overcome this issue by the so-called "scaling and squaring" method: We use that

$$\exp\left(\frac{M}{t}\right)^t = \exp(M)$$

holds for all $t \in \mathbb{N}$. We choose the smallest $s \in \mathbb{N}$ s.t.,

$$\frac{\|M\|}{2^s} \leq 1.$$

Next, we set $t = 2^s$ and compute $\exp\left(\frac{M}{t}\right)$ with the series method. In the end, we square the result s times and then we get $\exp(M)$.

This method works well in practice. It is also in compliance with [23] that this is generally the most robust and most efficient method for arbitrary matrices.

A.6. Convergence of the Newton Method with Armijo Damping

Consider the following minimization problem with smooth, strictly convex and coercive functional

$$L : \mathbb{R}^n \to \mathbb{R}$$

with a unique minimizer $x^* \in \mathbb{R}^n$. Let the Hesse matrix $\nabla^2 L$ elliptic, i.e.,

$$v^T \nabla^2 L(x) v \geq \varepsilon \|x\|_2^2$$

with a constant $\varepsilon > 0$ for all $x \in \mathbb{R}$. Starting from an initial value $x^* \in \mathbb{R}^n$ we define a sequence of iterates by the Newton method with Armijo damping. For a current iterate x^k we define the Newton direction d^k by

$$d^k := -\nabla^2 L(x^k)^{-1} \nabla L(x^k).$$

Then, for a fixed parameter $0 < \rho < 1$ we determine the smallest $\alpha^k \in \{2^{-l},\ l \in \mathbb{N}_0\}$, such that,

$$L(x^k + \alpha^k d^k) < L(x^k) + \rho \alpha^k \nabla L(x^k)^T d^k. \tag{A.5}$$

The next iterate is then given by

$$x^{k+1} := x^k + \alpha^k d^k.$$

Lemma A.6.1. *If $d^k \neq 0$, then there is a suitable $\alpha^k > 0$.*

Proof. Since $\nabla^2 L(x^k)$ is elliptic and $d^k \neq 0$, we always have

$$\nabla L(x^k)^T d^k = -\nabla L(x^k)^T \nabla^2 L(x^k) \nabla L(x^k) < 0.$$

If we rearrange (A.5) we get

$$\frac{L(x^k + \alpha^k d^k) - L(x^k)}{\alpha} \leq \rho \nabla L(x^k)^T d^k.$$

If α^k approaches zero the left hand side approaches $\nabla L(x^k)[d^k] = \nabla L(x^k)^T d^k$, which is not zero. Hence, we have to pass below the smaller portion $\rho \nabla L(x^k)^T d^k$ on the right hand side for some $\alpha^k > 0$. ☐

Remark A.6.2. *A consequence of the result above is that $\nabla L(x^k)^T d^k$ converges to zero if α^k converges to zero.*

According to the lemma above, we produce a descending sequence $L(x^k)$ that is bounded from below by $L(x^*)$. Thus, the sequence $L(x^k)$ converges. Now we have to show that we also have convergence $x^k \to x^*$.

Lemma A.6.3. *The Newton sequence (x^k), as described above, converges to x^* in $\mathbb{R}^n$.*

Proof. Let us assume that we are *not* converging to x^k, which implies from ellipticity of $\nabla^2 L(x^k)$ that both d^k and $\nabla L(x^k)$ do not converge to zero, as this is would be a sufficient criterion for the minimization. From convergence of $L(x^k)$ we conclude that $L(x^{k+1}) - L(x^k) \to 0$. Using the definition of x^{k+1} we get

$$L(x^k + \alpha^k d^k) - L(x^k) \to 0,$$

which implies

$$\rho \alpha^k \nabla L(x^k) d^k \to 0.$$

Since ρ is a positive constant, and $\nabla L(x^k)^T d^k$ does not converge to zero, the only possibility is that $\alpha^k \to 0$ as $k \to \infty$. On the other hand, by Remark A.6.2, we conclude that $\nabla L(x^k)^T d^k \to 0$, which is a contraction to the assumption. ☐

B. C++ Source Code

How to get the code and how to use it is described in detail in the following sections. Everything is implemented using the free and open source C++ library DUNE.[1] The research module with the specific implementations for the plasticity problems of this thesis can be found in the DUNE module `dune-plasticity`.[2] This module is based on previous works of Oliver Sander[3] on small-strain plasticity problems. Note that both modules are only visible internally for TU Dresden members. Nevertheless, a publicly available snapshot is available.[4]

The new implementations, especially for the finite-strain problems, are developed by the author of this thesis.

