Newton-Like Solver for Elastoplastic Problems with Hardening and its Local Super-Linear Convergence

Peter G. Gruber*, Jan Valdman*

Abstract

We discuss a new solution algorithm for quasi-static elastoplastic problems with hardening. Such problems are described by a time dependent variational inequality, where the displacement and the plastic strain fields serve as primal variables. After discretization in time, one variational inequality of the second kind is obtained per time step and can be reformulated as each one minimization problem with a convex energy functional, which depends smoothly on the displacement and non-smoothly on the plastic strain. There exists an explicit formula how to minimize the energy functional with respect to the plastic strain for a given displacement. Thus, by its substitution, an energy functional depending only on the displacement can be obtained. Our technique based on the well known theorem of Moreau from convex analysis shows that the energy functional is differentiable with an explicitely computable first derivative. The second derivative of the energy functional exists everywhere in the domain apart from the elastoplastic interface, which separates the deformed continuum in elastic and plastic parts. A Newton-like method exploiting slanting functions of the energy functional's first derivative is proposed and implemented numerically. The local super-linear convergence of the Newton-like method in the discrete case is shown and sufficient regularity assumptions are formulated to guarantee local super-linear convergence also in the continuous case.

1 Introduction

We consider a quasi-static initial-boundary value problem for small strain elastoplasticity with hardening. Throughout the paper, only the linear isotropic hardening is considered, however an extention to other kinds of linear hardening is straightforward. Several computation techniques for solving the elastoplastic problem with various kinds of hardening can

*Special Research Program SFB F013 'Numerical and Symbolic Scientific Computing', supported by the Austrian Science Fund 'Fonds zur Förderung der wissenschaftlichen Forschung (FWF)', at the Johannes Kepler University Linz, Altenbergerstrasse 69, A-4040 Linz; email: [peter.gruber, jan.valdman]@sfb013.uni-linz.ac.at

be found in [KL84, Bla97, SH98, ACZ00, KLV04, Kie06]. For the efficient solution of problems without hardening, i. e. perfect Prandtl Reuß plasticity, we refer to [Wie00, Wie06]. Combining the equilibrium of forces with the elastoplastic hardening law under the assumption of small deformations, we formulate a time-dependent variational inequality. The existence of a uniquene solution to such inequality has been for instance proved in [Joh76] utilizing results for general variation inequalities [DL76].

The traditional numerical methods for solving the time-dependent variational inequality were based on the explicit Euler time-discretization with respect to the loading history. In this case the idea of implicit return mapping discretization [SH98] turned out fruitful for calculations. By an implicit Euler time-discretization on the other side, the time-dependent inequality is approximated by a sequence of time-independent variational inequalities for the unknown displacement u and the plastic strain p. Each of these inequalities is equivalent [GLT81] to a minimization problem with the convex but non-smooth energy functional

$$\bar{J}(u, p) \to \min .$$

It has been already shown in [Car97] that a method of alternating minimization in the displacement and in the plastic strain convergences globally and linearly. The minimization in the plastic strain can be calculated locally using the explicitely known dependence [AC00] of the plastic strain on the total strain, i.e., $p = \tilde{p}(\varepsilon(u))$. Thus the equivalent energy minimization problem for the displacement u only

$$J(u) := \bar{J}(u, \tilde{p}(\varepsilon(u)) \to \min$$

can be defined. Since the dependencies of the energy functional on the plastic strain p, and of the minimizer $\tilde{p}$ on the total strain $\varepsilon(u)$ are continuous but non-smooth, the Fréchet derivate $D\,J(u)$ seems not to exist. Therefore, a Newton-like method introduced in [ACZ00] using damping theoretically converges globally but only linearly, the super-linear convergence is discussed but not proved there.

The main theoretical contribution of this paper is the extension of the analysis done in [ACZ00]. We show that the structure of the energy functional $J(u)$ satisfies the assumptions of Moreau's theorem from convex analysis and therefore the energy functional $J(u)$ is Fréchet differentiable (Proposition 1 on page 10) with the explicitely computable Fréchet derivative $D\,J(u)$. The second derivative of the energy functional $D^2 J(u)$ exists everywhere in the domain apart from the elastoplastic interface only, which separates the deformed continuum in elastically and plastically deformed parts.

Using the concept of slant differentiability we define a Newton-like method for which the super-linear convergence is investigated. The notion of slanting functions and slant differentiability was, recently, introduced by X. Chen, Z. Nashed and L. Qi in [CNQ01]. One of the key features of slanting functions is, that Newton-like methods utilizing slanting functions instead of classical Fréchet derivatives also converge locally super linearly. By detailed analysis we show that the second derivative of the elastoplastic energy functional $D^2 J(u)$ serves well as a slanting derivative $(D\,J)^o\,(u)$ and that its value on the elastoplastic interface can be chosen arbitrarily without effect on slant differentiability and on the super-linear convergence of the Newton-like method. This conclusion is easy to see in the spacial discrete case (e.g. after the FEM discretization) and it also provides an explanation to an open question of a rigorous proof of superlinear convergence formulated in Remark 7.5 of [ACZ00].

The continuous case is more complicated and requires some extra regularity assumptions for the trial stress in each Newton-like step. To the best knowledge of the authors, there are no theoretical results known, which would guarantee the required regularity properties. Already existing regularity results, e. g. such as in [FS00, BF02], concern the regularity of stress solution and displacement solution, but not of the trial stresses during the Newton-like iteration. Thus, this work may serve as a starting point for more advanced theoretical analysis on the regularity of elastoplastic problems.

Various numerical experiments conclude the paper. For the space-discretization, the finite element method of the lowest order with piece-wise linear nodal ansatz-functions for the displacement and the piece-wise constant plastic strain is used. The unknown discrete displacement u has to satisfy the necessary condition $D\,J(u) = 0$, which represents the system of nonlinear equations to solve. Three examples in two dimensions provide the following conclusions.

- The number of iteration steps is (almost) independent of the size of the discretization.
- The Newton-like method converges super-linearly, and even quadratically after the elastoplastic zones are identified sufficiently. This conclusion has also been explained theoretically for different types of Newton-like solvers [Bla97].

2 Mathematical Modeling

Let $\Theta := [0, T]$ be a time interval, and let Ω be a bounded Lipschitz domain in the space $\mathbb{R}^3$. The equilibrium of forces in the quasi-static case reads

$$-\operatorname{div}(\sigma(x,t)) = f(x,t) \quad \text{for} \;\; (x,t) \in \Omega \times \Theta\,, \tag{1}$$

where $\sigma(x,t) \in \mathbb{R}^{3\times 3}$ is called *Cauchy's stress tensor* and $f(x,t) \in \mathbb{R}^3$ is called the *volume force* acting at the material point $x \in \Omega$ at the time $t \in \Theta$. Let $u(x,t) \in \mathbb{R}^3$ be the *displacement* of the body, and let

$$\varepsilon(u) := \frac{1}{2}\left(\nabla u + (\nabla u)^T\right) \tag{2}$$

denotes the (linearized) *Green-St. Venant strain tensor*. In elastoplasticity, the strain ε is split additively into an elastic part e and a plastic part p, that is,

$$\varepsilon = e + p\,. \tag{3}$$

The stress-strain relation is given by Hook's law

$$\sigma = \mathbb{C}e, \tag{4}$$

where the fourth-order *elasticity tensor* $\mathbb{C} \in \mathbb{R}^{3\times3\times3\times3}$ is defined by $\mathbb{C}_{ijkl} := \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$. Here, $\lambda > 0$ and $\mu > 0$ denote the *Lamé* constants, and δ_{ij} is the *Kronecker-symbol*.

Let the boundary $\Gamma := \partial\Omega$ be split into a *Dirichlet-part* Γ_D and a *Neumann-part* Γ_N, which satisfy $\Gamma = \overline{\Gamma_D \cup \Gamma_N}$. We assume the boundary conditions

$$u = u_D \qquad \text{on } \Gamma_D\,, \tag{5}$$

$$\sigma\, n = g \qquad \text{on } \Gamma_N\,, \tag{6}$$

where $n(x,t)$ is the exterior unit normal, $u_D(x,t) \in \mathbb{R}^3$ denotes a prescribed displacement and $g(x,t) \in \mathbb{R}^3$ is a prescribed traction force. By neglecting the plastic term in (3), i.e. $p = 0$, the system (1) - (6) describes elastic behavior of the continuum Ω.

Another two properties incorporating the admissibility of a stress field σ with respect to a certain *hardening law* and the time evolution of the plastic strain p are required. Therefore, we introduce the *hardening parameter* α and call a tuple (σ,α) the *generalized stress*. Such generalized stress is called *admissible*, if for a given convex *yield functional* ϕ there holds

$$\phi(\sigma,\alpha) \le 0\,. \tag{7}$$

The explicit form of ϕ depends on the choice of the *hardening law*. In this paper we concentrate on the *isotropic* hardening law only, where the hardening parameter α is a scalar function $\alpha : \Omega \to \mathbb{R}$ and the yield functional ϕ is then defined by

$$\phi(\sigma,\alpha) := \begin{cases} \|\operatorname{dev}\sigma\|_F - \sigma_y(1 + H\alpha) & \text{if } \alpha \ge 0, \\ +\infty & \text{if } \alpha < 0. \end{cases} \tag{8}$$

Here, the matrix *Frobenius norm* $\|A\|_F := \langle A\,,\, A\rangle_F^{1/2}$ is defined via the matrix scalar product $\langle A\,,\, B\rangle_F := \sum_{ij} a_{ij}b_{ij}$ for $A = (a_{ij}) \in \mathbb{R}^{3\times3}$ and $B = (b_{ij}) \in \mathbb{R}^{3\times3}$ and the *deviator* is defined for square matrices as $\operatorname{dev} A = A - \frac{\operatorname{tr} A}{\operatorname{tr} I} I$, where the *trace* of a matrix is defined by $\operatorname{tr} A = \langle A\,,\, I\rangle_F$ with I denoting the identity matrix. The material constants σ_y and H are both positive real numbers and called *yield stress* and *modulus of hardening*, respectively. The second property addresses the time development of the *generalized plastic strain* $(p,-\alpha)$. There must hold the *normality condition*

$$\langle(\dot p, -\dot\alpha)\,,\,(\tau,\beta) - (\sigma,\alpha)\rangle_F \le 0 \qquad \forall(\tau,\beta) \text{ satisfying } \phi(\tau,\beta) \le 0\,, \tag{9}$$

where $\dot{p}$ and $\dot{\alpha}$ denote the first time derivatives of p and α. The initial conditions read

$$p(x,0) = p_0(x) \quad \text{and} \quad \alpha(x,0) = \alpha_0(x) \quad \forall x \in \Omega\,, \tag{10}$$

with given initial values $p_0 : \Omega \to \mathbb{R}^{3\times 3}_{\text{sym}}$ and $\alpha_0 : \Omega \to [0,\infty[$.

Problem 1 (classical formulation)**.** Find (u, p, α) such, that (1)–(7), (9) and (10) are satisfied.

Problem 1 is formal in the sense, that no function spaces have been specified so far. Convenient spaces turn out to be $V := \left[H^1(\Omega)\right]^3$ and $Q := [L_2(\Omega)]^{3\times 3}_{\text{sym}}$ with their associated scalar products and norms

$$\begin{aligned}
\langle u\,,\, v\rangle_V &:= \int_\Omega \left(\langle u\,,\, v\rangle_F + \langle \nabla u\,,\, \nabla v\rangle_F\right)\,\mathrm{d}x\,, & \|v\|_V &:= \langle v\,,\, v\rangle_V^{1/2}\,,\\
\langle p\,,\, q\rangle_Q &:= \int_\Omega \langle p\,,\, q\rangle_F\,\mathrm{d}x\,, & \|q\|_Q &:= \langle q\,,\, q\rangle_Q^{1/2}\,.
\end{aligned}$$

Further we denote $V_D := \{v \in V \mid v_{|\Gamma_D} = u_D\}$ and $V_0 := \{v \in V \mid v_{|\Gamma_D} = 0\}$.

Then it is possible to derive a time-dependent variational inequality for unknown displacement $u \in H^1(\Theta; V_D)$ and plastic strain $p \in H^1(\Theta; Q)$ from Problem 1, see [HR99] for details. It is know that such a time-dependent variational inequality has a unique solution. The numerical treatment requires a time discretization of the a time-dependent variational inequality. Therefore, let $N_\Theta \in \mathbb{N}$, $\tau := T/N_\Theta$ and $\Theta_\tau := \{t_k := k\tau \mid k \in \{0, \ldots, N_\Theta\}\}$ be a uniform discretization of the time interval $\Theta = [0, T]$. We introduce the notation

$$u_k := u(t_k)\ ,\quad p_k := p(t_k)\ ,\quad \alpha_k := \alpha(t_k)\ ,\quad f_k := f(t_k)\ ,\quad g_k := g(t_k)\ ,\quad \ldots\ ,$$

and approximate time derivatives by the backward difference quotients, that is,

$$\dot{p}_k \approx (p_k - p_{k-1})\,/\tau \quad \text{and} \quad \dot{\alpha}_k \approx (\alpha_k - \alpha_{k-1})\,/\tau.$$

Consequently, the time-dependent problem can be decomposed in a sequence of time independent variational inequalities of the second kind, each of which can be equivalently expressed by the minimization of a convex functional mapping to $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. The resulting time discretized minimization problem reads [Car97]:

Problem 2. Let $k \in \{1, \ldots, N_\Theta\}$ denote a given time step, $p_{k-1} \in Q$ and $\alpha_{k-1} \in L_2(\Omega)$ be given such, that $\alpha_{k-1} \geq 0$ almost everywhere. Define $\bar{J}_k : V \times Q \to \overline{\mathbb{R}}$ by $\bar{J}_k(v, q) := +\infty$ if $\operatorname{tr} q \neq \operatorname{tr} p_{k-1}$, else

$$\begin{aligned}
\bar{J}_k(v,q) := \frac{1}{2}\int_\Omega &\langle \mathbb{C}(\varepsilon(v) - q)\,,\, \varepsilon(v) - q\rangle_F + (\alpha_{k-1} + \sigma_y H\|q - p_{k-1}\|_F)^2\,\mathrm{d}x\\
&+ \int_\Omega \sigma_y \|q - p_{k-1}\|_F\,\mathrm{d}x - \int_\Omega f_k \cdot v\,\mathrm{d}x - \int_{\Gamma_N} g_k \cdot v\,\mathrm{d}s\,.
\end{aligned} \tag{11}$$

Find $(u_k, p_k) \in V_D \times Q$ such that $\bar{J}_k(u_k, p_k) \leq \bar{J}_k(v, q)$ holds for all $(v, q) \in V_D \times Q$.

Problem 2 represents a one time step problem. The convex functional $\bar{J}_k$ expresses the mechanical energy of the deformed system at the $k-$th time step. The goal is to find a displacement u_k and a plastic strain p_k such that the energy $\bar{J}_k$ is minimized. $\bar{J}_k$ is smooth with respect to the displacements v, but not with respect to the plastic strains q.

The hardening parameter $\alpha_k \in L_2(\Omega)$ does not appear in Problem 2 directly, but can be calculated analytically in dependence on the plastic strain by $\alpha_k = \tilde{\alpha}_k(p_k)$, where, in the case of isotropic hardening, $\tilde{\alpha}_k : Q \rightarrow L_2(\Omega)$ reads [Car97]

$$\tilde{\alpha}_k(q) = \alpha_{k-1} + \sigma_y H \|q - p_{k-1}\|_F \,. \tag{12}$$

3 Derivation of a Smooth Minimization Problem with Respect to the Displacement Only

Various strategies have been introduced to solve the minimization in Problem 2. C. Carstensen investigated a separated minimization in the displacement v and in the plastic strain q alternately and proved the linear convergence of the resulting method in [Car97]. Another interesting technique is to reduce Problem 2 to a minimization problem with respect to the displacements v only. We will make an important observation that such reduced minimization problem is smooth with respect to the displacements v and its derivative is explicitly computable. To discuss this issue, let us first introduce a more abstract formulation of (11). Therefore, we define the $\mathbb{C}$-scalarproduct, the $\mathbb{C}$-norm, a convex functional ψ_k and a linear functional l_k by

$$\langle q_1 \,, q_2 \rangle_{\mathbb{C}} \;:=\; \int_\Omega \langle \mathbb{C}\, q_1(x) \,, q_2(x) \rangle_F \,\mathrm{d}x \,, \qquad \|q\|_{\mathbb{C}} \;:=\; \langle q \,, q \rangle_{\mathbb{C}}^{1/2} \,, \tag{13}$$

$$\psi_k(q) \;:=\; \begin{cases} \int_\Omega \left(\frac{1}{2}\tilde{\alpha}_k(q)^2 + \sigma_y \|q - p_{k-1}\|_F \right) \,\mathrm{d}x & \text{if } \operatorname{tr} q = \operatorname{tr} p_{k-1} \,, \\ +\infty & \text{else} \,, \end{cases} \tag{14}$$

$$l_k(v) \;:=\; \int_\Omega f_k \cdot v \,\mathrm{d}x + \int_{\Gamma_N} g_k \cdot v \,\mathrm{d}s \,, \tag{15}$$

where $\tilde{\alpha}_k(q)$ is defined in (12). Then the functional $\bar{J}_k(v, q)$ in (11) simply rewrites:

$$\bar{J}_k(v, q) = \frac{1}{2} \|\varepsilon(v) - q\|_{\mathbb{C}}^2 + \psi_k(q) - l_k(v) \,. \tag{16}$$

The following results are formulated for functionals mapping from a Hilbert space $\mathcal{H}$ into the set of extended real numbers $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. The Hilbert space $\mathcal{H}$ provides a scalar product $\langle \circ \,, \diamond \rangle_{\mathcal{H}}$ and the norm $\|\cdot\|_{\mathcal{H}} := \langle \cdot \,, \cdot \rangle_{\mathcal{H}}^{1/2}$. The topological dual space of $\mathcal{H}$ is denoted by $\mathcal{H}^*$. Further, if a function F is Fréchet differentiable, we will denote its derivative in a point x by $D\,F(x)$ and its Gâteaux differential into the direction y by $D\,F(x\,;\,y)$.

