Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Duality results and regularization schemes for Prandtl–Reuss perfect plasticity

Michael Hintermüller[1,2], Simon Rösel[2]

submitted: February 23, 2017

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: michael.hintermueller@wias-berlin.de

[2] Department of Mathematics
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany
E-Mail: hint@math.hu-berlin.de
roesel@math.hu-berlin.de

No. 2376
Berlin 2017

2010 *Mathematics Subject Classification.* 74C05, 49M15, 49K20, 49M29.

Key words and phrases. Perfect plasticity, Prandtl–Reuss plasticity, small-strain, Fenchel duality, semismooth Newton.

This research was carried out in the framework of MATHEON supported by the Einstein Foundation Berlin within the ECMath projects OT1, SE5 and SE15 as well as project A-AP24. The authors further gratefully acknowledge the support of the DFG through the DFG-SPP 1962: Priority Programme "Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization“ within Projects 10, 11 and 13.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Duality results and regularization schemes for Prandtl–Reuss perfect plasticity

Michael Hintermüller, Simon Rösel

ABSTRACT. We consider the time-discretized problem of the quasi-static evolution problem in perfect plasticity posed in a non-reflexive Banach space and we derive an equivalent version in a reflexive Banach space. A primal-dual stabilization scheme is shown to be consistent with the initial problem. As a consequence, not only stresses, but also displacement and strains are shown to converge to a solution of the original problem in a suitable topology. This scheme gives rise to a well-defined Fenchel dual problem which is a modification of the usual stress problem in perfect plasticity. The dual problem has a simpler structure and turns out to be well-suited for numerical purposes. For the corresponding subproblems an efficient algorithmic approach in the infinite-dimensional setting based on the semismooth Newton method is proposed.

1. INTRODUCTION

The foundation of the mathematical analysis of the time-dependent problem of quasi-static small strain associative perfect plasticity or *Prandtl–Reuss plasticity* has its origin in [18, 34], where the latter reference contains the first existence result for the time-dependent case, which is extended in [39] to yield criteria varying in time. The fundamental difference to hardening plasticity lies in the possible presence of strain localization [36, 42]. On the mathematical level, this physical phenomenon entails that displacements may display discontinuities along surfaces, which necessitates a different functional analytic setting; cf. [53]. This framework essentially corresponds to the static problem usually referred to as *Hencky plasticity*; see [52, 54]. In the static case, existence results for the primal formulation in the displacement have been obtained on the basis of relaxation principles, which leads to the nonreflexive Banach space of functions with bounded deformation. Moreover, Fenchel duality yields the relation to the dual problem in terms of the mechanical stress [3, 54]. These developments build upon the suitable generalized pairing of stresses and strains from [35], which is not straightforward since the strain in perfect plasticity is just a measure.

Surprisingly, it was not until the rather recent work of Dal Maso, DeSimone, and Mora [15] that the corresponding *primal problem* of quasi-static perfect plasticity has been examined in a satisfying way. In this respect, the proper extension of the stress-strain duality from Hencky plasticity to the time-dependent case is the key to a (primal) problem formulation based on the abstract theory of energetic formulations for a very general class of rate-independent systems; see [37, 40]. In [15], it is further shown that a quasi-static evolution can be consistently approximated by a sequence of time-discrete problems. Moreover, the equivalence to the stress-based weak formulation from [34] is shown. The new formulation of perfect plasticity from [15] has gained increasing interest during the last decade, giving rise to several important extensions, for example to pressure-sensitive yield criteria [38], heterogeneous materials [21, 51], regularity theory [16] or coupled with other physical effects [7, 46]. Under minimal regularity, a quasi-static evolution in perfect plasticity can be obtained as an appropriate limit of plasticity problems with vanishing hardening [8].

On the numerical level, the approach from [8] can be coupled with a fully-discrete scheme using an implicit Euler time-discretization together with a standard Finite Element discretization to obtain convergence of displacement, stresses and strains, as mesh size, time step and hardening parameter go to zero. Regularization techniques have also been used earlier to obtain a convergence result for the discretized stresses for a suitable coupling of discretization and regularization parameter; see [45]. Adaptive methods for the static case are discussed, for instance, in [44], [49] and [12].