The Proximal Newton solver was developed in collaboration with Bastian Pötzl and is part of the DUNE module `dune-solvers`.[5] The Proximal Newton solver code is placed into a general solvers module available to the open audience. The reason is that Proximal Newton is a black box solver not specifically dedicated to plasticity problems.

B.1. Docker File

The easiest way to use the code from `dune-plasticity` is to create a container image from a Containerfile with `docker` or `podman`. To get the Containerfile you can either

- download it from
  `https://gitlab.mn.tu-dresden.de/jaap/dune-plasticity-snapshot/-/raw/main/Containerfile`
- scan the QR code below,
- or copy the contents from Figure B.1.

---

[1] `https://www.dune-project.org/`

[2] `https://gitlab.mn.tu-dresden.de/jaap/dune-plasticity`

[3] `https://gitlab.mn.tu-dresden.de/osander/dune-plasticity`

[4] `https://gitlab.mn.tu-dresden.de/jaap/dune-plasticity-snapshot`

[5] `https://git.imp.fu-berlin.de/agnumpde/dune-solvers/-/blob/master/dune/solvers/solvers/proximalnewtonsolver.hh` (this git repository will move to the TU Dresden in the future)

```
FROM docker.io/debian:bullseye

RUN \
export DEBIAN_FRONTEND=noninteractive; \
rm -f /etc/apt/apt.conf.d/docker-gzip-indexes \
&& rm -rf /var/lib/apt/lists/* \
&& apt-get update \
&& apt-get --no-install-recommends --yes dist-upgrade \
&& apt-get --no-install-recommends --yes install \
build-essential \
ca-certificates \
cmake \
gdb \
gfortran \
git \
libatlas-base-dev \
liblapack-dev \
libsuitesparse-dev \
libsuperlu-dev \
pkg-config \
python3-dev \
libadolc-dev \
nano \
vim \
&& apt-get clean && rm -rf /var/lib/apt/lists/*

WORKDIR /dune


# check out the snapshot of DUNE the time the thesis is published
RUN \
   git clone  https://gitlab.dune-project.org/core/dune-common.git           && cd dune-common         && git checkout 8eb66af05d0dc925b92f020fd7746b64e89d5a94 && cd .. \
&& git clone   https://gitlab.dune-project.org/core/dune-geometry.git        && cd dune-geometry       && git checkout 3b86089008aa6644c3da39aadd2b86e0425e24a4 && cd .. \
&& git clone   https://gitlab.dune-project.org/core/dune-grid.git            && cd dune-grid           && git checkout a0dd5fdb766e224c3212aeb70f6f84aeeb5813de && cd .. \
&& git clone   https://gitlab.dune-project.org/core/dune-istl.git            && cd dune-istl           && git checkout d65d7acc230250019770db62e6cb656e62bf2e6a && cd .. \
&& git clone   https://gitlab.dune-project.org/core/dune-localfunctions.git  && cd dune-localfunctions && git checkout 5f0256865c6fdd8cdfd1d0527ffa553a2a0ac5f2 && cd .. \
&& git clone   https://gitlab.dune-project.org/staging/dune-functions.git    && cd dune-functions      && git checkout 20d59c97bc0e19a5753ff38474752880d6eac0fd && cd .. \
&& git clone   https://gitlab.dune-project.org/staging/dune-typetree.git     && cd dune-typetree       && git checkout d3345d13804d9b6a26973272d799ca345abfd49c && cd .. \
&& git clone   https://gitlab.mn.tu-dresden.de/jaap/dune-elasticity.git      && cd dune-elasticity     && git checkout dbc113136412a119a0da30886b031196379c0397 && cd .. \
&& git clone   https://gitlab.dune-project.org/fufem/dune-fufem.git          && cd dune-fufem          && git checkout fc762036bf8d6003811a3251518e5a820786d107 && cd .. \
&& git clone   https://git.imp.fu-berlin.de/agnumpde/dune-solvers.git        && cd dune-solvers        && git checkout 4d821d6180d48e58d3f7215bbd64078b1b1874fa && cd .. \
&& git clone   https://gitlab.dune-project.org/fufem/dune-matrix-vector.git  && cd dune-matrix-vector  && git checkout c073d2a0d8feb3de23a48668219895ffcfb7e53c && cd .. \
&& git clone   https://gitlab.mn.tu-dresden.de/iwr/dune-gmsh4.git            && cd dune-gmsh4          && git checkout 180f26425adff80252bde351484f274a284c2fbc && cd .. \
&& git clone   https://gitlab.dune-project.org/staging/dune-uggrid.git       && cd dune-uggrid         && git checkout d0762d9b1c56048ab738eba80dbdd65604b26708 && cd .. \
&& git clone   https://git.imp.fu-berlin.de/agnumpde/dune-tnnmg.git          && cd dune-tnnmg          && git checkout 3bba09523b10238dd6f59d614616f15efd912897 && cd ..

# get the snapshot of the diss code
RUN git clone   https://gitlab.mn.tu-dresden.de/jaap/dune-plasticity-snapshot.git  dune-plasticity


RUN echo "CMAKE_FLAGS=\"-DCMAKE_CXX_FLAGS='-O3 -DNDEBUG'\"" > dune.opts

RUN /dune/dune-common/bin/dunecontrol --opts=dune.opts all

ENTRYPOINT ["/bin/bash", "--login", "--rcfile", "-i"]

LABEL run podman run --interactive --tty
```