Definition 1 (convexity)**.** Let F be a mapping of $\mathcal{H}$ into $\overline{\mathbb{R}}$. F is said to be *convex* if, for every x and y in $\mathcal{H}$, we have

$$F(tx + (1-t)y) \leq tF(x) + (1-t)F(y) \qquad \forall t \in [0,1] \ , \tag{17}$$

whenever the right hand side is defined.

Definition 2 (strict convexity)**.** Let F be a mapping of $\mathcal{H}$ into $\mathbb{R}$. F is said to be *strictly convex* if it is convex and the strict inequality holds in (17) for all $x, y \in \mathcal{H}$ with $x \neq y$ and for all $t \in \,]0,1[$.

Definition 3 (proper function, effective domain)**.** Let F be a mapping of $\mathcal{H}$ into $\overline{\mathbb{R}}$. F is said to be *proper* if there exists $x \in \mathcal{H}$ such that $F(x) < +\infty$ and if for all $y \in \mathcal{H}$ there holds $F(y) > -\infty$. The set $\{x \in \mathcal{H} \mid F(x) < +\infty\}$ is said to be the *effective domain* of F and denoted by $\operatorname{dom} F$.

Definition 4 (subdifferential)**.** Let F be a mapping of $\mathcal{H}$ into $\overline{\mathbb{R}}$. F is said to be *subdifferentiable* at the point $x \in \mathcal{H}$ if there exists $x^* \in \mathcal{H}^*$ such that $F(x+y) \geq F(x) + \langle x^* \, , \, y \rangle_{\mathcal{H}}$ holds for all $y \in \mathcal{H}$. We call x^* a *subgradient*, and the set of all subgradients in x is said to be the *subdifferential* of F in x and denoted by $\partial F(x)$.

The following lemma summarizes three well known results from convex analysis which will be frequently used later.

Lemma 1. *Let $F : \mathcal{H} \to \mathbb{R}$ be a convex function. Then the following two properties hold:*

a) F is continuous in $\mathcal{H}$ if and only if there exists a non-empty open subset $U \subset \mathcal{H}$ on which F is bounded above by a constant $C \in \mathbb{R}$.

b) If F is continuous, then F is subdifferentiable in $\mathcal{H}$.

c) If F is continuous and has a unique subgradient at $y \in \mathcal{H}$, then F is Fréchet differentiable at y and its derivative is identical to the subgradient.

Proof. Ad a), see [ET99, Proposition 2.5]. Ad b), see [ET99, Proposition 5.2]. Ad c), see [ET99, Proposition 5.3]. ☐

Now we can formulate a theorem, which can be seen as a generalization of a work of J. J. Moreau [Mor65]. The precise difference is discussed later in Remark 1 on page 10.

Theorem 1. *Let $\Phi : \mathcal{H} \to \mathbb{R}$ be a convex and Fréchet differentiable function with the derivative $D\,\Phi \in \mathcal{H}^*$, and let $\Psi : \mathcal{H} \to \overline{\mathbb{R}}$ be a convex and proper function. We define the functions $f : \mathcal{H} \times \mathcal{H} \to \overline{\mathbb{R}}$ and $F : \mathcal{H} \to \mathbb{R}$ by*

$$f(x,y) := \Phi(x-y) + \Psi(x) \quad \textit{and} \quad F(y) := \inf_{x \in \mathcal{H}} f(x,y)\,. \tag{18}$$

Let us assume additionally, that the infimum $F(y)$ is attained for all $y \in \mathcal{H}$, that is, there exists a function $\tilde{x} : \mathcal{H} \to \mathcal{H}$ such that $F(y) = f(\tilde{x}(y), y)$. Then the following statements are valid:

1. *F is convex and continuous in $\mathcal{H}$. If either Φ is strictly convex is $\mathcal{H}$ or Ψ is strictly convex in its effective domain, then F is strictly convex in $\mathcal{H}$.*
2. *The subdifferential of F writes $\partial F(y) = \{-D\,\Phi(\tilde{x}(y) - y)\}$ for all $y \in \mathcal{H}$.*

Proof. Hence Φ is finite and Ψ is proper, the function $f(\cdot, y) = \Phi(\cdot - y) + \Psi(\cdot)$ is proper with respect to the first argument for all $y \in \mathcal{H}$. Due to the minimization property of $\tilde{x}$, there holds that $\Psi(\tilde{x}(y))$ and $F(y)$ are finite for all $y \in \mathcal{H}$. Thus, F in (18) is well defined as a mapping of $\mathcal{H}$ into $\mathbb{R}$. Moreover, we note that $f(\tilde{x}(y), z)$ is finite for all y and z in $\mathcal{H}$. For the convexity of F, we must check that

$$F(ty_1 + (1-t)y_2) \leq tF(y_1) + (1-t)F(y_2)$$

for all $y_1 \in \mathcal{H}$, $y_2 \in \mathcal{H}$ and $t \in [0, 1]$. Let $\overline{y} := ty_1 + (1-t)y_2$ and $\overline{x} := t\tilde{x}(y_1) + (1-t)\tilde{x}(y_2)$. Utilizing the minimization property of $\tilde{x}$ we obtain

$$F(ty_1 + (1-t)y_2) = F(\overline{y}) = f(\tilde{x}(\overline{y}), \overline{y}) \leq f(\overline{x}, \overline{y}). \tag{19}$$

Using the structure $f(x, y) = \Phi(x - y) + \Psi(y)$ and the convexity of Φ and Ψ, elementary calculations yield

$$f(\overline{x}, \overline{y}) \leq tf(\tilde{x}(y_1), y_1) + (1-t)f(\tilde{x}(y_2), y_2) = tF(y_1) + (1-t)F(y_2). \tag{20}$$

The substitution of (20) in (19) proves the convexity of F. If either Φ or $\Psi_{|\mathrm{dom}\,\Psi}$ was strictly convex, the inequality in (20) would hold strictly for $y_1 \neq y_2$ and $t \in \,]0, 1[$. As a result, F would be strictly convex. It remains to show, that F is continuous in $\mathcal{H}$. We arbitrarily fix $\hat{x} \in \mathrm{dom}\,\Psi$, $\hat{y} \in \mathcal{H}$, and $\epsilon > 0$. Then, obviously

$$F(y) = \inf_{x \in \mathcal{H}} (\Phi(x - y) + \Psi(x)) \leq \Phi(\hat{x} - y) + \Psi(\hat{x}).$$

Since Φ is continuous in $\hat{x} - \hat{y}$, there exists $\delta > 0$, such that for all $y : \|y - \hat{y}\|_{\mathcal{H}} < \delta$ there holds $\Phi(\hat{x} - y) + \Psi(\hat{x}) \leq \Phi(\hat{x} - \hat{y}) + \epsilon + \Psi(\hat{x})$. Thus, F is bounded above on the non-empty open set $U := \{y : \|y - \hat{y}\|_{\mathcal{H}} < \delta\}$ and Lemma 1 a) concludes the continuity of F in $\mathcal{H}$.

Note, that due to Lemma 1 b), the function F is subdifferentiable. Let $y \in \mathcal{H}$, and $G \in \partial F(y)$ be arbitrary. By the definition of the subdifferential, there holds

$$F(y + z) \geq F(y) + \langle G\,,\, z\rangle_{\mathcal{H}} \tag{21}$$

for all $z \in \mathcal{H}$. On the other hand, for all $z \in \mathcal{H}$, there holds

$$F(y + z) = f(\tilde{x}(y + z), y + z) \leq f(\tilde{x}(y), y + z)\,. \tag{22}$$

Since $f(x,y) = \Phi(x-y) + \Psi(x)$ and Φ is Fréchet differentiable, there exists a function $r : \mathcal{H} \to \mathbb{R}$ with the property $\lim_{z\to 0}|r(z)|/\|z\|_{\mathcal{H}} = 0$ such that

$$f(\tilde{x}(y), y+z) = \underbrace{f(\tilde{x}(y), y)}_{=F(y)} - \langle D\,\Phi(\tilde{x}(y) - y)\,,\, z\rangle_{\mathcal{H}} + r(z)\,. \tag{23}$$

Combining (22) and (23) we obtain

$$-F(y+z) \geq -F(y) + \langle D\,\Phi(\tilde{x}(y) - y)\,,\, z\rangle_{\mathcal{H}} - r(z)\,. \tag{24}$$

Summation of (21) and (24) yields $r(z) \geq \langle G + D\,\Phi(\tilde{x}(y)-y)\,,\, z\rangle_{\mathcal{H}} \geq -r(-z)$ for all $z \in \mathcal{H}$, and thus there holds $\lim_{z\to 0} \frac{\langle G + D\,\Phi(\tilde{x}(y)-y)\,,\, z\rangle_{\mathcal{H}}}{\|z\|_{\mathcal{H}}} = 0$, which implies $G = -D\,\Phi(\tilde{x}(y) - y)$. Since G was chosen arbitrarily in $\partial F(y)$, we end up with $\partial F(y) = \{-D\,\Phi(\tilde{x}(y) - y)\}$. $\square$

Notice, the subdifferential $\partial F(y)$ does not necessarily contain one element only, but depends on the set of functions $\tilde{x}$ satisfying $F(y) = f(\tilde{x}(y), y)$. If $\tilde{x}$ was unique, then $\partial F(y)$ would contain only one subgradient identical to derivative $D\,F(y)$ according to Lemma 1 *c*). We formulate a sufficient condition for the (unique) existence of $\tilde{x}$ under the assumptions of coercivity and lower semicontinuity.

Definition 5 (coercivity)**.** Let F be a mapping of $\mathcal{H}$ into $\overline{\mathbb{R}}$. F is said to be *coercive*, if for all $C \in \mathbb{R}$ there exists $K \in \mathbb{R}$ such that for all $x \in \mathcal{H}$ there holds

$$F(x) \leq C \Rightarrow \|x\|_{\mathcal{H}} \leq K\,.$$

Definition 6 (lower semicontinuity)**.** Let F be a mapping of $\mathcal{H}$ into $\overline{\mathbb{R}}$. F is said to be *lower semi continuous (l.s.c. for short)* at $x \in \mathcal{H}$ if

$$\lim_{y\to x} F(y) \geq F(x)\,.$$

F is said to be *l.s.c. in* $\mathcal{H}$ if F is l.s.c. at all $x \in \mathcal{H}$.

Theorem 2. *Let $F : \mathcal{H} \to \overline{\mathbb{R}}$ be l.s.c., proper, convex and coercive. Then there exists $\hat{x} \in \mathcal{H}$ such that $F(\hat{x}) = \inf_{x\in\mathcal{H}} F(x)$. If F is strictly convex, then $\hat{x}$ is unique.*

Proof. See [ET99, Proposition 1.2 of Chapter II]. $\square$

Corollary 1 (Moreau)**.** *Let the function $f : \mathcal{H} \times \mathcal{H} \to \overline{\mathbb{R}}$ be defined*

$$f(x,y) = \frac{1}{2}\|x - y\|^2_{\mathcal{H}} + \psi(x) \tag{25}$$

where ψ is a convex, proper, l.s.c. and coercive function of $\mathcal{H}$ into $\overline{\mathbb{R}}$. Then $F(y) := \inf_{x\in\mathcal{H}} f(x,y)$ defines a mapping $F : \mathcal{H} \to \mathbb{R}$ and there exists a unique function $\tilde{x} : \mathcal{H} \to \mathcal{H}$ such, that $F(y) = f(\tilde{x}(y), y)$ for all $y \in \mathcal{H}$, and there holds:

1. *F is strictly convex and continuous in $\mathcal{H}$.*

2. F is Fréchet differentiable with

$$D\,F(y) = \langle y - \tilde{x}(y)\,,\,\cdot\rangle_{\mathcal{H}} \in \mathcal{H}^* \quad \text{for all } y \in \mathcal{H}\,. \tag{26}$$

Proof. Let $y \in \mathcal{H}$ be fixed arbitrarily. Then, $f(\cdot, y)$ satisfies the assumptions of Theorem 2. Thus, there exists a unique element $\tilde{x}(y) \in \mathcal{H}$ such that $f(\tilde{x}(y), y) = F(y)$. Theorem 1 (by choosing $\Phi(z) := \frac{1}{2}\|z\|^2_{\mathcal{H}}$) states, that F is strictly convex, continuous and subdifferentiable with a unique subgradient $\langle y - \tilde{x}(y)\,,\,\cdot\rangle_{\mathcal{H}}$. Together with Lemma 1 c) we conclude that F is Fréchet differentiable with $D\,F(\cdot)$ as in (26). $\square$

Remark 1. This corollary was first formulated and proved in 1965 by J. J. Moreau [Mor65, 7.d. Proposition], and can be interpreted as an immediate consequence of Theorem 1 and Theorem 2.

Now, we apply Corollary 1 to Problem 2 to obtain the following proposition.

Proposition 1. *Let $k \in \{1, \ldots, N_\Theta\}$ denote the time step, and let $\bar{J}_k$ be defined as in (11). Then there exists a unique mapping $\tilde{p}_k : Q \to Q$ satisfying*

$$\bar{J}_k\left(v, \tilde{p}_k\left(\varepsilon\left(v\right)\right)\right) = \inf_{q \in Q} \bar{J}_k\left(v, q\right) \quad \forall v \in V_D\,. \tag{27}$$

Let J_k be a mapping of V_D into $\mathbb{R}$ defined as

$$J_k(v) := \bar{J}_k(v, \tilde{p}_k(\varepsilon(v))) \quad \forall v \in V_D\,. \tag{28}$$

Then, J_k is strictly convex and Fréchet differentiable. The associated Gâteaux differential reads

$$D\,J_k(v\,;\,w) = \langle \varepsilon(v) - \tilde{p}_k(\varepsilon(v))\,,\,\varepsilon(w)\rangle_{\mathbb{C}} - l_k(w) \quad \forall w \in V_0 \tag{29}$$

with the scalar product $\langle \circ\,,\,\diamond\rangle_{\mathbb{C}}$ defined in (13) and l_k defined in (15).

Proof. Recall, that the functional $\bar{J}_k : V \times Q \to \overline{\mathbb{R}}$ defined in (16) using (13), (14), and (15) can be decomposed as $\bar{J}_k(v, q) = f_k(\varepsilon(v), q) - l_k(v)$, where the functional $f_k : Q \times Q \to \overline{\mathbb{R}}$ reads

$$f_k(s, q) := \frac{1}{2}\|q - s\|^2_{\mathbb{C}} + \psi_k(q).$$

Then, Corollary 1 states an existence of a unique minimizer $\tilde{p}_k : Q \to Q$ which satisfies the condition $f_k(\varepsilon(v), \tilde{p}_k(\varepsilon(v))) = \inf_{q \in Q} f_k(\varepsilon(v), q)$, where the functional

$$F_k(\varepsilon(v)) := f_k(\varepsilon(v), \tilde{p}_k(\varepsilon(v)))$$

is strictly convex and differentiable with respect to $\varepsilon(v) \in Q$. Since $\varepsilon : v \to \varepsilon(v)$ is a Fréchet differentiable, linear and injective mapping of V_D into Q, the compound functional $F_k(\varepsilon(v))$ is Fréchet differentiable and strictly convex with respect to $v \in V_D$. Considering the Fréchet differentiability and linearity of l_k with respect to $v \in V_D$, we can conclude the strictly convexity and Fréchet differentiability (in V_D) of the functional J_k defined in (28). The explicit form of the Gâteaux differential $D\,J_k(v\,;\,w)$ in (29) results from the linearity of the two mappings l_k and ε, and the Fréchet derivative $DF_k(\varepsilon(v)\,;\,\cdot) = \langle \varepsilon(v) - \tilde{p}_k(\varepsilon(v))\,,\,\cdot\rangle_{\mathbb{C}}$ as in (26), combined using the chain rule for functionals. $\square$

Proposition 1 tells us, that for each displacement v there exists exactly one plastic strain $\tilde{p}_k(\varepsilon(v))$, such that the energy functional $J(v,q)$ attains its minimum $J(v,\tilde{p}_k(\varepsilon(v)))$. By the definition of $\Delta\tilde{p}_k(\cdot) := \tilde{p}_k(\cdot) - p_{k-1}$, there holds that for fixed $v \in V_D$ finding the minimizer $\tilde{p}_k(\varepsilon(v))$ of functional $\bar{J}_k(v,q)$ in (11) with respect to q is equivalent to finding the minimizer $\Delta\tilde{p}_k(\varepsilon(v))$ of the functional

$$\frac{1}{2}\left(2\mu + \sigma_y^2 H^2\right)\|q\|_Q^2 - \langle \mathbb{C}\left(\varepsilon(v) - p_{k-1}\right)\,,\,q\rangle_Q + \langle \sigma_y\left(1 + \alpha_{k-1} H\right)\,,\,\|q\|_F\rangle_{L_2} \tag{30}$$

amongst trace-free elements $q \in Q$.