DOI 10.20347/WIAS.PREPRINT.2376 Berlin 2017

As for algorithmic approaches to the time-discrete problems of perfect plasticity, we mention the standard *return mapping* algorithm from [50] and the superlinear convergence of this generalized Newton method can explained by the semismoothness of the plastic response function [48]. Other approaches comprise SQP [56] and multigrid techniques [55] and typically depend on the smoothness of the yield surface. However, there is no convergence result for the discrete solutions under minimal regularity. Due to the lack of a well-defined infinite-dimensional iteration, the solvers usually display a high degree of mesh-dependence leading to extensive computational overhead on fine meshes. We refer to [19] for a survey on the various complications in both theoretical and computational Prandtl–Reuss plasticity.

For these reason it appears to be worthwhile to develop solvers that have a well-defined infinite-dimensional counterpart. In this regard, the application of an infinite-dimensional augmented Lagrangian method in the vein of [33] to perfect plasticity has been discussed in [47]. However, this method requires the solution of a sequence of visco-plastic problems and the convergence depends on the higher regularity of the strain, which is not the case in perfect plasticity.

The outline of the paper is as follows. In section 3, we recall the system of equations of the Prandtl–Reuss model of perfect plasticity. Thereupon, the properties of the different weak formulations, their interrelation and the generalized stress-strain duality are recalled. The time-incremental problem of quasi-static evolution in perfect plasticity, which involves a convex minimization over the cartesian product of the space of functions of bounded deformation and the space of Borel measures, is considered in Section 3. For this problem we derive an equivalent inf-sup formulation that is posed in a usual separable and reflexive Lebesgue space. The alternative function space setting of the reduced problem further allows to characterize the classical incremental stress problem as a Fenchel dual problem, and we obtain necessary and sufficient optimality conditions for the time-discrete problems. The last section is devoted to a primal-dual regularization scheme that combines the visco-plastic regularization with a penalty type approach with respect to the mechanical equilibrium condition. We further prove the consistency of this regularization approach by showing that displacements, stresses and strains converge to a solution of the initial problem. The Fenchel dual problem of the regularized problem represents a modification of the usual stress problem in perfect plasticity, which may be well-suited for numerical purposes. For the corresponding subproblems, we propose a Tikhonov regularization-based semismooth Newton approach, and we include a convergence result for the regularized problems. Finally, we discuss some open questions related to suitable discretized versions of the approach presented in this paper.

2. Prandtl–Reuss Plasticity and Weak Formulations

In this paper we consider the quasi-static evolution of an elastic-perfectly plastic body subject to (s.t.) a given external loading procedure in the time interval $[0,T]$, $T > 0$. The elasto-plastic material is represented by a bounded domain $\Omega \subset \mathbb{R}^N$, $N \in \{2,3\}$, and it is assumed to be fixed on a nonempty boundary portion $\Gamma_0 \subset \partial\Omega$. The material behavior is described by the displacement u, the mechanical stress σ and the strain tensors e and p describing elastic and plastic strains, respectively.

The time-dependent loading is induced by a volume force $f = f(t,x)$ acting on Ω, and a surface force $g = g(t,x)$ acting on the complement Γ_1 of Γ_0 in $\partial\Omega$. We also adopt the small strain assumption, i.e., the total strain is expected to be reasonably well approximated by the infinitesimal strain tensor

$$\varepsilon(u) = \tfrac{1}{2}(\nabla u + \nabla u^\top).$$

Assuming linear elastic behavior, the relation between elastic strain and stress,

$$e = \mathbb{C}^{-1}\sigma,$$

is determined by a fourth order elasticity tensor $\mathbb{C} \in \mathbb{R}^{(N\times N)^2}$, which is assumed to be symmetric,

$$\mathbb{C}_{ijkl} = \mathbb{C}_{klij} = \mathbb{C}_{jikl}$$

and positive definite,

$$\exists\, \kappa_{\mathbb{C}} > 0 : \quad \mathbb{C}\sigma : \sigma \geq \kappa_{\mathbb{C}} |\sigma|_F^2, \quad \forall\, \sigma \in \mathbb{M}^{N\times N}.$$