Figure B.1.: Contents of the Containerfile

This image is based on Debian Bullseye and contains all necessary DUNE modules and external dependencies to run all the numerical tests of this thesis. It checks the snapshots of all DUNE modules at the time of this writing. Once the Containerfile is downloaded, run

```
podman image build -t dune-plasticity .
```

to create the image `dune-plasticity`. You can also use `docker` instead of `podman`. If the image is build successfully, use

```
podman run -it dune-plasticity:latest
```

to start the image in the interactive mode. The executable code related to this thesis can be found in the directory `/dune/dune-plasticity/build-cmake/src`

B.2. Small-Strain Tests

If you wish to run small-strain tests, start the executable `primalplasticity-p1p0` with a parameter file. A canonical parameter file can be found in `/dune/dune-plasticity/src/dune-plasticity-parset`. Hence, for a quick test, run

```
./primalplasticity-p1p0 ../../src/dune-plasticity.parset
```

Feel free to use other parameters in the parameter set. The available grids can be found in `/dune/dune-plasticity/doc/grids`. You can edit the files with `vim` or `nano` directly in the image. If you want to change the grid dimension, you need to set this in

```
/dune/dune-plasticity/src/primalplasticity-p1p0.cc
```

in the `main` function. Afterwards, you need to recompile the code with `make` in the directory `/dune/dune-plasticity/build-cmake/src`. The tests of Chapter 6 can be found in `dune/dune-plasticity/benchmarks`.

B.3. Finite-Strain Tests

The finite-strain code is completely different from the small-strain code. The executable is `finite-strain-plasticity` in `/dune/dune-plasticity/build-cmake/src`. A canonical parameter file can be found in `/dune/dune-plasticity/problems/finite-strain-plasticity.parset`. Hence, you can directly call

```
./finite-strain-plasticity ../../problems/finite-strain-plasticity.parset
```

in the directory `/dune/dune-plasticity/build-cmake/src`. Feel free to discover the different parameter choices. Remember to recompile the code if you change a compile-time parameter in the file

```
/dune/dune-plasticity/src/finite-strain-plasticity.cc
```

by calling `make` in the directory `/dune/dune-plasticity/build-cmake/src`. The benchmarks of Chapter 11 can be found in `/dune/dune-plasticity/src/benchmarks-finite-strain`.

Bibliography

[1] H. H. Bauschke and P. L. Combettes. “Lower Semicontinuous Convex Functions”. In: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Cham: Springer International Publishing, 2017, pp. 157–176. DOI: `10.1007/978-3-319-48311-5_9`.

[2] C. Carstensen. “Domain decomposition for a non-smooth convex minimization problem and its application to plasticity”. In: *Numerical linear algebra with applications* 4.3 (1997), pp. 177–190.

[3] Y. Chen, T. Davis, W. Hager, and S. Rajamanickam. “Algorithm 887: CHOLMOD, Supernodal Sparse Cholesky Factorization and Update/Downdate.” In: *ACM Trans. Math. Softw.* 35 (Jan. 2008).

[4] P. G. Ciarlet. *Mathematical Elasticity: Three-Dimensional Elasticity.* Society for Industrial and Applied Mathematics, Jan. 2021. DOI: `10.1137/1.9781611976786`.