The explicit form of $\Delta\tilde{p}_k$ is presented in the following theorem, which is a generalization of [ACZ99, Proposition 7.1] in the sense we analyse the plastic strain field instead of the pointwise value. The validity of the pointwise equalities and inequalities occurring there, has to be understood in accordance with Lebesque spaces as almost everywhere (denoted a. e.), i.e. up to a set of a zero measure.

Theorem 3. *Let $Q = L_2(\Omega)^{3\times 3}_{sym}$, $A \in Q$, $b \in L_2(\Omega)$ with $b(x) > 0$ in Ω, and $\xi \in \mathbb{R}$ with $\xi > 0$. Then there exists exactly one $p \in Q$ with $\|\operatorname{tr} p\|_{L_2(\Omega)} = 0$, that satisfies*

$$\langle A - \xi p\,,\,q - p\rangle_Q \leq \langle b\,,\,\|q\|_F - \|p\|_F\rangle_{L_2} \tag{31}$$

for all $q \in Q$ with $\|\operatorname{tr} q = 0\|_{L_2(\Omega)}$. This p is characterized as the minimizer of

$$\frac{\xi}{2}\|q\|_Q^2 - \langle A\,,\,q\rangle_Q + \langle b\,,\,\|q\|_F\rangle_{L_2} \tag{32}$$

amongst trace-free elements $q \in Q$, and reads

$$p = \frac{1}{\xi}\max\{0, \|\operatorname{dev} A\|_F - b\}\frac{\operatorname{dev} A}{\|\operatorname{dev} A\|_F} \qquad \text{on } \Omega. \tag{33}$$

The minimal value of (32), attained for p as in (33), is

$$-\frac{1}{2\xi}\|\max\{0, \|\operatorname{dev} A\|_F - b\}\|_{L_2}^2\,. \tag{34}$$

Proof. According to Definition 4, expression (31) states that

$$A - \xi p \in b\,\partial\|\cdot\|_F(p) \tag{35}$$

where $\partial\|\cdot\|_F$ denotes the subgradient of the Frobenius norm, and only trace-free arguments are under consideration. The Frobenius norm $\|\cdot\|_F : Q \to \mathbb{R}$ is a convex functional and so is (32). The identity (35) is equivalent to 0 belonging to the subgradient of (32), which characterizes the minimizers of (32). Moreover, there holds $\langle A\,,\,q\rangle_Q = \langle \operatorname{dev} A\,,\,q\rangle_Q$ for all trace-free elements $q \in Q$, whence the matrix A can be replaced by the matrix $\operatorname{dev} A$ in (31) and (32).

Let us separate the domain Ω into three disjoint subdomains

$$\begin{aligned}\Omega_e &:= \{x \in \Omega \mid \exists \text{ open } \omega \subset \Omega :\ x \in \omega \wedge \|\operatorname{dev} A\|_F - b \le 0 \text{ in } \omega\}\,,\\ \Omega_p &:= \Omega \setminus \overline{\Omega_e}\,, \quad \Gamma_{ep} := \Omega \setminus (\Omega_e \cup \Omega_p)\,.\end{aligned}$$

Note that Ω_e and Ω_p are open, Γ_{ep} has zero measure, it holds $\|\operatorname{dev} A\|_F - b \le 0$ on Ω_e and $\|\operatorname{dev} A\|_F - b > 0$ on Ω_p. Consequently, the minimization of (32) results in finding $p \in Q$ with $\|\operatorname{tr} p\|_{L_2(\Omega)} = 0$, such that the functionals

$$J_i(p) := \frac{\xi}{2}\int_{\Omega_i} \|p\|_F^2 \,\mathrm{d}x - \int_{\Omega_i} \langle \operatorname{dev} A\,,\,p\rangle_F \,\mathrm{d}x + \int_{\Omega_i} b\|p\|_F \,\mathrm{d}x \quad i \in \{e,p\} \tag{36}$$

are minimized, or equivalently the inequalities

$$\int_{\Omega_i} \langle \operatorname{dev} A - \xi p\,,\,q - p\rangle_F \,\mathrm{d}x \le \int_{\Omega_i} b\left(\|q\|_F - \|p\|_F\right)\,\mathrm{d}x \quad i \in \{e,p\} \tag{37}$$

are satisfied for all $q \in Q$ with $\|\operatorname{tr} q\|_{L_2(\Omega)} = 0$.

We will show identity (33). An application of the pointwise Cauchy-Schwarz inequality $\langle \operatorname{dev} A\,,\,p\rangle_F \le \|\operatorname{dev} A\|_F\|p\|_F$ yields

$$J_e(p) \ge \frac{\xi}{2}\int_{\Omega_e} \|p\|_F^2 \,\mathrm{d}x + \int_{\Omega_e} \underbrace{(b - \|\operatorname{dev} A\|_F)}_{\ge 0}\|p\|_F \,\mathrm{d}x \ge 0.$$

By choosing $p = 0$ on Ω_e we obtain $J_e(p) = 0$. Therefore,

$$p = 0 \qquad \text{on } \Omega_e \tag{38}$$

minimizes J_e in (36). Moreover, there holds $p(x) \ne 0$ on Ω_p which we show by contradiction. Choose $\Omega' \subset \Omega_p$ arbitrary and fix. Assuming, that $p = 0$ on Ω' and plugging it into (37) for $i = p$ would yield

$$\int_{\Omega'} \langle \operatorname{dev} A\,,\,q\rangle_F \,\mathrm{d}x \le \int_{\Omega'} b\|q\|_F \,\mathrm{d}x$$

for all trace-free elements $q \in Q$, which satisfy $q = p$ on $\Omega_p \setminus \Omega'$. By the choice of $q = \operatorname{dev} A$ on Ω' one obtains $\int_{\Omega'} \|\operatorname{dev} A\|_F - b \,\mathrm{d}x \le 0$ and this would be a contradiction to the definition of Ω_p.

Thus there holds $p(x) \ne 0$ and consequently $\partial\|\cdot\|_F(p) = \{p/\|p\|_F\}$ on Ω_p, whence (37) with $i = p$ rewrites

$$\int_{\Omega_p} \left(\operatorname{dev} A - \xi\, p - b\,\frac{p}{\|p\|_F}\right) : q \,\mathrm{d}x = 0 \qquad \forall q \in Q\,,\ \|\operatorname{tr} q\|_{L_2(\Omega)} = 0\,.$$

Necessarily, there must hold

$$\operatorname{dev} A - \xi\, p - b\, \frac{p}{\|p\|_F} = 0 \qquad \text{on } \Omega_p\,, \tag{39}$$

whence we conclude

$$\|p\|_F = \frac{1}{\xi}\left(\|\operatorname{dev} A\|_F - b\right)\,. \tag{40}$$

Plugging (40) into (39) yields

$$p = \frac{1}{\xi}\left(\|\operatorname{dev} A\|_F - b\right)\frac{\operatorname{dev} A}{\|\operatorname{dev} A\|_F} \text{ on } \Omega_p. \tag{41}$$

Combining the formulae (38) and (41) we obtain (33). Finally, plugging (33) into (32) yields (34). □

We define the trial stress $\tilde{\sigma}_k : Q \to Q$ at the kth time step and the yield function $\phi_{k-1} : Q \to \mathbb{R}$ (cf. (8)) at the $k-1$st time step by

$$\tilde{\sigma}_k(q) := \mathbb{C}(q - p_{k-1}) \quad \text{and} \quad \phi_{k-1}(\sigma) := \|\operatorname{dev}\sigma\|_F - \sigma_y(1 + H\,\alpha_{k-1})\,. \tag{42}$$

After using the substitution $\Delta\tilde{p}_k(\varepsilon(v)) = \tilde{p}_k(\varepsilon(v)) - p_{k-1}$, Theorem 3 tells us that for a fixed displacement $v \in V_D$ the minimizer $\tilde{p}_k(\varepsilon(v))$ of (11) reads

$$\tilde{p}_k(\varepsilon(v)) = \frac{1}{2\mu + \sigma_y^2 H^2}\max\{0, \phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v)))\}\frac{\operatorname{dev}\tilde{\sigma}_k(\varepsilon(v))}{\|\operatorname{dev}\tilde{\sigma}_k(\varepsilon(v))\|_F} + p_{k-1}\,. \tag{43}$$

Therefore, if the minimizer $u_k \in V_D$ of the functional $J_k(\cdot) = \bar{J}_k(\cdot, \tilde{p}_k(\varepsilon(\cdot)))$ in (28) is known, then the plastic strain p_k at the time step k is provided by the formula (43) as $p_k = \tilde{p}_k(\varepsilon(u_k))$. Notice that the formula (43) also satisfies the necessary condition $\operatorname{tr} p_k = \operatorname{tr} p_{k-1}$ to guarantee the minimization property $J_k(u_k) = \bar{J}_k(u_k, p_k) < +\infty$ (cf. (16) and (14)).

At each time step k the domain Ω can be decomposed into three disjoint parts (see Figure 1), analogously to the decomposition we used in the proof to Theorem 3:

- $\Omega_k^e(v) := \{x \in \Omega \mid \exists \text{ open } \omega \subset \Omega:\ x \in \omega \wedge \phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v))) \le 0 \ a.\ e. \text{ in } \omega\}$, which is the set of elastic increment points,
- the set of plastic increment points $\Omega_k^p(v) := \Omega \setminus \overline{\Omega_k^e}(v)$,
- and the set of elastoplastic interface points $\Gamma_k^{ep}(v) := \Omega \setminus (\Omega_k^p(v) \cup \Omega_k^e(v))$.

Obviously, both sets $\Omega_k^e(v)$ and $\Omega_k^p(v)$ are open, $\Gamma_k^{ep}(v)$ has zero measure, and that

$$\begin{aligned} \phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v))) &\le 0 \quad \text{a. e. in } \Omega_k^e(v),\\ \phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v))) &> 0 \quad \text{a. e. in } \Omega_k^p(v). \end{aligned} \tag{44}$$

For a one-time step problem, the sets $\Omega^e(v) := \Omega_1^e(v)$ and $\Omega^p(v) := \Omega_1^p(v)$ specify elastically and plastically deformed parts of the continuum, respectively.

We obtain a smooth minimization problem with respect to the displacement field u_k only:

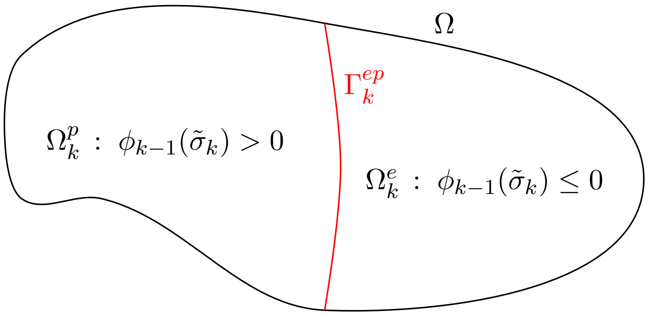

Figure 1: Domain decomposition of Ω at the kth time step, generated by the trial stress $\tilde\sigma_k(\varepsilon(v))(x)$ with $x \in \Omega$, as an argument of the yield functional ϕ_{k-1} (cf. 42).

Problem 3. Let $k \in \{1, \ldots, N_\Theta\}$ denote the time step. Let $p_{k-1} \in Q$ and $\alpha_{k-1} \in L_2(\Omega)$ be given, such that $\alpha_{k-1} \geq 0$ almost everywhere. Find $u_k \in V_D$ such that for all $v \in V_D$ there holds $J_k(u_k) \leq J_k(v)$ with the strictly convex and Fréchet differentiable functional J_k defined in (28) using $\tilde p_k$ as in (43). The Gâteaux differential of J_k is presented in (29).

Remark 2 (unique existence of a solution). We know, that there exists a unique solution (u_k, p_k) to Problem 2, and the second component p_k can be calculated by the identity $p_k = \tilde p_k(\varepsilon(u_k))$ explicitely. This implies that, due to the definition $J_k(\cdot) = \bar J_k(\cdot, \tilde p_k(\cdot))$, there holds $J_k(u_k) \leq J_k(v)$ for all $v \in V_D$. Thus, there exists a solution, namely $u_k \in V_D$, to Problem 3. The uniqueness of the solution follows from the strict convexity of the energy functional J_k, as it is shown in Proposition 1.

4 Computing a Solution of the Smooth Problem by Means of a Newton-like Method

The minimizer $\tilde p_k$ in (43) is a continuous mapping of Q into Q. Thus, $D\,J_k(v\,;w)$ in (29) is continuous with respect to v as well, and a gradient method could be used for a numerical solution. Instead, we investigate the existence of the second derivative of $J_k(v)$, which would allow the use of Newton's method or at least some Newton-like method.

4.1 An Attempt to Calculate the Second Derivative of J_k

The Gâteaux differential of $D\,J_k$ defined in (29) reads

$$D^2 J_k(v\,;w_1,w_2) = \langle \varepsilon(w_1) - D\,\tilde p_k(\varepsilon(v)\,;\varepsilon(w_1))\,,\,\varepsilon(w_2)\rangle_{\mathbb{C}} \quad \forall w_1, w_2 \in V_0$$

provided that the Gâteaux differential $D\,\tilde p_k(\varepsilon(v)\,;\varepsilon(w_1)) \in Q$ of the plastic strain minimizer $\tilde p_k(\varepsilon(v))$ defined in (43) exists in the whole domain Ω.

In the set of elastic increment points $\Omega_k^e(v)$, where $\phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v))) \le 0$ (cf. (42)), there obviously holds

$$D\,\tilde{p}_k(\varepsilon(v)\,;\,q) = 0 \tag{45}$$

for all $q \in Q$, and therefore we obtain the formula known from theory of elasticity

$$D^{\,2} J_k(v\,;\,w_1, w_2) = \langle \varepsilon(w_1)\,,\,\varepsilon(w_2)\rangle_{\mathbb{C}} \quad \forall w_1, w_2 \in V_0.$$

In the set of plastic increment points $\Omega_k^p(v)$, where $\phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v))) > 0$ holds a. e., the plastic strain reads

$$\tilde{p}_k(\varepsilon) = (2\mu + \sigma_y^2 H^2)^{-1} \phi_{k-1}(\tilde{\sigma}_k(\varepsilon))\,\frac{\operatorname{dev}\tilde{\sigma}_k(\varepsilon)}{\|\operatorname{dev}\tilde{\sigma}_k(\varepsilon)\|_F}\,.$$

For the moment, we omit the dependency of ε on v in our notation, and calculate the Gâteaux differential of $\tilde{p}_k$ with respect to ε. By using the product and the chain rules, we obtain

$$\begin{aligned} D\,\tilde{p}_k(\varepsilon\,;\,q) = (2\mu + \sigma_y^2 H^2)^{-1} \Big(& D\,\phi_{k-1}(\tilde{\sigma}_k(\varepsilon)\,;\,D\,\tilde{\sigma}_k(\varepsilon\,;\,q))\frac{\operatorname{dev}\tilde{\sigma}_k(\varepsilon)}{\|\operatorname{dev}\tilde{\sigma}_k(\varepsilon)\|_F} \\ & + \phi_{k-1}(\tilde{\sigma}_k(\varepsilon)) D\,\frac{(\cdot)}{\|\cdot\|_F}(\operatorname{dev}\tilde{\sigma}_k(\varepsilon)\,;\,D\,\operatorname{dev}\tilde{\sigma}_k(\varepsilon\,;\,q))\Big)\,. \end{aligned}$$

Using the derivatives rules (cf. (42))

$$D\,\tilde{\sigma}_k(\varepsilon\,;\,q) = D\,\tilde{\sigma}_k(\,q) = \mathbb{C}\,q, \quad D\,\operatorname{dev}\tilde{\sigma}_k(\varepsilon\,;\,q) = D\,\operatorname{dev}\tilde{\sigma}_k(\,q) = 2\mu\operatorname{dev} q$$

and

$$D\,\phi_{k-1}(\sigma\,;\,\tau) = \frac{\langle \operatorname{dev}\sigma\,,\,D\,\operatorname{dev}(\sigma\,;\,\tau)\rangle_F}{\|\sigma\|_F}\,, \qquad D\,\frac{(\cdot)}{\|\cdot\|_F}(\sigma\,;\,\tau) = \frac{\tau}{\|\sigma\|_F} - \frac{\sigma\langle\sigma\,,\,\tau\rangle_F}{\|\sigma\|_F^3},$$

we end up with the formula

$$\begin{aligned} D\,\tilde{p}_k(\varepsilon\,;\,q) = \frac{2\mu}{2\mu + \sigma_y^2 H^2} \Bigg(& \frac{\phi_{k-1}(\varepsilon)}{\|\operatorname{dev}\tilde{\sigma}_k(\varepsilon)\|_F}\operatorname{dev} q \\ & + \left(1 - \frac{\phi_{k-1}(\varepsilon)}{\|\operatorname{dev}\tilde{\sigma}_k(\varepsilon)\|_F}\right)\frac{\langle\operatorname{dev}\tilde{\sigma}_k(\varepsilon)\,,\,\operatorname{dev} q\rangle_F}{\|\operatorname{dev}\tilde{\sigma}_k(\varepsilon)\|_F^2}\operatorname{dev}\tilde{\sigma}_k(\varepsilon)\Bigg)\,. \end{aligned} \tag{46}$$

The set of elastoplastic interface points $\Gamma_k^{ep}(v)$ represents the only part of the domain Ω where $\tilde{p}_k$ in (43), due to the term $\max\{0, \phi_{k-1}\}$, is not differentiable.