Under a uniform positive definiteness assumption on $\mathbb{C}$, the extension to heterogeneous elasticity is immediate. Here, we denote by $\mathbb{M}^{N\times N}$ the space of symmetric $N \times N$-matrices endowed with the Frobenius norm defined by

$$|\sigma|_F = (\sum_{i,j=1}^{N} \sigma_{ij}^2)^{1/2}, \quad \sigma \in \mathbb{M}^{N\times N}.$$

The subspace of symmetric matrices with vanishing trace is indicated by

$$\mathbb{M}_0^{N\times N} := \{\sigma \in \mathbb{M}^{N\times N} : \mathrm{tr}(\sigma) := \sum_{i=1}^{N} \sigma_{ii} = 0\}.$$

Together with an initial condition, the quasi-static evolution of an elastic-perfectly plastic material is described by the following set of conditions.

2.1. **Pressure-insensitive Prandtl–Reuss plasticity.** Given $f : [0, T] \times \Omega \to \mathbb{R}^N$ and $g : [0, T] \times \Gamma_1 \to \mathbb{R}^N$ with $f(0, x) = 0$ in Ω and $g(0, x) = 0$ on Γ_1, find

$$[u, p, \sigma] : [0, T] \times \Omega \to \mathbb{R}^N \times \mathbb{M}_0^{N\times N} \times \mathbb{M}^{N\times N},$$

with

$$[u, p, \sigma](0, x) = 0 \quad \text{in } \Omega \tag{2.1}$$

such that

$$u(t, x) = 0 \quad \text{on } \Gamma_0, \tag{2.2}$$

$$\sigma\nu(t, x) = g(t, x) \quad \text{on } \Gamma_1, \tag{2.3}$$

$$-\operatorname{Div} \sigma(t, x) = f(t, x) \quad \text{in } \Omega, \tag{2.4}$$

$$\varepsilon(u)(t, x) = \mathbb{C}^{-1}\sigma(t, x) + p(t, x) \quad \text{in } \Omega, \tag{2.5}$$

$$\operatorname{dev} \sigma(t, x) \in \mathbb{K}_0 \quad \text{in } \Omega \tag{2.6}$$

$$i^*_{\mathbb{K}_0}(\dot{p}(t, x)) = p(t, x) : \operatorname{dev} \sigma(t, x) \quad \text{in } \Omega, \tag{2.7}$$

for all $t \in [0, T]$.

Here, ν is the unit outer normal to $\partial\Omega$. The set of admissible stresses $\mathbb{K}_0 \subset \mathbb{M}_0^{N\times N}$ is assumed to be a nonempty, compact and convex neighborhood of the origin in $\mathbb{M}_0^{N\times N}$. Its support support function is defined as the convex conjugate $i^*_{\mathbb{K}_0}$ of the indicator function

$$i_{\mathbb{K}_0} : \mathbb{M}_0^{N\times N} \to \mathbb{R} \cup \{+\infty\},$$

of $\mathbb{K}_0 \subset \mathbb{M}_0^{N\times N}$.

The conditions (2.1)-(2.7) are interpreted as follows. Equation (2.4) represents the usual mechanical equilibrium condition neglecting inertial effects. The additive split of the total strain into an elastic part $e = \mathbb{C}^{-1}\sigma$ and an inelastic part p is given in (2.5). Under the assumption that the yield criterion is pressure-insensitive, the set of admissible stresses is given by the constraint (2.6) on the deviatoric part of the stress,

$$\operatorname{dev} \sigma := \sigma - \tfrac{\mathrm{tr}(\sigma)}{N} I_N.$$

The last condition (2.7) is equivalent to the associative flow law

$$\dot{p}(t, x) \in N_{\mathbb{K}_0}(\operatorname{dev} \sigma(t, x)), \tag{2.8}$$

where $N_{\mathbb{K}_0}(\operatorname{dev} \sigma(t, x))$ denotes the normal cone to $\mathbb{K}_0$ at $\operatorname{dev} \sigma(t, x)$. At this point we emphasize that we do not assume the yield surface $\partial\mathbb{K}_0$ to be smooth. The system is supplemented by an initial condition (2.1) and by the mixed boundary conditions (2.2)-(2.3). We proceed by discussing weak formulations of (2.1)-(2.7).