[5] Y. F. Dafalias. “Plastic spin: necessity or redundancy?” In: *International Journal of Plasticity* 14.9 (1998), pp. 909–931. ISSN: 0749-6419. DOI: `https://doi.org/10.1016/S0749-6419(98)00036-9`.

[6] P. Deuflhard. *Numerische Mathematik 2 : gewöhnliche Differentialgleichungen.* Berlin Boston: Walter de Gruyter GmbH & Co., KG, 2013.

[7] H. Dörrie. *Kubische und biquadratische Gleichungen.* De Gruyter, Dec. 1948. DOI: `10.1515/9783486775990`.

[8] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems.* Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 1999.

[9] V. Elser. “Matrix product constraints by projection methods”. In: *Journal of Global Optimization* 68.2 (2017), pp. 329–355.

[10] C. Gräser, D. Kienle, and O. Sander. “Truncated nonsmooth newton multigrid for phase-field brittle-fracture problems”. In: *arXiv preprint arXiv:2007.12290* (2020).

[11] C. Gräser and O. Sander. “Truncated Nonsmooth Newton Multigrid Methods for Block-Separable Minimization Problems”. In: (2017). URL: `https://arxiv.org/abs/1709.04992`.

[12] S. Gratton, E. Simon, and P. L. Toint. “An algorithm for the minimization of nonsmooth nonconvex functions using inexact evaluations and its worst-case complexity”. In: *Mathematical Programming* 187.1-2 (Jan. 2020), pp. 1–24. DOI: `10.1007/s10107-020-01466-5`.

[13] A. Griewank and A. Walther. *Evaluating Derivatives.* Second. Society for Industrial and Applied Mathematics, 2008. DOI: `10.1137/1.9780898717761`.

[14] W. Hackbusch. *Iterative Solution of Large Sparse Systems of Equations.* Springer New York, 1994. DOI: `10.1007/978-1-4612-4288-8`.

[15] K. Hackl. “Generalized standard media and variational principles in classical and finite strain elastoplasticity”. In: *Journal of the Mechanics and Physics of Solids* 45.5 (May 1997), pp. 667–688. DOI: `10.1016/s0022-5096(96)00110-x`.

[16] W. Han and B. D. Reddy. *Plasticity: mathematical theory and numerical analysis.* Vol. 9. Springer Science & Business Media, 2012.

[17] H. Hardering and O. Sander. “Geometric Finite Elements”. In: *Handbook of Variational Methods for Nonlinear Geometric Data.* Ed. by P. Grohs, M. Holler, and A. Weinmann. Cham: Springer International Publishing, 2020, pp. 3–49. DOI: `10.1007/978-3-030-31351-7_1`.

[18] N. J. Higham. *Functions of matrices: theory and computation.* SIAM, 2008.

[19] P. Jaap, B. Pötzl, and O. Sander. “in prep.: An Inexact Proximal Newton Method for Rate-Independent Finite-Strain Plasticity Problems”. In: ().

[20] E. Lee and D. Liu. “Finite-strain elastic—plastic theory with application to plane-wave analysis”. In: *Journal of applied physics* 38.1 (1967), pp. 19–27.

[21] A. Mielke. “Finite elastoplasticity Lie groups and geodesics on SL (d)”. In: *Geometry, mechanics, and dynamics.* Springer, 2002, pp. 61–90.

[22] A. Mielke and T. Roubíček. *Rate-Independent Systems.* Springer New York, 2015. DOI: `10.1007/978-1-4939-2706-7`.

[23] C. Moler and C. Van Loan. “Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later”. In: *SIAM Review* 45.1 (2003), pp. 3–49. DOI: `10.1137/S00361445024180`.

[24] B. Pötzl. *in prep.: Inexact Proximal Newton Methods for Finite-strain Plasticity (dissertation).*

[25] B. Pötzl, A. Schiela, and P. Jaap. *Inexact Proximal Newton methods in Hilbert spaces.* 2022. DOI: `10.48550/ARXIV.2204.12168`.

[26] B. Pötzl, A. Schiela, and P. Jaap. “Second order semi-smooth Proximal Newton methods in Hilbert spaces”. In: *Computational Optimization and Applications* (Apr. 2022). DOI: `10.1007/s10589-022-00369-9`.

[27] F. Rellich. “Störungstheorie der Spektralzerlegung. IV”. In: *Mathematische Annalen* 117.1 (1940), pp. 356–382.