To summarize it, the second derivative $D^{\,2}J_k(v)$ exists everywhere in the sets of elastic and plastic increment points, but is not computable on the elastoplastic interface (see Figure 1). No matter that the elastoplastic interface is a set of zero measure, a classical Newton method is not applicable to Problem 3.

4.2 Concept of Slant Differentiability

Our goal here is to solve Problem 3 by means of a Newton-like method which replaces the requirement of the second derivative $D^2 J_k(v)$ on the elastoplastic interface in a way that the local superlinear convergence rate can be shown.

The main tool here to overcome the non-differentiability of $D\,J_k$ due to the mapping $\max\{0,\cdot\}$ is the concept of *slant differentiability*, which was recently introduced by X. Chen, Z. Nashed and L. Qi in [CNQ01]. Other concepts of semi smoothness, e. g. [Ulb03], or the regularization of the non-differentiable terms, e. g. [Kie06], are not discussed here and might be considered for alternate analysis of elastoplastic problems.

Henceforth, let X, Y, and Z be Banach spaces, and $\mathcal{L}(\circ,\diamond)$ denote the set of all linear mappings of the set $\circ$ into the set $\diamond$.

Definition 7 (slant differentiability pointwise)**.** Let $U \subseteq X$ be an open subset and $x \in U$. A function $F : U \to Y$ is said to be *slantly differentiable at x* if there exist

1. mappings $F^o : U \to \mathcal{L}(X,Y)$ and $r : X \to Y$ with $\lim_{h\to 0} \frac{\|r(h)\|}{\|h\|} = 0$ such, that
$$F(x+h) = F(x) + F^o(x+h)\,h + r(h)$$
holds for all $h \in X$ satisfying $(x+h) \in U$, and
2. constants $\delta > 0$ and $C > 0$ such that for all $h \in X$ with $\|h\| < \delta$ there holds
$$\|F^o(x+h)\| := \sup_{y\in X\setminus\{0\}} \frac{\|F^o(x+h)\,y\|}{\|y\|} \le C\,.$$

We say, that $F^o(x)$ is a *slanting function for F at x.*

Definition 8 (slant differentiability in an open set)**.** Let $U \subseteq X$ be an open subset. A function $F : U \to Y$ is said to be *slantly differentiable in U* if there exists $F^o : U \to \mathcal{L}(X,Y)$ such that F^o is a slanting function for F at every point $x \in U$. F^o is said to be a *slanting function for F in U*. The set of all functions which are slantly differentiable in U and map to Y is denoted by $\mathcal{S}(U;Y)$.

Remark 3. In analogy to the relation between Gâteaux differential and Gâteaux derivative, we define the *slanting differential for F^o at x along the direction h* by $\tilde{F}^o : U \times X \to Y$ with $\tilde{F}^o(x\,;\,h) := F^o(x)h$. Since the mappings F^o and $\tilde{F}^o$ are taking a different number of arguments, it is sufficient, if we characterize both by the same denomination F^o and forget about $\tilde{F}^o$. In other words, we shall write $F^o(\cdot)$ for a slanting function and $F^o(\circ\,;\,\diamond)$ for the appropriate slanting differential for F.

Theorem 4. *Let $U \subseteq X$ be an open subset, and $F : U \to Y$ be a slantly differentiable function with a slanting function $F^o : U \to \mathcal{L}(X,Y)$. We suppose, that $x^* \in U$ is a* *solution to the nonlinear problem $F(x) = 0$. If $F^o(x)$ is non-singular for all $x \in U$ and $\{\|F^o(x)^{-1}\| : x \in U\}$ is bounded, then the Newton-like iteration*

$$x^{j+1} = x^j - F^o(x^j)^{-1} F(x^j) \tag{47}$$

converges super-linearly to x^*, provided that $\|x^0 - x^*\|$ is sufficiently small.

Proof. See [CNQ01, Theorem 3.4] or [HIK02, Theorem 1.1]. $\square$

We solve the smooth minimization problem in the displacement (Problem 3) by finding $u_k \in V_D$ such, that $D\,J_k(u_k; w) = 0$ for all $w \in V_0$ with $D\,J_k$ as in (29). Therefore, we use the Newton-like method (47) with the choice

$$X = V\,, \qquad Y = {V_0}^*\,, \qquad U = V_D\,, \qquad F = D\,J_k\,, \qquad x^j = v^j\,, \qquad \text{and} \qquad x^* = u_k\,.$$

The iteration scheme for the Newton-like method is formulated either as an identity in ${V_0}^*$ or in $\mathbb{R}$:

$$\text{Find } v^{j+1} \text{ in } V_D \quad (D\,J_k)^o\,(v^j\,;\, v^{j+1} - v^j) = -D\,J_k(v^j)\,. \tag{48}$$

$$\text{Find } v^{j+1} \text{ in } V_D \quad (D\,J_k)^o\,(v^j\,;\, v^{j+1} - v^j\,,\, w) = -D\,J_k(v^j\,;\, w) \quad \forall w \in V_0\,. \tag{49}$$

4.3 Slanting Functions for $\tilde{p}_k$ and $D\,J_k$

Let us now calculate a slanting function $(D\,J_k)^o$ for $D\,J_k$ in V_D. Henceforth we will use the following property, which is easy to verify: A Fréchet differentiable function is slantly differentiable, with the Fréchet derivative serving as a slanting function, and the Gâteaux differential serving as a slanting differential. Due to the chain rule for slanting functions (Theorem 7 in the Appendix) we obtain

$$(D\,J_k)^o\,(v\,;\, w_1, w_2) = \langle \varepsilon(w_1) - \tilde{p_k}^o(\varepsilon(v)\,;\, \varepsilon(w_1))\,,\, \varepsilon(w_2)\rangle_{\mathbb{C}} \quad \forall w_1, w_2 \in V_0\,. \tag{50}$$

It remains to calculate the slanting function $\tilde{p_k}^o$. Taking to account, that a Fréchet derivative serves as a slanting function, we obtain from (45) and (46), that

$$\tilde{p_k}^o(\varepsilon(v)\,;\, q) = \begin{cases} 0 & \text{in } \Omega_k^e(v)\,, \\ \xi\left(\beta_k \operatorname{dev} q + (1-\beta_k)\,\frac{\langle \operatorname{dev}\tilde{\sigma}_k\,,\, \operatorname{dev} q\rangle_F}{\|\operatorname{dev}\tilde{\sigma}_k\|_F^2} \operatorname{dev}\tilde{\sigma}_k\right) & \text{in } \Omega_k^p(v)\,, \end{cases}$$

where the abbreviations

$$\xi := \frac{2\mu}{2\mu + \sigma_y^2 H^2}, \quad \beta_k := \frac{\phi_{k-1}(\tilde{\sigma}_k)}{\|\operatorname{dev}\tilde{\sigma}_k\|_F}, \quad \tilde{\sigma}_k := \tilde{\sigma}_k(\varepsilon(v)) \tag{51}$$

with the mappings ϕ_{k-1} and $\tilde{\sigma}_k$ defined in (42) are used. Since the modulus of hardening H, the yield stress σ_y, and the Lamé parameter μ are positive and due to (12), (42) and (44), we always have

$$\xi \in \,]0,1[\qquad \text{and} \qquad \beta_k : \Omega_k^p(v) \to \,]0,1[\;. \tag{52}$$

The minimizer $\tilde{p}_k$ is not differentiable on the whole domain Ω, since it is not differentiable on $\Gamma_k^{ep}(v)$ due to the term $\max\{0, \phi_k\}$ in (43).

M. Hintermüller, K. Ito and K. Kunisch discuss the slant differentiability of the mapping $\max\{0, y\}$ for certain Banach spaces, that is, for the finite dimensional case $y \in \mathbb{R}^n$ in [HIK02, Lemma 3.1], and the infinite dimensional case $y \in L_q(\Omega)$ in [HIK02, Proposition 4.1]. Let us summarize their results in the following two theorems.

Theorem 5 (The finite dimensional case)**.** *Let $n \in N$ be arbitrary, and F be a mapping of $\mathbb{R}^n$ into $\mathbb{R}^n$ defined as $F(y) := \max\{0, y\}$. Then, F is slantly differentiable, and, for all $\gamma \in \mathbb{R}^n$, the matrix valued function*

$$F^o(y) := \operatorname{diag}\left(f_i(y_i)\right)_{i=1}^n \quad \text{with} \quad f_i(z) = \begin{cases} 0 & \text{if } z < 0\,, \\ 1 & \text{if } z > 0\,, \\ \gamma_i & \text{if } z = 0 \end{cases} \tag{53}$$

serves as a slanting function.

The next theorem addresses the slant differentiability of the mapping $\max\{0, y\}$ in the infinite dimensional case $y \in L_q(\Omega)$. Therefore we require a decomposition of the domain Ω into three distinct subspaces $\Omega = \Omega_{\leq} \cup \Gamma_{|} \cup \Omega_{>}$, where $\Omega_{>}$ denotes the union of all open subsets of Ω satisfying $y(x) > 0$ a. e., $\Omega_{\leq}$ is the interior of the complement of $\Omega_{>}$ with respect to Ω, and $\Gamma_{|}$ denotes the interface between $\Omega_{>}$ and $\Omega_{\leq}$.

Theorem 6 (The infinite dimensional case)**.** *Let p and q in $\mathbb{R}$ be fixed arbitrarily such that $1 \leq p \leq q \leq +\infty$ is satisfied, and let F be a mapping of $L_q(\Omega)$ into $L_p(\Omega)$ defined as $F(y) := \max\{0, y\}$. Then there holds, that for γ fixed arbitrarily in $\mathbb{R}$, the function*

$$F^o(y)(x) := \begin{cases} 0 & \text{on } \Omega_{\leq}\,, \\ 1 & \text{on } \Omega_{>}\,, \\ \gamma & \text{on } \Gamma_{|}\,. \end{cases} \tag{54}$$

serves as a slanting function for F if $p < q$, but F^o does in general not serve as a slanting function for F if $p = q$.

We apply the last two theorems to find a slanting function for the minimizer $\tilde{p}_k(\varepsilon)$ defined in (43) each in the continuous and the spatially discretized case.

The task turns out to be trivial in the latter case (see Section 5), but some further regularity assumptions are required in the continuous case due to the following considerations: The minimizer $\tilde{p}_k$ works as a mapping $Q \to Q$ in order to keep the energy functional J_k in (28) well-defined. The explicit formula (43) says, that $\tilde{p}_k$ maps into Q if and only if

$$\max\{0, \phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v)))\}$$

maps into $L_2(\Omega)$, where ϕ_{k-1} and $\tilde{\sigma}_k$ are defined in (42).

To apply Theorem 6 to the slant differentiation of the max-term measured in the $L_2(\Omega)$-norm, we must guarantee, that its argument

$$\phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v)))$$

is bounded in the $L_{2+\epsilon}(\Omega)$-norm for some $\epsilon > 0$ and for all $v \in V_D$, or at least for those $v \in V_D$ which are run through by the Newton-like method. This issue is not further discussed in this work, but left as an open question for theoretical analysis on regularities of elastoplastic problems. See Table 5 in the Appendix on page 36 for a compact summary of the still open issue.

Thus, under the assumption $\phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v))) \in L_{2+\epsilon}(\Omega)$, we can formulate an immediate result as the combination of the chain rule, Theorem 6 (with the setting $\gamma = 0$), and the explicit formula (43).

Corollary 2. *Let $k \in \{1, \ldots, N_\Theta\}$ and $v \in V_D$ be arbitrarily fixed. If there exists $\epsilon > 0$ such that $\phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v)))$, as defined in (42), is in $L_{2+\epsilon}(\Omega)$, then the mapping $\tilde{p}_k : Q \to Q$ defined in (43) is slantly differentiable at $\varepsilon(v)$. The mapping*

$$\tilde{p_k}^o(\varepsilon(v)\,;\,q) = \begin{cases} \xi \left(\beta_k \operatorname{dev} q + (1-\beta_k)\, \frac{\langle \operatorname{dev} \tilde{\sigma}_k \,,\, \operatorname{dev} q \rangle_F}{\|\operatorname{dev} \tilde{\sigma}_k\|_F^2} \operatorname{dev} \tilde{\sigma}_k \right) & \text{in } \Omega_k^p(v)\,, \\ 0 & \text{else}\,, \end{cases} \tag{55}$$

for all $q \in Q$ serves as a slanting function for $\tilde{p}_k$ at $\varepsilon(v)$, wherein the abbreviations (51) *together with the definitions* (42) *are used. Moreover, the functional $D\,J_k(v)$ is slantly differentiable with the slanting function $(D\,J_k)^o\,(v)$ as in (50).*

Corollary 2 corresponds to Corollary 3 in Section 5 on page 22, which states the slant differentiability of the energy functional's first derivative $D\,J_k$ and the plastic strain minimizer $\tilde{p}_k$ in finite dimensional FE-spaces. Unlike the infinite dimensional case, no additional assumptions will be necessary in the finite dimensional case (cf. Theorem 5 and Theorem 6).

4.4 Local Superlinear Convergence Rate of the Algorithm

In order to apply Theorem 4, the existence and boundedness of the inverse operator $[(D\,J_k)^o]^{-1}$ is required. It is proved in detail in Proposition 2 on page 21, which uses the boundedness and ellipticity of the bilinear form $(D\,J_k)^o\,(v) := (D\,J_k)^o\,(v\,;\,\diamond,\circ)$ from the following lemma.

Lemma 2. *Let $k \in \{1, \ldots, N_\Theta\}$ and $v \in V_D$ be fixed arbitrarily, and let the mapping $(D\,J_k)^o : V_D \to \mathcal{L}(V_0, V_0{}^*)$ be defined $(D\,J_k)^o\,(v) := (D\,J_k)^o\,(v\,;\,\diamond,\circ)$ as in (50) with the* *mapping $\tilde{p_k}^o$ as in (55). Then there exist positive constants κ_1 and κ_2 which satisfy*

$$(D\,J_k)^o\,(v\,;\,w,w) \quad \geq \quad \kappa_1 \|w\|_V^2 \quad \forall w \in V_0 \quad \textit{(ellipticity)}, \tag{56}$$

$$(D\,J_k)^o\,(v\,;\,w,\overline{w}) \quad \leq \quad \kappa_2 \|w\|_V \|\overline{w}\|_V \quad \forall w,\overline{w} \in V_0 \quad \textit{(boundedness)}. \tag{57}$$

Proof. Let us recall the definition of $(D\,J_k)^o$ in (50), i.e.,

$$(D\,J_k)^o\,(v\,;\,w,w) = \langle \varepsilon(w) - \tilde{p_k}^o(\varepsilon(v)\,;\,\varepsilon(w))\,,\,\varepsilon(w)\rangle_{\mathbb{C}}\,. \tag{58}$$

First, we prove the contractivity of the operator $p_k{}^o(\varepsilon(v),\cdot)$ defined in (55) with respect to its second argument:

$$\begin{aligned}
\|\tilde{p_k}^o(\varepsilon(v)\,;\,q)\|_{\mathbb{C}}^2 &= \int_\Omega \langle \mathbb{C} p_k{}^o(\varepsilon(v)\,;\,q)\,,\,p_k{}^o(\varepsilon(v)\,;\,q)\rangle_F \,\mathrm{d}x = 2\mu \int_\Omega \|p_k{}^o(\varepsilon(v)\,;\,q)\|_F^2 \,\mathrm{d}x \\
&= \xi^2\, 2\mu \int_{\Omega_k^p(v)} \|\beta_k \operatorname{dev} q + (1-\beta_k)\, \frac{\langle \operatorname{dev}\tilde{\sigma}_k\,,\,\operatorname{dev} q\rangle_F}{\|\operatorname{dev}\tilde{\sigma}_k\|_F^2} \operatorname{dev}\tilde{\sigma}_k\|_F^2 \,\mathrm{d}x \\
&\leq \xi^2\, 2\mu \int_\Omega \|\operatorname{dev} q\|_F^2 = \xi^2 \int_\Omega \langle \mathbb{C} \operatorname{dev} q\,,\,\operatorname{dev} q\rangle_F \,\mathrm{d}x \\
&\leq \xi^2 \int_\Omega \langle \mathbb{C} q\,,\,q\rangle_F \,\mathrm{d}x = \xi^2 \|q\|_{\mathbb{C}}^2 \quad \forall q \in Q.
\end{aligned}$$