2.2. **Functional analytic setting of weak formulations.** On the mathematical level, a fundamental difference to elasto-plastic problems with hardening [25] is that optimal displacements a priori cannot be expected to lie in the Sobolev space

$$V := H^1_{0,\Gamma_0}(\Omega)^N := \{u \in H^1(\Omega)^N : u|_{\Gamma_0} = 0\}.$$

This is a consequence of the fact that, in the absence of hardening, the material may form shear bands. From the mathematical point of view, this is reflected by the observation that displacements may exhibit discontinuities on $(N-1)$-dimensional submanifolds, which rules out the usual Sobolev setting. The appropriate relaxation, which goes back to [53], requires that the displacement is sought in the space of functions with bounded deformation, which is defined as

$$BD(\Omega) = \{u \in L^1(\Omega)^N : \varepsilon(u) \in M(\Omega; \mathbb{M}^{N\times N})\}.$$

Here, for any Borel set $B \subset \mathbb{R}^N$, $M(B;\mathbb{R}^d)$ denotes the space of $\mathbb{R}^d$-valued Borel measures (that is an $\mathbb{R}^d$-valued σ-additive measure). The space $M(B;\mathbb{R}^d)$ is equipped with the total variation norm and the Riesz-Alexandrov Theorem provides an isometric isomorphism between $M(B;\mathbb{R}^d)$ and $[C_0(B;\mathbb{R}^d)]^*$, the topological dual of the space of continuous functions vanishing at the boundary of B; see, e.g., [2, Prop. 1.47]. Using this identification, we consider $\varepsilon(u)$ for $u \in BD(\Omega)$ as an $\mathbb{M}^{N\times N}$-valued distribution that is also continuous on $C_c^\infty(\Omega;\mathbb{M}^{N\times N})$ equipped with the supremum norm. As a result, the total variation norm on $M(B;\mathbb{M}^{N\times N})$ is given by

$$\|\mu\|_{M(B;\mathbb{M}^{N\times N})} = \sup\{\langle \mu, \varphi\rangle : \varphi \in C_0(B;\mathbb{M}^{N\times N}),\ |\varphi(x)|_F \le 1\ \ \forall\, x \in B\}.$$

The norm on the space $M(B;\mathbb{M}_0^{N\times N})$ is defined analogously. The space $BD(\Omega)$ is equipped with the standard norm

$$\|u\|_{BD(\Omega)} = \|u\|_{L^1(\Omega;\mathbb{R}^N)} + \|\varepsilon(u)\|_{M(\Omega;\mathbb{M}^{N\times N})}.$$

We recall that

$$BD(\Omega) \hookrightarrow L^{N/(N-1)}(\Omega;\mathbb{R}^N), \tag{2.9}$$

i.e., $BD(\Omega)$ embeds continuously into $L^{N/(N-1)}(\Omega;\mathbb{R}^N)$. Under the condition that $\partial\Omega$ is sufficiently smooth, functions in $BD(\Omega)$ admit an integrable trace on the boundary, i.e., $u \in L^1(\partial\Omega)$, and the following Green's formula for functions $u \in BD(\Omega)$ and $\varphi \in C^1(\overline{\Omega})$ is available;

$$\int_\Omega \varphi\, \varepsilon_{ij}(u) = -\tfrac{1}{2}\int_\Omega u_i\partial_j\varphi + u_j\partial_i\varphi\, dx + \int_{\partial\Omega} [u \odot \nu]_{ij}\, \varphi\, d\mathcal{H}^{N-1}). \tag{2.10}$$

Here, the symmetrized outer product of two vectors a and b of the same length is denoted by

$$a \odot b = \tfrac{1}{2}(ab^\top + ba^\top).$$

Moreover, $BD(\Omega)$ can be characterized as the dual space of a separable normed space. This give rise to a weak*-topology on $BD(\Omega)$ for which bounded subsets are sequentially compact. For a sequence $(u_k) \subset BD(\Omega)$ it is known that (u_k) converges weakly* to u in $BD(\Omega)$ if and only if

$$u_k \to u \text{ in } L^1(\Omega), \quad \varepsilon(u_k) \overset{*}{\rightharpoonup} \varepsilon(u) \text{ in } M(\Omega). \tag{2.11}$$