[28] L. Rencontre, W. Bird, and J. Martin. “Internal variable formulation of a backward difference corrector algorithm for piecewise linear yield surfaces”. In: *Meccanica* 27.1 (1992), pp. 13–24.

[29] O. Sander. “Solving primal plasticity increment problems in the time of a single predictor-corrector iteration”. In: (2017). URL: `https://arxiv.org/abs/1707.03733`.

[30] O. Sander and P. Jaap. “Solving primal plasticity increment problems in the time of a single predictor–corrector iteration”. In: *Computational Mechanics* 65.3 (Oct. 2019), pp. 663–685. DOI: `10.1007/s00466-019-01788-y`.

[31] J. C. Simo and T. J. Hughes. *Computational inelasticity*. Vol. 7. Springer Science & Business Media, 2006.

[32] E. Stein, E. Ramm, E. Rank, R. Rannacher, K. Schweizerhof, E. Stein, W. Wendland, G. Wittum, P. Wriggers, and W. Wunderlich. *Error-controlled Adaptive Finite Elements in Solid Mechanics*. John Wiley & Sons, 2003.

[33] M. Weiser, P. Deuflhard, and B. Erdmann. “Affine conjugate adaptive Newton methods for nonlinear elastomechanics”. In: *Optimization Methods and Software* 22.3 (June 2007), pp. 413–431. DOI: `10.1080/10556780600605129`.

Danksagung

Auf meinem Weg bis zu dieser Dissertation haben mich viele Menschen begleitet, ohne die ich diese Arbeit nicht hätte erstellen können. Mir ist klar, dass man nie alle auflisten kann, die mir in den letzten Jahren Unterstützung geleistet haben. Ich bedanke mich insbesondere bei:

Meinem Betreuer Oliver Sander, der mir während der gesamten Promotion stets mit Rat und Tat zur Seite stand. Danke für die Einladung nach Dresden, das tolle Forschungsthema, die Begeisterung an freier Software und die unzähligen fachlichen Diskussionen beim Mittagessen.

Meinen Kolleginnen und Kollegen Ansgar, Lisa, Klaus, Robert, Hanne und Katja aus der Arbeitsgruppe. Ohne euch wäre ich nicht so oft nach Dresden gefahren. Vielen Dank für die Aufheiterungen im Büro, den schönen schwarzen Kaffee und viele tolle Restaurantbesuche!

Allen anderen Kolleginnen und Kollegen des Instituts. Danke für den stets herzlichen Umgang, die leckeren Kuchen und die vielen schönen Quizabende.

Meinem tollen Projektpartner Bastian Pötzl aus Bayreuth, der mir gezeigt hat, wie gut eine wissenschaftliche Zusammenarbeit aussehen kann. Ohne ihn wäre der 2. Teil der Dissertation niemals zu Stande gekommen. Danke für unsere hoffnungslose Abhängigkeit, für die unzähligen Onlinegespräche am frühsten Morgen, und die leckere Pizza in Bayreuth.

In Dresden danke ich auch dem Team der Mensen, da das Essen oft Grund genug war an die Uni zu kommen. Außerdem danke ich der DB Lounge am Dresdner Hauptbahnhof, die mich immer herzlich aufgenommen hat, wenn der Eurocity nach Berlin mal wieder nicht kam, wie er sollte.

Außerhalb der Universität danke ich meiner Familie sowie meinen Freundinnen und Freunden. Es tut gut, ein so schönes Umfeld zu haben und mit euch den Kopf frei zu bekommen. Wenn ich versucht habe die Dinge mit einfachen Worten zu erklären, sind mir aber dank euch viele Ungereimtheiten klar geworden.

Ein großer Dank geht auch an Township Rebellion für die tollen DJ-Sets, ohne die ich auf vielen Zugfahren und im Homeoffice die Krise bekommen hätte.

Zu guter Letzt danke ich natürlich meiner Partnerin Anne, die mich schon durch das ganze Mathematikstudium hindurch begleitet hat. Danke, dass du mir das Leben so schön machst. Danke, dass du mir einen Grund gibst, ständig zurück nach Berlin zu pendeln. Danke, dass du meine unbeholfenen Erklärungsversuche in Bezug auf die ganze Mathematik hinter dieser Arbeit ertragen hast. Und vielen Dank für das sich immer wieder auffüllende Süßigkeitenglas, das mir die Kraft für die letzten Meter der Dissertation gegeben hat.