Then the substitution of this estimate to (58) yields

$$(D\,J_k)^o\,(v\,;\,w,w) \geq (1-\xi)\|\varepsilon(w)\|_{\mathbb{C}}^2\,,$$

which together with Korn's inequality from the theory of linear elasticity (there exists a constant $\kappa_1^{\mathrm{e}} > 0$ such, that $\|\varepsilon(w)\|_{\mathbb{C}}^2 \geq \kappa_1^{\mathrm{e}} \|w\|_V^2$ holds for all w in V_0) already provides the ellipticity with the constant

$$\kappa_1 := (1-\xi)\,\kappa_1^{\mathrm{e}}.$$

We show the boundedness (57). The Cauchy-Schwarz inequality reads

$$(D\,J_k)^o\,(v\,;\,w,\overline{w}) \leq \|\varepsilon(w) - \tilde{p_k}^o(\varepsilon(v)\,;\,\varepsilon(w))\|_{\mathbb{C}} \|\varepsilon(\overline{w})\|_{\mathbb{C}} \quad \forall w,\overline{w} \in V_0\,. \tag{59}$$

Then the triangle inequality and the contractivity of $\tilde{p_k}^o$ provide

$$(D\,J_k)^o\,(v\,;\,w,\overline{w}) \leq (1+\xi)\,\|\varepsilon(w)\|_{\mathbb{C}} \|\varepsilon(\overline{w})\|_{\mathbb{C}} \quad \forall w,\overline{w} \in V_0. \tag{60}$$

It is well know from the theory of linear elasticity, that there exists a constant κ_2^{e}, which satisfies $\|\varepsilon(w)\|_{\mathbb{C}} \|\varepsilon(\overline{w})\|_{\mathbb{C}} \leq \kappa_2^{\mathrm{e}} \|w\|_V \|\overline{w}\|_V$. Thus, (57) holds with

$$\kappa_2 = (1+\xi)\,\kappa_2^{\mathrm{e}}. \tag{61}$$

□

Remark 4. By exploiting the structure of the slanting function $\tilde{p_k}^o(\varepsilon(v);\varepsilon(w))$ the boundedness constant κ_2 from (61) can be further improved to

$$\kappa_2 = \kappa_2^{\mathrm{e}}. \tag{62}$$

Let us check that for all $w \in V_0$ there holds a. e. in $\Omega_k^p(v)$:

$$\begin{aligned}
\|\tilde{p_k}^o(\varepsilon(v)\,;\,\varepsilon(w))\|_F^2 &= \xi^2 \left(\beta_k^2 \|\operatorname{dev}\varepsilon(w)\|_F^2 + (1+\beta_k)(1-\beta_k) \frac{\langle \operatorname{dev}\tilde{\sigma}_k\,,\,\operatorname{dev}\varepsilon(w)\rangle_F^2}{\|\operatorname{dev}\tilde{\sigma}_k\|_F^2} \right) \\
&\le \xi \left(\beta_k \|\operatorname{dev}\varepsilon(w)\|_F^2 + 2(1-\beta_k) \frac{\langle \operatorname{dev}\tilde{\sigma}_k\,,\,\operatorname{dev}\varepsilon(w)\rangle_F^2}{\|\operatorname{dev}\tilde{\sigma}_k\|_F^2} \right) \\
&\le 2\xi \left(\beta_k \|\operatorname{dev}\varepsilon(w)\|_F^2 + (1-\beta_k) \frac{\langle \operatorname{dev}\tilde{\sigma}_k\,,\,\operatorname{dev}\varepsilon(w)\rangle_F^2}{\|\operatorname{dev}\tilde{\sigma}_k\|_F^2} \right) \\
&= 2\langle \operatorname{dev}\varepsilon(w)\,,\,\tilde{p_k}^o(\varepsilon(v)\,;\,\varepsilon(w))\rangle_F\,.
\end{aligned}$$

This inequality holds trivially a. e. in $\Omega_k^e(v)$, where $\tilde{p_k}^o(\varepsilon(v)\,;\,\cdot) \equiv 0$. Using the scalar product $\langle \circ\,,\,\diamond\rangle_Q = \int_\Omega \langle \circ\,,\,\diamond\rangle_F \,\mathrm{d}x$, we obtain

$$\|\tilde{p_k}^o(\varepsilon(v)\,;\,\varepsilon(w))\|_Q^2 \le 2\langle \operatorname{dev}\varepsilon(w)\,,\,\tilde{p_k}^o(\varepsilon(v)\,;\,\varepsilon(w))\rangle_Q,$$

which is equivalent thanks to Lemma 3 to

$$\|\tilde{p_k}^o(\varepsilon(v)\,;\,\varepsilon(w))\|_{\mathbb{C}}^2 \le 2\langle \varepsilon(w)\,,\,\tilde{p_k}^o(\varepsilon(v)\,;\,\varepsilon(w))\rangle_{\mathbb{C}}. \tag{63}$$

Due to (63), there holds $\|\varepsilon(w) - \tilde{p_k}^o(\varepsilon(v)\,;\,\varepsilon(w))\|_{\mathbb{C}}^2 \le \|\varepsilon(w)\|_{\mathbb{C}}^2$, which applied to the inequality (59) improves the inequality (60) and provides the sharper constant (62).

Proposition 2. *Let $k \in \{1,\dots,N_\Theta\}$ be fixed and the assumptions of Corollary 2 be fulfilled. Let the mapping $D\,J_k : V_D \to {V_0}^*$ be defined $D\,J_k(v) := D\,J_k(v\,;\,\circ)$ as in (29), and $(D\,J_k)^o : V_D \to \mathcal{L}(V_0, {V_0}^*)$ be defined $(D\,J_k)^o\,(v) := (D\,J_k)^o\,(v\,;\,\diamond,\circ)$ as in (50). Then, the Newton-like iteration*

$$v^{j+1} = v^j - \left[(D\,J_k)^o\,(v^j)\right]^{-1} D\,J_k(v^j)$$

converges superlinearly to the solution u_k of Problem 3, provided that $\|v^0 - u_k\|_V$ is sufficiently small.

Proof. We check the assumptions of Theorem 4 for the choice $F = D\,J_k$. Let $v \in V_D$ be arbitrarily fixed. The mapping $(D\,J_k)^o\,(v) : V_0 \to {V_0}^*$ serves as a slanting function for $D\,J_k$ at v. Moreover, $(D\,J_k)^o\,(v) : V_0 \to {V_0}^*$ is bijective if and only if there exists a unique element w in V_0 such, that for arbitrary but fixed $f \in {V_0}^*$ there holds

$$(D\,J_k)^o\,(v\,;\,w,\overline{w}) = f(\overline{w}) \quad \forall \overline{w} \in V_0\,. \tag{64}$$

Since the bilinear form $(D\,J_k)^o\,(v)$ is elliptic and bounded (Lemma 2), we apply the Lax-Milgram Theorem to ensure the existence of a unique solution to (64). Finally, the uniform boundedness of $[(D\,J_k)^o\,(\cdot)]^{-1}$ follows from the estimate

$$\begin{aligned}\|[(D\,J_k)^o\,(v)]^{-1}\|_{L({V_0}^*,V_0)} &= \sup_{w^*\in {V_0}^*} \frac{\|[(D\,J_k)^o\,(v)]^{-1}\;w^*\|_V}{\|w^*\|_{{V_0}^*}} = \sup_{w\in V_0} \frac{\|w\|_V}{\|(D\,J_k)^o\,(v\,;\,w,\cdot)\|_{{V_0}^*}} \\ &= \sup_{w\in V_0}\inf_{\overline{w}\in V_0} \frac{\|w\|_V\|\overline{w}\|_V}{|(D\,J_k)^o\,(v\,;\,w,\overline{w})|} \le \sup_{w\in V_0} \frac{\|w\|_V^2}{|(D\,J_k)^o\,(v\,;\,w,w)|} \le \frac{1}{\kappa_1}\,,\end{aligned}$$

with κ_1 denoting the v-independent ellipticity constant from Lemma 2. $\square$

5 Spatial Discretization

We decompose the domain Ω by a shape-regular triangulation $\mathcal{T} = \{T \text{ open} \subset \Omega\}$, such that $\overline{\bigcup T} = \overline{\Omega}$ and $\bigcap T = \emptyset$. We approximate the infinite-dimensional space V by the finite-dimensional subspace $V_h := \{u_h \in V \mid u_h \in C^1(T)^3\ \forall T \in \mathcal{T}\}$ and define $V_{hD} := V_h \cup V_D$ and $V_{h0} := V_h \cup V_0$. Analogous results to Corollary 2 and Proposition 2 can be shown for the finite-dimensional subspace V_h without any additional assumptions:

Corollary 3. *Let $k \in \{1,\dots,N_\Theta\}$ and $v_h \in V_{hD}$ be arbitrarily fixed. Let $DJ_k : V_{hD} \to {V_{h0}}^*$ and $\tilde{p}_k : C(T)^{3\times3}_{sym} \to C(T)^{3\times3}_{sym}$ for all $T \in \mathcal{T}$ be defined as in* (29) *and* (43)*. Then, DJ_k is slantly differentiable at v_h and $\tilde{p}_k$ is slantly differentiable at $\varepsilon(v_h)$ with the slanting functions*

$$(D\,J_k)^o\,(v_h\,;\,w_h,\overline{w}_h) = \sum_{T\in\mathcal{T}} \int_T \mathbb{C}\left(\varepsilon(w_h) - \tilde{p_k}^o(\varepsilon(v_h)\,;\,\varepsilon(w_h))\right) : \varepsilon(\overline{w}_h)\,\mathrm{d}x\,, \qquad (65)$$

$$\tilde{p_k}^o(\varepsilon(v_h)\,;\,\varepsilon(w_h)) = \begin{cases} \xi\left(\beta_k\,\operatorname{dev}\varepsilon(w_h) + (1-\beta_k)\,\frac{\langle \operatorname{dev}\tilde{\sigma}_k\,,\operatorname{dev}\varepsilon(w_h)\rangle_F}{\|\operatorname{dev}\tilde{\sigma}_k\|_F^2}\operatorname{dev}\tilde{\sigma}_k\right) & \textit{in } \Omega_k^p(v_h)\,,\\ 0 & \textit{else}\,,\end{cases}$$

for all $w_h, \overline{w}_h \in V_{h0}$. Herein the abbreviations (51) *together with the definitions* (42) *are used.*

Proof. The result follows due to the piecewise continuously differentiability of v_h, which implies that $\operatorname{dev}\tilde{\sigma}_k(\varepsilon(v_h))$ and $\phi_{k-1}(\operatorname{dev}\tilde{\sigma}_k(\varepsilon(v_h)))$ in (42) are continuous mappings, and thus Theorem 5 is applicable (where we choose $\gamma = 0$). $\square$

Proposition 3. *Let $k \in \{1,\dots,N_\Theta\}$ be fixed and the assumptions of Corollary 2 be fulfilled. Let the mapping $D\,J_k : V_{hD} \to {V_{h0}}^*$ be defined $D\,J_k(v_h) := D\,J_k(v_h\,;\,\circ)$ as in (29), and $(D\,J_k)^o : V_{hD} \to \mathcal{L}(V_{h0}, {V_{h0}}^*)$ be defined $(D\,J_k)^o\,(v_h) := (D\,J_k)^o\,(v_h\,;\,\diamond,\circ)$ as in (65). Then, the Newton-like iteration*

$$v_h^{j+1} = v_h^j - \left[(D\,J_k)^o\,(v_h^j)\right]^{-1} D\,J_k(v_h^j)$$

converges superlinearly to the solution u_{hk} of Problem 3, provided that $\|v_h^0 - u_{hk}\|_V$ is sufficiently small.

Proof. The proof is analogous to the proof of Proposition 2 since V_h is a subspace of V. □

5.1 Vector Representation

This subsection is based on [ACFK02]. Here we consider the 2D case only. The additional information about the implementation including Matlab code can be found in [GV06].

We approximate the possibly non-polygonal 2D domain Ω by a polygonal 2D domain Ω' with the boundary Γ' split into the approximated Dirichlet and Neumann part Γ'_D and Γ'_N. Let $\mathcal{T} = \{T \text{ open} \subset \Omega'\}$ be a shape-regular triangulation of Ω', where all T are triangles, $\mathcal{E} = \{E\}$ be a set of edges and $\mathcal{E}_N = \mathcal{E} \cap \Gamma'_N$ be its intersection with the approximated Neumann boundary Γ'_N. The vertices of all triangles are collected in the set $\mathcal{N} = \{\mathbf{x} \in \mathbb{R}^2 \mid \exists T \in \mathcal{T} : \mathbf{x} \text{ is vertex of } T\}$. Let $\varphi_i : \Omega' \to \mathbb{R}$ be affine linear on each element $T \in \mathcal{T}$ such that for an arbitrary node $\mathbf{x}_l$ the condition $\varphi_i(\mathbf{x}_l) = \delta_{il}$ is satisfied for all $i, l \in \{1, \ldots, |\mathcal{N}|\}$. Further, let e_j denote the j-th unit vector. Then, u_h can be expressed by $u_h(x) := \sum_{i,j} u_{i,j}\varphi_i(x)e_j$, where $u_{i,j} := (u(\mathbf{x}_i))_j$, or for short, we can write $u_h(x) = \Phi(x)^T \mathbf{u}$ by defining $\Phi(x) := (\varphi_i(x)\, e_j)_{i\in\{1,\ldots,|\mathcal{N}|\}, j\in\{1,2\}} \in \mathbb{R}^{2|\mathcal{N}|}$ and $\mathbf{u} := (u_{i,j})_{i\in\{1,\ldots,|\mathcal{N}|\}, j\in\{1,2\}} \in \mathbb{R}^{2|\mathcal{N}|}$. We then lead the infinite-dimensional space V into a finite-dimensional subspace $V_h := \{u_h \in V \mid u_h = \Phi^T\mathbf{u},\ \mathbf{u} \in \mathbb{R}^{2|\mathcal{N}|}\}$.

We consider the domain Ω to be thin with respect to one of the three space dimensions. Thus, the strain ε (plain strain model) or the stress σ (plain stress model) have zero components in that direction. The following formulations hold for the plain strain model only, the plain stress model can be expressed analogously. We assume the total strain ε, the plastic strain p and the stress tensor σ in forms

$$\varepsilon = \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & 0 \\ \varepsilon_{12} & \varepsilon_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad p = \begin{pmatrix} p_{11} & p_{12} & 0 \\ p_{12} & p_{22} & 0 \\ 0 & 0 & p_{33} \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & 0 \\ \sigma_{12} & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix}.$$

The certain structure of ε in the plain strain case implies the certain structure of the plastic strain p by the application of the minimizer property (43). Moreover, the structure of the stress σ in the plain strain case follows from Hook's Law (4). it is sufficient to store the information about ε, p and σ in the vectors $\boldsymbol{\varepsilon} := (\varepsilon_{11}, \varepsilon_{22}, 2\varepsilon_{12})^T$, $\mathbf{p} := (p_{11}, p_{22}, p_{12})^T$ and $\boldsymbol{\sigma} := (\sigma_{11}, \sigma_{22}, \sigma_{12})^T$. Analogous operations in tensor and vector representation, such as norms, traces and deviators, are summarized in Table 1. It follows $\langle \sigma_\varepsilon\, , \, \varepsilon \rangle_F = (\boldsymbol{\sigma}_\varepsilon)^T \boldsymbol{\varepsilon}$ and $\langle \sigma_p\, , \, \varepsilon \rangle_F = (\boldsymbol{\sigma}_p)^T \boldsymbol{\varepsilon}$. Let R_T and R_E be operators which restrict the global vector $\mathbf{u}$ onto a local element T by

$$\mathbf{u}_T = R_T\mathbf{u}\,, \quad \mathbf{u}_E = R_E\mathbf{u}\,. \tag{66}$$

Let the fixed triangle $T \in \mathcal{T}$ have the vertices $(\mathbf{x}_\alpha, \mathbf{x}_\beta, \mathbf{x}_\gamma)$ with the coordinates

$$((x_{\alpha,1}, x_{\alpha,2}), (x_{\beta,1}, x_{\beta,2}), (x_{\gamma,1}, x_{\gamma,2}))\,.$$

Common (Tensor) Representation	Vector Representation
$\varepsilon := \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & 0 \\ \varepsilon_{12} & \varepsilon_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\boldsymbol{\varepsilon} := \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ 2\,\varepsilon_{12} \end{pmatrix}$
$\sigma_\varepsilon := \mathbb{C}\,\varepsilon = \begin{pmatrix} \sigma_{\varepsilon,11} & \sigma_{\varepsilon,12} & 0 \\ \sigma_{\varepsilon,12} & \sigma_{\varepsilon,22} & 0 \\ 0 & 0 & \sigma_{\varepsilon,33} \end{pmatrix}$ with $\quad \mathbb{C}\,\varepsilon = 2\mu\,\varepsilon + \lambda\,\operatorname{tr}\varepsilon\, I$	$\boldsymbol{\sigma}_\varepsilon := \begin{pmatrix} \sigma_{\varepsilon,11} \\ \sigma_{\varepsilon,22} \\ \sigma_{\varepsilon,12} \end{pmatrix} = \underbrace{\begin{pmatrix} \lambda + 2\,\mu & \lambda & 0 \\ \lambda & \lambda + 2\,\mu & 0 \\ 0 & 0 & \mu \end{pmatrix}}_{=:C} \boldsymbol{\varepsilon}\,,$ $\sigma_{\varepsilon,33} = \underbrace{\frac{\lambda}{2\,(\lambda+\mu)}}_{=\nu} \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \boldsymbol{\sigma}_\varepsilon, \;\; \operatorname{tr}\sigma_\varepsilon = \frac{\nu+1}{\nu}\,\sigma_{\varepsilon,33}$
$\operatorname{dev}\sigma_\varepsilon = \sigma_\varepsilon - \frac{\operatorname{tr}\sigma_\varepsilon}{\operatorname{tr} I}\, I$	$\mathbf{dev}\boldsymbol{\sigma}_\varepsilon := \begin{pmatrix} (\operatorname{dev}\sigma_\varepsilon)_{11} \\ (\operatorname{dev}\sigma_\varepsilon)_{22} \\ (\operatorname{dev}\sigma_\varepsilon)_{12} \end{pmatrix} = \boldsymbol{\sigma}_\varepsilon - \frac{\operatorname{tr}\sigma_\varepsilon}{\operatorname{tr} I} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$ thus, $\;\mathbf{dev}\boldsymbol{\sigma}_\varepsilon = \underbrace{\left(I - \frac{\nu+1}{\dim(\sigma_\varepsilon)} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right)}_{=:K} \boldsymbol{\sigma}_\varepsilon$
$p = \begin{pmatrix} p_{11} & p_{12} & 0 \\ p_{12} & p_{22} & 0 \\ 0 & 0 & -\,(p_{11}+p_{22}) \end{pmatrix}$	$\mathbf{p} := \begin{pmatrix} p_{11} \\ p_{22} \\ p_{12} \end{pmatrix}, \;\; \lVert\mathbf{p}\rVert_N^2 := \mathbf{p}^T \underbrace{\begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}}_{=:N} \mathbf{p}\,,$ then: $\quad \lVert\mathbf{p}\rVert_N = \lVert p\rVert_F$
$\sigma_p := \mathbb{C}\,p = \begin{pmatrix} \sigma_{\varepsilon,11} & \sigma_{\varepsilon,12} & 0 \\ \sigma_{\varepsilon,12} & \sigma_{\varepsilon,22} & 0 \\ 0 & 0 & \sigma_{\varepsilon,33} \end{pmatrix}$ with $\quad \mathbb{C}\,p = 2\,\mu\,p + \lambda\,\underbrace{\operatorname{tr} p}_{=0}\, I = 2\,\mu\,p$	$\boldsymbol{\sigma}_p := \begin{pmatrix} \sigma_{p,11} \\ \sigma_{p,22} \\ \sigma_{p,12} \end{pmatrix} = 2\,\mu\,\mathbf{p}$ and $\quad \sigma_{p,33} = -\begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \boldsymbol{\sigma}_p$
$\sigma = \mathbb{C}\,(\varepsilon - p) = \sigma_\varepsilon - \sigma_p$	$\boldsymbol{\sigma} = \boldsymbol{\sigma}_\varepsilon - \boldsymbol{\sigma}_p$ and $\sigma_{33} = \sigma_{\varepsilon,33} - \sigma_{p,33}$
$\operatorname{dev}\sigma = \operatorname{dev}\sigma_\varepsilon - \underbrace{\operatorname{dev}\sigma_p}_{=\sigma_p},$ $\lVert\operatorname{dev}\sigma\rVert_F^2 = \sum_{i,j} (\operatorname{dev}\sigma)_{ij}^2$	$\mathbf{dev}\boldsymbol{\sigma} = \mathbf{dev}\boldsymbol{\sigma}_\varepsilon - \boldsymbol{\sigma}_p\,, \;\; \lVert\operatorname{dev}\sigma\rVert_F = \lVert\mathbf{dev}\boldsymbol{\sigma}\rVert_N\,,$ $(\operatorname{dev}\sigma)_{33} = -\begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \mathbf{dev}\boldsymbol{\sigma}$

Table 1: Table of Vector Representation regarding the Plain Strain Model.

Then $\boldsymbol{\varepsilon}(u_h)$ can be calculated on T by

$$\boldsymbol{\varepsilon}(u_h)(\mathbf{x})_{|T} = \begin{pmatrix} \partial_1\varphi_\alpha & 0 & \partial_1\varphi_\beta & 0 & \partial_1\varphi_\gamma & 0 \\ 0 & \partial_2\varphi_\alpha & 0 & \partial_2\varphi_\beta & 0 & \partial_2\varphi_\gamma \\ \partial_2\varphi_\alpha & \partial_1\varphi_\alpha & \partial_2\varphi_\beta & \partial_1\varphi_\beta & \partial_2\varphi_\gamma & \partial_1\varphi_\gamma \end{pmatrix} \begin{pmatrix} u_{\alpha,1} \\ u_{\alpha,2} \\ u_{\beta,1} \\ u_{\beta,2} \\ u_{\gamma,1} \\ u_{\gamma,2} \end{pmatrix},$$

or in a more compact way,

$$\boldsymbol{\varepsilon}(u_h)(x)_{|T} = B\,\mathbf{u}_T\,, \tag{67}$$

where the partial derivatives of φ_α, φ_β, and φ_γ can be obtained by

$$\nabla \begin{pmatrix} \varphi_\alpha \\ \varphi_\beta \\ \varphi_\gamma \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ x_{\alpha,1} & x_{\beta,1} & x_{\gamma,1} \\ x_{\alpha,2} & x_{\beta,2} & x_{\gamma,2} \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Integration over body and surface forces may be realized by the midpoint rule. We approximate f_k and g_k by $f_T := f_k(\overline{x}_T)$ and $g_E := g_k(\overline{x}_E)$, where $\overline{x}_T$ and $\overline{x}_E$ denote the center of mass of the element T, and the edge E, respectively. Defining

$$\mathbf{f}_T := \frac{|T|}{3}\,R_T^T\,f_T\,, \quad \text{and} \quad \mathbf{g}_E := \frac{|E|}{2}\,R_E^T\,g_E\,,$$

on each $T \in \mathcal{T}$ and on each $E \in \mathcal{E}_N$ there hold

$$\int_T f^T v_h\,\mathrm{d}x \approx \mathbf{f}_T^T\mathbf{v}\,, \quad \text{and} \quad \int_E g^T v_h\,\mathrm{d}s \approx \mathbf{g}_E^T\mathbf{v}\,. \tag{68}$$

5.2 Derivatives and Slanting Functions in Vector Representation

The whole integral over Ω can be split into a sum of integrals on single elements $T \in \mathcal{T}$. Therefore, by combining (66), (67) and (68) we obtain from (29) the discrete formulation of the energy functional's Gâteaux-differential

$$D\,J_k(\mathbf{u}\,;\,\mathbf{v}) := \sum_{T\in\mathcal{T}} \Big[|T|\ \left(C\,B\,\mathbf{u}_T - 2\mu\,\tilde{\mathbf{p}}_k(B\,\mathbf{u}_T)\right)^T B\,R_T - \mathbf{f}_T^T\Big]\,\mathbf{v} - \sum_{E\in\mathcal{E}_N} \mathbf{g}_E^T\,\mathbf{v}$$

with

$$\tilde{\mathbf{p}}_k(B\,\mathbf{u}_T) := \frac{\max\{0, \phi_{k-1}(\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k(B\,\mathbf{u}_T))\}}{2\mu + \sigma_y^2 H^2}\,\frac{\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k(B\,\mathbf{u}_T)}{\|\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k(B\,\mathbf{u}_T)\|_N} + \mathbf{p}_{k-1}\,, \tag{69}$$

where

$$\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k(B\,\mathbf{u}_T) \;:=\; KCB\,\mathbf{u}_T - 2\,\mu\,\mathbf{p}_{k-1}\,, \tag{70}$$

$$\phi_{k-1}(\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k(B\,\mathbf{u}_T)) \;:=\; \|\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k(B\,\mathbf{u}_T)\|_N - \sigma_y(1 + H\alpha_{k-1})\,. \tag{71}$$

Since $D\,J_k(\mathbf{u}\,;\,\mathbf{v})$ is linear in $\mathbf{v}$, there exists the Fréchet-derivative

$$D\,J_k(\mathbf{u}) = \sum_{T\in\mathcal{T}} \Big(|T|\;(CB\,\mathbf{u}_T - 2\mu\,\tilde{\mathbf{p}}_k(B\,\mathbf{u}_T))^T\,B\,R_T - \mathbf{f}_T\Big) - \sum_{E\in\mathcal{E}_N} \mathbf{g}_E\,. \tag{72}$$

Due to Corollary 3, the mapping $D\,J_k$ is slantly differentiable with

$$(D\,J_k)^o\,(\mathbf{u}) = \sum_{T\in\mathcal{T}} |T|\,R_T^T\,B^T\,(C - 2\mu\,\tilde{\mathbf{p}}_k^o(B\,\mathbf{u}_T))^T\,B\,R_T\;,$$

where

$$\tilde{\mathbf{p}}_k^o(B\,\mathbf{u}_T) = \begin{cases} \xi\,\Big((1-\beta_k)\,\frac{\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k^T N}{\|\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k\|_N^2} + \beta_k I\Big)\,KC & \text{if } \phi_k(\tilde{\boldsymbol{\sigma}}_k) > 0\,,\\ 0 & \text{else}\,. \end{cases} \tag{73}$$

serves as a slanting function for $\tilde{\mathbf{p}}_k$ defined in (69). Here, the definitions $\xi := \frac{1}{2\mu+\sigma_y^2 H^2}$ and $\beta_k := \frac{\phi_{k-1}(\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k)}{\|\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k\|_N}$, and the abbreviation $\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k$ for $\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k(B\,\mathbf{u}_T)$ as in (70) are used.

5.3 The Newton-like Method for the Discrete Problem

The Newton-like method is applied for the calculation of $\mathbf{u}\in\mathbb{R}^{d\cdot|\mathcal{N}|}$ such that $DJ_k(\mathbf{u}) = 0$ and $\mathbf{u}$ satisfies the Dirichlet boundary condition:

$$\mathbf{u}_i = \mathbf{u}_{i-1} + \Delta\mathbf{u}_i \qquad (\forall i\in\mathbb{N}), \tag{74}$$

where $\Delta\mathbf{u}_i$ solves

$$-\,(D\,J_k)^o\,(\mathbf{u}_{i-1})\,\Delta\mathbf{u}_i = D\,J_k(\mathbf{u}_{i-1})\;.$$

Note, that $\mathbf{u}_i$ must satisfy (generally inhomogeneous) Dirichlet boundary conditions for all $i\in\mathbb{N}$. Therefore, it is sufficient for the initial approximation $\mathbf{u}_0$ to satisfy the inhomogeneous Dirichlet conditions, and for $\Delta\mathbf{u}_i$ to solve the homogeneous Dirichlet conditions. For the termination of the Newton-like method we check, if

$$\frac{|u_{h,i} - u_{h,i-1}|_\varepsilon}{|u_{h,i}|_\varepsilon + |u_{h,i-1}|_\varepsilon} \tag{75}$$

with $|\cdot|_\varepsilon := (\int_\Omega \|\varepsilon(\cdot)\|_F^2\,\mathrm{d}x)^{1/2}$, is smaller than a given prescribed bound $\epsilon > 0$.

6 Numerical Examples

The following tests were calculated on a computer with 2.4 GHz CPU, 2 GB RAM using Matlab© version 7.0 on Linux OS. We define 'DOF' as the short form of *degrees of freedom*, and 'VPZ' to be the short form of *variation in plastic zones* which is calculated as follows: In the i-th iteration step, the vector $\mathbf{w}^i$ stores the information about which elements are plastic and which are not by defining its components $w_j^i := 1$ if T_j is deformed plastically

$(\phi_{k-1}(\mathbf{dev}\tilde{\boldsymbol{\sigma}}_k(B\,\mathbf{u}_T)) > 0)$, and $w_j^i := 1$ else. Let the starting vector $\mathbf{w}^0 = 0$. Variation in plastic zones VPZ_{i-1}^i from the $(i-1)$-st to the i-th iteration step is defined by

$$\mathrm{VPZ}_{i-1}^i = \frac{100}{|\mathcal{T}|} \sum_{j=1}^{|\mathcal{T}|} |w_j^i - w_j^{i-1}|. \tag{76}$$

At all numerical examples, the termination bound $\epsilon = 1e-12$ was used.

Example 1 (L-Shape)**.** This example is taken from [ACFK02] and its geometry and the coarse grid triangulation are displayed in Figure 2. We assume non-homogeneous Dirichlet boundary conditions in polar coordinates r, θ

$$\begin{aligned} u_r(r,\theta) =& \frac{1}{2\mu} r^\alpha \left[-(\alpha+1)\cos((\alpha+1)\,\theta) + (C_2 - (\alpha+1))\,C_1 \cos((\alpha-1)\,\theta) \right], \\ u_\theta(r,\theta) =& \frac{1}{2\mu} r^\alpha \left[(\alpha+1)\sin((\alpha+1)\,\theta) + (C_2 + (\alpha-1))\,C_1 \sin((\alpha-1)\,\theta) \right]. \end{aligned} \tag{77}$$

The critical exponent $\alpha \approx 0.544483737$ is the solution to the equation

$$\alpha \sin(2\omega) + \sin(2\omega\alpha) = 0$$

with $C_1 = -(\cos((\alpha+1)\omega))/\cos((\alpha-1)\omega)$, $C_2 = (2(\lambda+2\mu))/(\lambda+\mu)$ and $\omega = \frac{3\pi}{4}$. It can be shown that the formulae (77) describe the solution to the purely elastic problem with the same non-homogeneous Dirichlet boundary conditions also in the interior of the Lshape domain. Thus there is a strain-singularity in the reentrant corner, which can also be expected for the elastoplastic case. The material parameters are defined as

$$E = 1e5,\ \nu = 0.3,\ \sigma_Y = 2.2,\ H = 1.$$

Figure 3 shows the yield function (right) and the elastoplastic zones (left), where purely elastic zones are colored green (light gray in case of a non-color print respectively), and elastoplastic zones are colored pink (dark grey respectively). The domain's displacement is multiplied by factor $3e3$. Table 2 reports on convergence behavior of Newton-like method for graduated uniform meshes.

Example 2 (Wrench)**.** This example simulates the deformation of a screw-wrench under pressure. Problem geometry is shown in Figure 4: A screw-wrench *sticks* on a screw (homogeneous Dirichlet boundary condition) and a surface load g is applied to a part of the wrench's handhold in interior normal direction (Neumann boundary condition, cf. 4). The material parameters are set

$$E = 2e8\,,\ \nu = 0.3\,,\ \sigma_Y = 2e6\,,\ H = 0.001$$

and the traction intensity amounts $|g| = 6e4$. Figure 5 shows the yield function (right) and the elastoplastic zones (left), where purely elastic zones are colored green (light gray

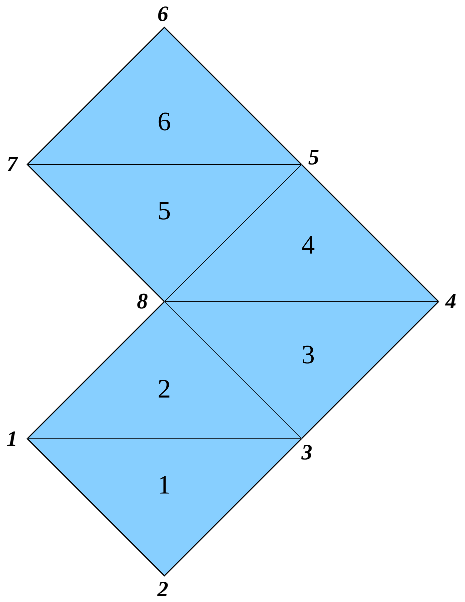

Figure 2: Problem geometry and the coarse triangulation of Example 1. The L-shape domain Ω is described by the polygon $(-1,-1)$, $(0,-2)$, $(2,0)$, $(0,2)$, $(-1,1)$, $(0,0)$.

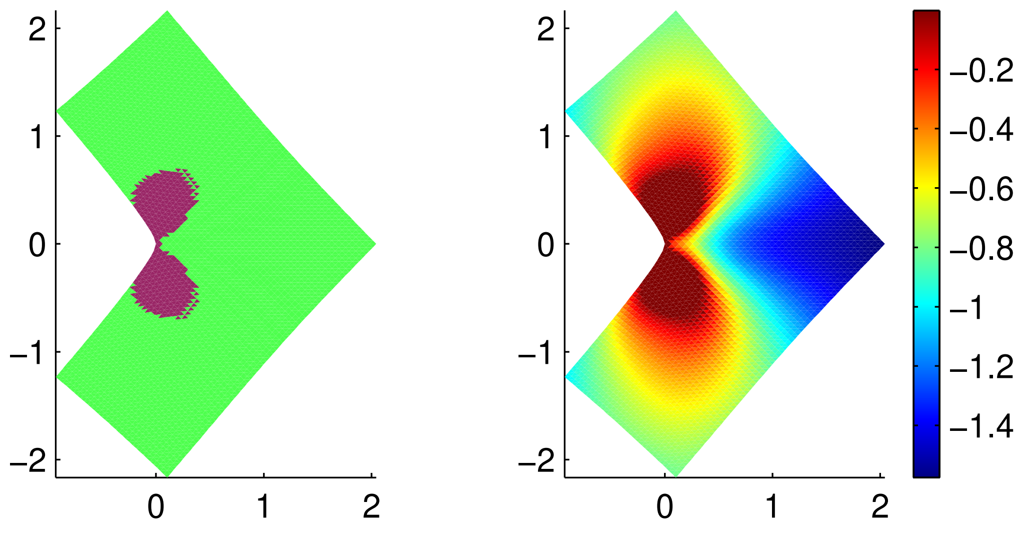

Figure 3: Elastoplastic zones (left) and yield function (right) of the deformed domain in Example 1. The displacement is magnified by factor $3e3$.

DOF	10	66	...	20466	97282	391170
relative error:						
step 1	2.8383e-02	3.9827e-02	...	7.2243e-02	7.0236e-02	6.8321e-02
step 2	1.0467e-04	1.2352e-03	...	1.1004e-02	1.1063e-02	1.1022e-02
step 3	2.3781e-09	6.1409e-07	...	1.1453e-03	1.2746e-03	1.3552e-03
step 4	1.0944e-16	2.9589e-13	...	2.0826e-05	4.0743e-05	5.9611e-05
step 5			...	6.8005e-09	5.1957e-08	2.0693e-07
step 6			...	5.2211e-15	1.3866e-13	4.3361e-12
step 7			...			1.8774e-14
VPZ (%):						
step 0-1	16.67	10.42	...	10.59	10.61	10.62
step 1-2	0	2.083	...	2.873	2.816	2.752
step 2-3	0	0	...	0.2686	0.2218	0.1638
step 3-4	0	0	...	0.04069	0.02848	0.01882
step 4-5			...	0	0	0
step 5-6			...	0	0	0
step 6-7			...			0
Time (sec.)	2.00537	2.25042	...	142.29	590.106	2692.87

Table 2: Convergence table in Example 1 **(Lshape)**. The table displays the relative error in displacements (75) and the variation of plastic zones (VPZ) per iteration step for various uniformly refined meshes.

Level	0	1	...	5	6	7
DOF	60	202	...	41662	165246	658174
relative error:						
step 1	2.3834e-14	3.6169e-03	...	1.3194e-01	1.4872e-01	1.5846e-01
step 2		2.3598e-06	...	5.6966e-02	6.9302e-02	7.9603e-02
step 3		1.5324e-11	...	7.5805e-03	1.3223e-02	2.9909e-02
step 4		4.5752e-15	...	4.0307e-04	2.4344e-03	3.5626e-03
step 5			...	5.9665e-06	2.1840e-04	1.2013e-04
step 6			...	2.9485e-10	1.5089e-05	1.0364e-05
step 7			...	7.8696e-14	3.8914e-09	1.1642e-09
step 8			...		1.5508e-13	2.9988e-13
VPZ (%):						
step 0-1	0	1.25	...	1.819	1.83	1.828
step 1-2		0	...	0.9741	1.168	1.27
step 2-3		0	...	0.3564	0.5591	0.7588
step 3-4		0	...	0.05127	0.1501	0.1418
step 4-5			...	0.002441	0.02563	0.02319
step 5-6			...	0	0.00183	0.004425
step 6-7			...	0	0	0
step 7-8			...		0	0
Time (sec.)	1.31385	2.58625	...	262.304	1177.64	4892

Table 3: Convergence table in Example 2 **(wrench)**. The table displays the relative error in displacements (75) and the variation of plastic zones (VPZ) per iteration step for various uniformly refined meshes.

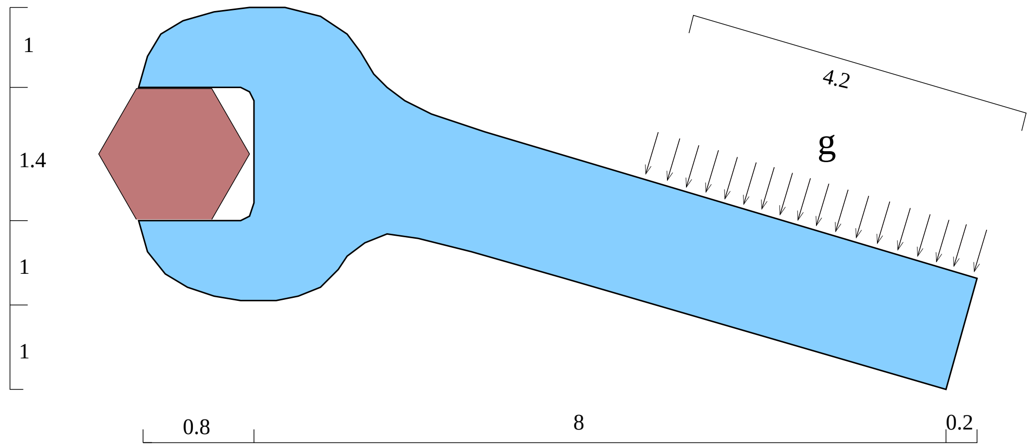

Figure 4: Problem geometry in Example 2.

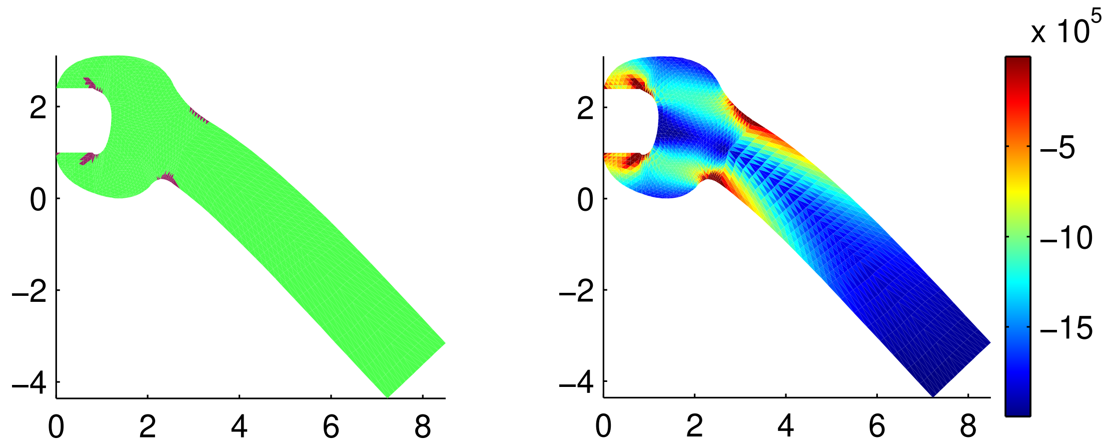

Figure 5: Elastoplastic zones (left) and yield function (right) of the deformed domain in Example 2. The displacement is magnified by factor 10.

Level	0	1	...	3	4	5
DOF	245	940	...	14560	57920	231040
relative error:						
step 1	2.1826e-02	3.5365e-02	...	4.5238e-02	4.6300e-02	4.6603e-02
step 2	2.2225e-03	5.8553e-03	...	8.0839e-03	8.3886e-03	8.5454e-03
step 3	1.0478e-04	1.6539e-04	...	3.4440e-04	4.0032e-04	4.1602e-04
step 4	1.4404e-08	3.9755e-08	...	1.5206e-05	1.2050e-05	1.3944e-05
step 5	7.2634e-16	6.9728e-15	...	2.4947e-07	7.2972e-07	3.2631e-07
step 6			...	3.5062e-13	5.3972e-12	1.6473e-12
step 7			...		7.2441e-15	1.4518e-14
VPZ (%):						
step 0-1	4.889	5.889	...	7.042	7.116	7.129
step 1-2	1.333	4.222	...	5.444	5.549	5.546
step 2-3	0.8889	1.222	...	1.056	1.125	1.098
step 3-4	0	0	...	0.1597	0.1233	0.1215
step 4-5	0	0	...	0.01389	0.01042	0.008247
step 5-6			...	0	0	0
step 6-7			...		0	0
Time (sec.)	2	4.6	...	64	286	1195

Table 4: Convergence table in Example 3 **(plate with a hole)**. The convergence table displays the relative error in displacements (75) and the variation of plastic zones VPZ (76) per iteration step for various uniformly refined meshes.

in case of a non-color print respectively), and elastoplastic zones are colored pink (dark grey respectively). The displacement of the domain is multiplied by factor 10. Table 3 reports on the convergence of the Newton-like method for graduated uniform meshes.

Example 3 (Plate with a hole)**.** The example is taken from [Ste03] and serves as a benchmark problem in computational plasticity. In difference to the original problem setup, we choose H to be non-zero, thus hardening effects are considered. The calculation of the original perfect plastic problem can be found in [GV06, GV07]. We consider a thin plate represented by the square $(-10, 10) \times (-10, 10)$ with a circular hole of the radius $r = 1$ in the middle, as can be seen in Figure 6. A surface load g is applied on the plate's upper and lower edge with the intensity $|g| = 450$. Due to the domain's symmetry, only the right upper quarter is discretized. Therefore it is necessary to incorporate homogeneous Dirichlet boundary conditions in the normal direction (gliding conditions) to both symmetry axes. The material parameters are set

$$E = 206900,\ \nu = 0.29,\ \sigma_Y = \sqrt{\frac{2}{3}}\,450,\ H = \frac{1}{2}\,.$$

Figure 7 shows the yield function (right) and the elastic-plastic zones, where purely elastic zones are colored green (light gray in case of a non-color print respectively), and elastic-plastic zones are colored pink (dark grey respectively). The displacement is multiplied by 100. Table 4 reports on the convergence of the Newton-like method.

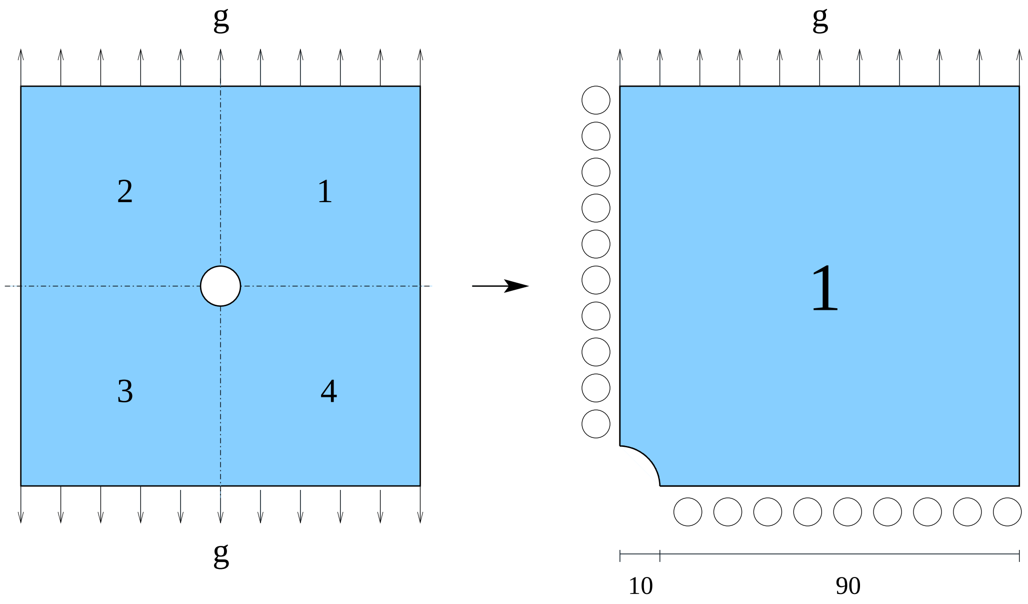

Figure 6: Problem geometry in Example 3.

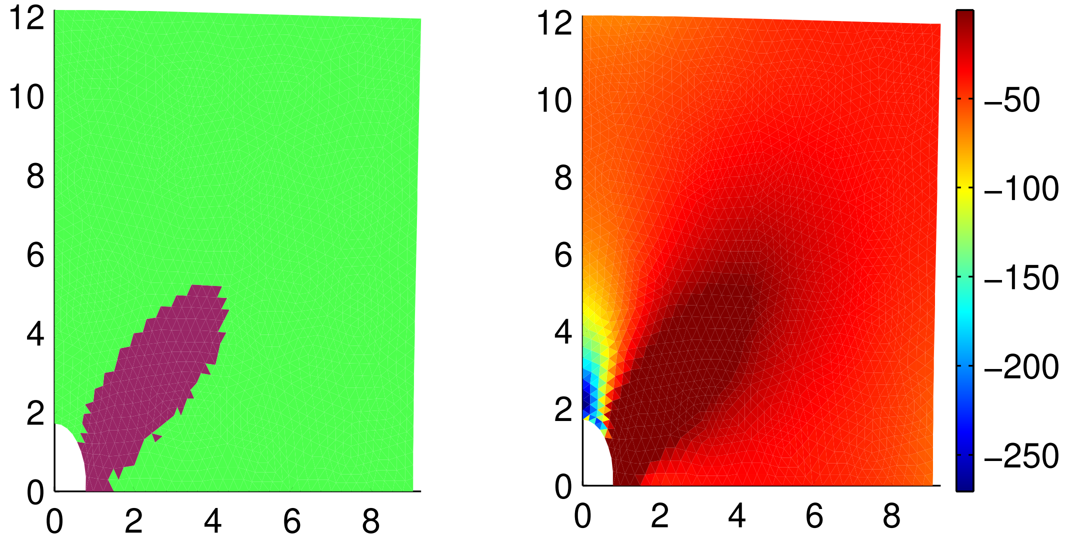

Figure 7: The two plots show the elastoplastic zones (left) and the yield function (right) of the deformed domain in Example 3. The displacement is magnified by the factor 100.

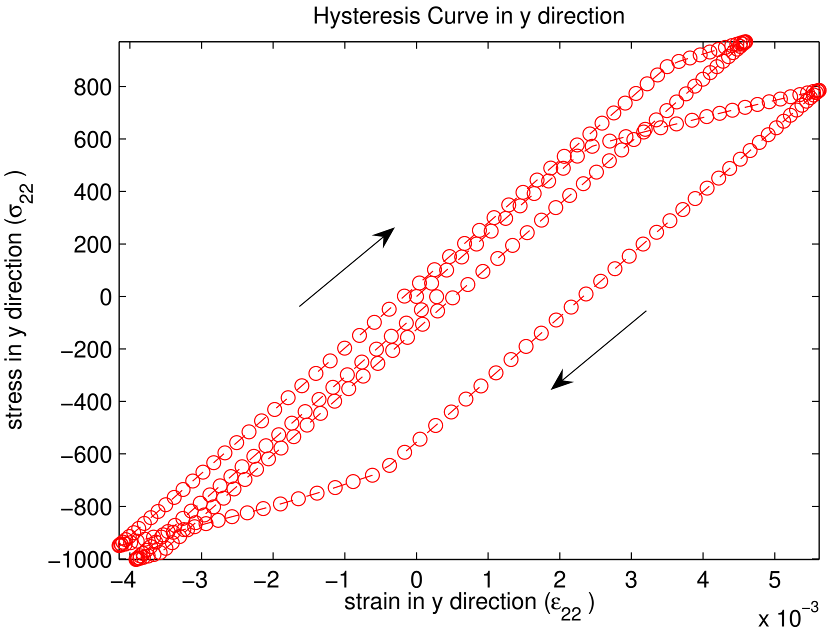

Figure 8: Hysteresis curve for Example 3 with respect to the time dependent surface load $g(t) = (0, \sin(\pi t))$ for $t \in [0, 4]$. At the material point with coordinates roughly (2,2), the stress component σ_{22} is plotted versus the strain component ε_{22}. Both quantifiers are set to zero at $t = 0$. The time development takes place in direction of the arrows.

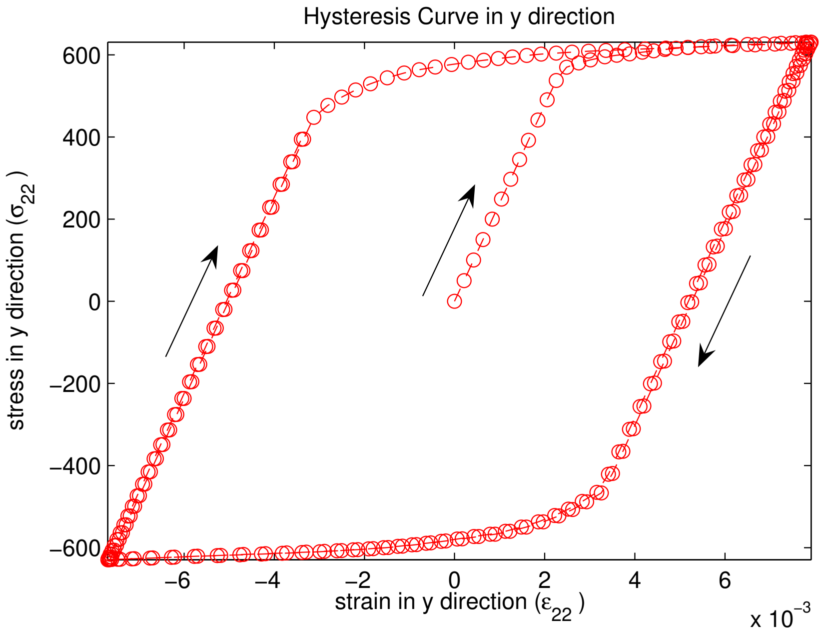

Figure 9: Same hysteresis curve as in Figure 8, except for $H = 0$ which models a perfect plastic material behaviour.

7 Appendix

A few simple properties concerning the deviator are summarized in the following lemma, and extently used throughout this work.

Lemma 3. *Let $n \in \mathbb{N}$, $\lambda \in \mathbb{R}$ with $\lambda > 0$, $\mu \in \mathbb{R}$ with $\mu > 0$, and I denote the identity matrix in $\mathbb{R}^{n\times n}$. Let the mappings* $\operatorname{dev} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ *and* $\mathbb{C} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ *be defined*

$$\operatorname{dev} x := x - \frac{\langle x\,,\, I\rangle_F}{\langle I\,,\, I\rangle_F}\, I\,, \quad \mathbb{C}x := \mu(x + x^T) + \lambda\langle x\,,\, I\rangle_F\, I\,.$$

Then, for all matrices x and y in $\mathbb{R}^{n\times n}$, the the following properties hold:

1. $\langle \operatorname{dev} x\,,\, y\rangle_F = \langle x\,,\, \operatorname{dev} y\rangle_F\,,$
2. $\operatorname{dev} I = 0\,,$
3. $\langle \operatorname{dev} x\,,\, I\rangle_F = 0\,,$
4. $\operatorname{dev}\operatorname{dev} x = \operatorname{dev} x\,,$
5. $\operatorname{dev}\mathbb{C}x = \mu\left(\operatorname{dev} x + \operatorname{dev} x^T\right)\,,$
6. $\mathbb{C}\operatorname{dev} x = \mu\left(\operatorname{dev} x + \operatorname{dev} x^T\right)\,,$
7. $\langle \mathbb{C}x\,,\, I\rangle_F = (2\mu + \langle I\,,\, I\rangle_F\lambda)\langle x\,,\, I\rangle_F\,,$
8. $\langle \mathbb{C}\operatorname{dev} x\,,\, \operatorname{dev} x\rangle_F \le \langle \mathbb{C}x\,,\, x\rangle_F\,.$

Proof. The first and the second property follow from the definition of the deviator:

$$\langle \operatorname{dev} x\,,\, y\rangle_F = \langle x\,,\, y\rangle_F - \frac{\langle x\,,\, I\rangle_F\langle y\,,\, I\rangle_F}{\langle I\,,\, I\rangle_F} = \langle x\,,\, \operatorname{dev} y\rangle_F\,,$$

$$\operatorname{dev} I = I - \frac{\langle I\,,\, I\rangle_F}{\langle I\,,\, I\rangle_F} I = 0\,.$$

The third property follows from the first two properties:

$$\langle \operatorname{dev} x\,,\, I\rangle_F = \langle x\,,\, \operatorname{dev} I\rangle_F = 0\,.$$

The fourth property holds due to the third property:

$$\operatorname{dev}\operatorname{dev} x = \operatorname{dev} x - \frac{\langle \operatorname{dev} x\,,\, I\rangle_F}{\langle I\,,\, I\rangle_F}\, I = \operatorname{dev} x\,.$$

The fifth property relies on the second property,

$$\operatorname{dev}\mathbb{C}x = \mu(\operatorname{dev} x + \operatorname{dev} x^T) + \lambda\langle x\,,\, I\rangle_F\, \operatorname{dev} I = \mu\left(\operatorname{dev} x + \operatorname{dev} x^T\right)\,,$$

and the sixth property relies on the third property,

$$\mathbb{C}\operatorname{dev} x = \mu\left(\operatorname{dev} x + \operatorname{dev} x^T\right) + \lambda\langle \operatorname{dev} x\,,\, I\rangle_F\, I = \mu\left(\operatorname{dev} x + \operatorname{dev} x^T\right)\,.$$

The seventh property follows from the definition of the mapping $\mathbb{C}$:

$$\langle \mathbb{C}x\,,\, I\rangle_F = \mu(\langle x\,,\, I\rangle_F + \langle x^T\,,\, I\rangle_F) + \langle I\,,\, I\rangle_F\lambda\langle x\,,\, I\rangle_F = (2\mu + \langle I\,,\, I\rangle_F\lambda)\langle x\,,\, I\rangle_F\,.$$

The eighth property can be shown by

$$\begin{aligned}\langle \mathbb{C}\operatorname{dev} x\,,\, \operatorname{dev} x\rangle_F &= \langle \operatorname{dev}\mathbb{C}x\,,\, \operatorname{dev} x\rangle_F = \langle \mathbb{C}x\,,\, \operatorname{dev}\operatorname{dev} x\rangle_F = \langle \mathbb{C}x\,,\, \operatorname{dev} x\rangle_F \\ &= \langle \mathbb{C}x\,,\, x\rangle_F - \frac{\langle \mathbb{C}x\,,\, I\rangle_F\langle x\,,\, I\rangle_F}{\langle I\,,\, I\rangle_F} \le \langle \mathbb{C}x\,,\, x\rangle_F\,.\end{aligned}$$

$\square$

The chain rule for slanting functions is provided by the next theorem. Herein we use, that a slantly differentiable function is continuous on a Banach space X, since

$$\lim_{h\to 0}\left(F(x+h)-F(x)\right)=\lim_{h\to 0}\left(F^o(x+h)h+r(h)\right)=0 \qquad \forall x\in X\,,$$

for $\lim_{h\to 0} F^o(x+h)$ is bounded.

Theorem 7. *(chain rule) Let $U\subseteq X$ and $V\subseteq Y$ be open subsets. Let $F\in\mathcal{S}(U;Y)$ such that $F(U)\subseteq V$ and $G\in\mathcal{S}(V;Z)$. Let F^o be a slanting function for F in U and G^o be a slanting function for G in V. Then there holds $G\circ F\in\mathcal{S}(U;Z)$ where*

$$\left(G\circ F\right)^o(x):=G^o(F(x))\,F^o(x) \qquad \forall x\in U$$

serves as a slanting function for $G\circ F$ in U.

Proof. Let $x\in U$ be arbitrary. Since U is open, there exists an open neighborhood $\mathcal{N}\subseteq X$ centered at zero, such that $(x+h)\in U$ if $h\in\mathcal{N}$. The function F is slantly differentiable in U with the slanting function F^o. That is, there exists a mapping $r:X\to Y$ with $\lim_{h\to 0}\frac{\|r(h)\|}{\|h\|}=0$ such that, for all $h\in\mathcal{N}$, there holds

$$F(x+h)=F(x)+F^o(x+h)\,h+r(h)\,.$$

Alike, the function G is slantly differentiable in V with the slanting function G^o. That is, there exists a mapping $s:Y\to Z$ with $\lim_{k\to 0}\frac{\|s(k)\|}{\|k\|}=0$ such that

$$G(y+k)=G(y)+G^o(y+k)\,k+s(k) \tag{78}$$

holds for all $y\in V$ and $k\in Y$, which satisfy $(y+k)\in V$. The certain choice of $y:=F(x)$ and $k(h):=F(x+h)-F(x)=F^o(x+h)\,h+r(h)$ for $h\in\mathcal{N}$ satisfies $y\in V$ and $(y+k(h))\in V$, and yields

$$G(F(x+h))=G(F(x))+G^o(F(x+h))\,F^o(x+h)\,h+t(h)\,,$$

where $t(h):=G^o(F(x+h))\,r(h)+s(k(h))$, due to (78). It remains to show, that

$$\lim_{h\to 0}\frac{\|t(h)\|}{\|h\|}=0\,.$$

Let $\varepsilon>0$ be arbitrary. Since $\lim_{h\to 0}\|F^o(x+h)\|$ is bounded, $\lim_{h\to 0} k(h)=0$, and $\lim_{k\to 0}\frac{\|s(k)\|}{\|k\|}=0$, there holds

$$\lim_{h\to 0}\left(\left(\|F^o(x+h)\|+\varepsilon\right)\frac{\|s(k(h))\|}{\|k(h)\|}\right)=0\,. \tag{79}$$

There exists $\delta>0$, such that for all $h\in\mathcal{N}$ with $\|h\|<\delta$ there holds

$$\left(\|F^o(x+h)\|+\varepsilon\right)\|h\|>\|F^o(x+h)\|\|h\|+\|r(h)\|\geq\|F^o(x+h)\,h+r(h)\|=\|k(h)\|\,.$$

Using this together with (79), we obtain $\lim_{h\to 0} \frac{\|s(k(h))\|}{\|h\|} = 0$. Hence a slantly differentiable function is continuous, the function F is continuous. Thus, the limit $\lim_{h\to 0}\|G^o(F(x+h))\|$ is bounded, and we conclude

$$\lim_{h\to 0} \frac{\|t(h)\|}{\|h\|} \le \lim_{h\to 0} \left(\|G^o(F(x+h))\| \frac{\|r(h)\|}{\|h\|} \right) + \lim_{h\to 0} \left(\frac{\|s(k(h))\|}{\|h\|} \right) = 0\,.$$

□

Acknowledgments

The authors are pleased to acknowledge support by the Austrian Science Fund 'Fonds zur Förderung der wissenschaftlichen Forschung (FWF)' for their support under grant SFB F013/F1306 in Linz, Austria. The idea of looking at the elastoplastic formulation in terms of the Moreau Theorem came out during working progresses with H. Gfrerer, J. Kienesberger and U. Langer. We would also like to thank D. Knees and C. Wieners for an extent overview on literature about regularity results in elastoplasticity.

Let $j \in \mathbb{N}_0$ denote the Newton-like iteration index and $k \in \mathbb{N}_0$ the time index.
Choose a real and strictly decreasing sequence (ϵ_j), which satisfies $\lim_{j\to\infty} \epsilon_j > 0$.
Let $\phi_{k-1}(\sigma) = \|\mathrm{dev}\,\sigma\|_F - \sigma_y(1 + H\alpha_{k-1})$ and $\tilde{\sigma}_k(\varepsilon(v)) = \mathbb{C}\,(\varepsilon(v) - p_{k-1})$.
Let $v_j \in V_D$ be such, that $\phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v_j)) \in L_{2+\epsilon_j}(\Omega)$.
Let $\xi = \frac{2\mu}{2\mu+\sigma_y^2 H^2}$ and $\beta_k = \frac{\phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v)))}{\|\mathrm{dev}\,\tilde{\sigma}_k(\varepsilon(v))\|_F}$.
Then there hold (by the substitution $v = v_j$):
$\tilde{p}_k(\varepsilon(v)) = \frac{1}{2\mu+\sigma_y^2 H^2} \max\{0, \phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v)))\} \frac{\mathrm{dev}\,\tilde{\sigma}_k(\varepsilon(v))}{\|\mathrm{dev}\,\tilde{\sigma}_k(\varepsilon(v))\|_F} + p_{k-1}$,
$D\,J_k(v\,;\,w) = \langle \varepsilon(v) - \tilde{p}_k(\varepsilon(v))\,,\,\varepsilon(w)\rangle_{\mathbb{C}} - \int_\Omega f\cdot w\,\mathrm{d}x - \int_{\Gamma_N} g\cdot w \quad \forall w \in V_0$,
$\tilde{p}_k^o(\varepsilon(v);\varepsilon(w)) = \begin{cases} \xi \left(\beta_k\,\mathrm{dev}\,\varepsilon(w) + (1-\beta_k)\,\frac{\langle \mathrm{dev}\,\tilde{\sigma}_k\,,\,\mathrm{dev}\,\varepsilon(w)\rangle_F}{\|\mathrm{dev}\,\tilde{\sigma}_k\|_F^2}\,\mathrm{dev}\,\tilde{\sigma}_k \right) & \text{in } \Omega_k^p(v)\,, \\ 0 & \text{else}\,, \end{cases}$
$(D\,J_k)^o\,(v\,;\,w_1, w_2) = \langle \varepsilon(w_1) - \tilde{p}_k^o(\varepsilon(v)\,;\,\varepsilon(w_1))\,,\,\varepsilon(w_2)\rangle_{\mathbb{C}} \quad \forall w_1, w_2 \in V_0$.
Notice, that the integrability of p_{k-1} and α_{k-1} underlies the solution $u_{k-1} \in V_D$, for $p_{k-1} = \tilde{p}_{k-1}(\varepsilon(u_{k-1}))$ and $\alpha_{k-1} = \tilde{\alpha}_{k-1}(p_{k-1})$ with $\tilde{\alpha}_k(q) = \alpha_{k-1} + \sigma_y H\|q - p_{k-1}\|_F$.
Using the above defined quantifiers, the $j+1$st Newton-like step reads:
Calculate $v_{j+1} \in V_D$ by solving
$(D\,J_k)^o\,(v_j\,;\,v_{j+1} - v_j, w) = -D\,J_k(v_j\,;\,w)$ for all $w \in V_0$.
The task is now to show, that there holds $\phi_{k-1}(\tilde{\sigma}_k(\varepsilon(v_{j+1}))) \in L_{2+\epsilon_{j+1}}(\Omega)$.

Table 5: A summary of the still open regularity problem, as discussed in subsection 4.3.

References

[AC00] J. Alberty and C. Carstensen, *Numerical analysis of time-depending primal elastoplasticity with hardening*, SIAM J. Numer. Anal. **37** (2000), no. 4, 1271–1294.

[ACFK02] J. Alberty, C. Carstensen, S. A. Funken, and R. Klose, *Matlab implementation of the finite element method in elasticity*, Computing **69** (2002), no. 3, 239–263. MR 1 954 562

[ACZ99] J. Alberty, C. Carstensen, and D. Zarrabi, *Adaptive numerical analysis in primal elastoplasticity with hardening*, Comput. Methods Appl. Mech. Eng. **171** (1999), no. 3-4, 175–204.

[ACZ00] ———, *Numerical analysis in primal elastoplasticity with hardening*, ZAMM, Z. Angew. Math. Mech. (2000), no. 80.

[BF02] A. Bensoussan and J. Frehse, *Regularity results for nonlinear elliptic systems and applications.*, Applied Mathematical Sciences 151, Springer, Berlin, 2002.

[Bla97] R. Blaheta, *Numerical methods in elasto-plasticity*, Comp. Meth. Appl. Mech. Engrg. **147** (1997), 167–185.

[Car97] C. Carstensen, *Domain decomposition for a non-smooth convex minimalization problems and its application to plasticity*, Numerical Linear Algebra with Applications **4** (1997), no. 3, 177–190.

[CNQ01] X. Chen, Z. Nashed, and L. Qi, *Smoothing methods and semismooth methods for nondifferentiable operator equations*, SIAM J. Numer. Anal. **38** (2001), no. 4, 1200–1216.

[DL76] G. Duvaut and Lions J. L., *Numerical analysis of variational inequalities*, Springer-Verlag Berlin Heidelberg, 1976.

[ET99] I. Ekeland and R. Témam, *Convex analysis and variational problems*, SIAM, 1999.

[FS00] M. Fuchs and G. Seregin, *Variational methods for problems from plasticity theory and for generalized Newtonian fluids.*, Lecture Notes in Mathematics 1749, Springer, Berlin, 2000.

[GLT81] R. Glowinskii, J. L. Lions, and R. Trémolières, *Numerical analysis of variational inequalities*, North-Holland, Amsterdam, 1981.

[GV06] P. Gruber and J. Valdman, *Solution of Elastoplastic Problem based on the Moreau-Yosida Theorem*, SFB Report 2006-05, Johannes Kepler University Linz, SFB "Numerical and Symbolic Scientific Computing", 2006.

[GV07] ______, *Implementation of an Elastoplastic Solver based on the MoreauYosida Theorem*, Math. Comput. Simul. (2007).

[HIK02] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth newton method*, SIAM J. Optim. **13** (2002), no. 3, 865–888.

[HR99] W. Han and B. D. Reddy, *Plasticity: Mathematical theory and numerical analysis*, Springer-Verlag New York, 1999.

[Joh76] C. Johnson, *Existence theorems for plasticity problems*, J. math. pures et appl. **55** (1976), 431–444.

[Kie06] J. Kienesberger, *Efficient solution algorithms for elastoplastic problems*, Ph.D. thesis, Johannes-Kepler-Universität Linz, 2006.

[KL84] V. G. Korneev and U. Langer, *Approximate solution of plastic flow theory problems*, Teubner-Texte zur Mathematik, vol. 69, Teubner-Verlag, Leipzig, 1984.

[KLV04] J. Kienesberger, U. Langer, and J. Valdman, *On a robust multigrid-preconditioned solver for incremental plasticity problems*, Proceedings of IMET 2004 - Iterative Methods, Preconditioning & Numerical PDEs, 2004, pp. 84–87.

[Mor65] J. J. Moreau, *Proximité et dualité dans un espace hilbertien*, Bulletin de la Société Mathématique de France **93** (1965), 273–299.

[SH98] J. C. Simo and T. J. R. Hughes, *Computational inelasticity*, Springer-Verlag New York, 1998.

[Ste03] E. Stein, *Error-controlled adaptive finite elements in solid mechanics*, Wiley, Chichester, 2003.

[Ulb03] M. Ulbrich, *Semismooth Newton methods for operator equations in function spaces.*, SIAM J. Optim. **13** (2003), no. 3, 805–841.

[Wie00] C. Wieners, *Multigrid methods for finite elements and the application to solid mechanics. Theorie und Numerik der Prandtl-Reuß Plastizität.*, 2000, Habilitationsschrift.

[Wie06] ______, *Nonlinear solution methods for infinitesimal perfect plasticity*, Tech. report, IWRMM-06/11, Universität Karlsruhe, 2006.