For these results and further details on the space $BD(\Omega)$ we refer to [54, 52]. As a consequence of the low regularity of the displacement, the plastic strains are only Borel measures. Furthermore, the appropriate relaxation of the Dirichlet boundary condition (2.2) entails that the plastic strains may also be supported on Γ_0; cf. [54, 15]. Consequently, the proper function space for the plastic strain in the weak formulation of (2.1)-(2.7) is given by $M(\Omega \cup \Gamma_0; \mathbb{M}_0^{N\times N})$, the space of $\mathbb{M}_0^{N\times N}$-valued Borel measures on $\Omega \cup \Gamma_0$, and the set of admissible states W_{ad} is defined as

$$W_{\text{ad}} := \{(u, e, p) \in BD(\Omega) \times Q \times M(\Omega \cup \Gamma_0; \mathbb{M}_0^{N\times N}) : \tag{2.12}$$
$$\varepsilon(u) = p\lfloor\Omega + e,\ p\lfloor\Gamma_0 = -(u \odot \nu)\mathcal{H}^{n-1}\},$$

where $p\lfloor\Omega$ and $p\lfloor\Gamma_0$ designate the restriction of the measure p to Ω and Γ_0, respectively.

In order to describe the set of admissible stresses and to properly define the stress-strain duality for perfect plasticity, we further define the following stress-related spaces,

$$Q := L^2(\Omega; \mathbb{M}^{N\times N}), \quad H(\mathrm{Div}; \Omega) := \{\sigma \in Q : \mathrm{Div}\,\sigma \in L^2(\Omega; \mathbb{R}^N)\},$$
$$\Sigma(\mathrm{Div}; \Omega) := \{\sigma \in Q : \mathrm{Div}\,\sigma \in L^N(\Omega; \mathbb{R}^N)\},$$

together with their standard norms

$$||q||_Q := \left(\int_\Omega |q|_F^2\, dx\right)^{1/2}, \quad ||\sigma||_{H(\mathrm{Div};\Omega)} := ||\sigma||_Q + ||\,\mathrm{Div}\,\sigma||_{L^2(\Omega,\mathbb{R}^N)},$$
$$||\sigma||_{\Sigma(\mathrm{Div};\Omega)} := ||\sigma||_Q + ||\,\mathrm{Div}\,\sigma||_{L^N(\Omega,\mathbb{R}^N)}.$$

For the sake of accuracy, we distinguish between the usual (distributional) divergence operator div and its vector-valued version Div. We also recall that the (normal) trace of any element $\sigma \in H(\mathrm{Div}; \Omega)$ on the boundary $\partial\Omega$ is defined by the trace operator

$$\tau_\nu : H(\mathrm{Div}; \Omega) \to H^{-1/2}(\partial\Omega)^N, \quad \sigma \mapsto \sigma\nu = \tau_\nu(\sigma),$$

with values in $H^{-1/2}(\partial\Omega)^N$, where $H^{-1/2}(\partial\Omega)$ is the dual space of the trace space $H^{1/2}(\partial\Omega)$. The trace mapping is defined by extension in the usual way. In the same vein, one may also define a trace on a (sufficiently regular) subset of $\Gamma_1 \subset \partial\Omega$; the appropriate trace operator is given by

$$\tau_\nu^{\Gamma_1} : H(\mathrm{Div}; \Omega) \to H_{00}^{-1/2}(\Gamma_1)^N, \quad \sigma \mapsto \sigma\nu|_{\Gamma_1} = \tau_\nu^{\Gamma_1}(\sigma).$$

Here, the image space involves the dual space $H_{00}^{-1/2}(\Gamma_1) := H_{00}^{1/2}(\Gamma_1)^*$ of the trace space

$$H_{00}^{1/2}(\Gamma_1) := \{v \in L^2(\Gamma_1) :\ \exists\, \tilde{v} \in H^1(\Omega), \tilde{v}|_{\Gamma_0} = 0, \tilde{v}|_{\Gamma_1} = v\}.$$

For details on trace spaces we refer to [6, 43]. Given a fixed subspace $X(\Omega) \subset Q$, the set of admissible stresses in $X(\Omega)$ is denoted by

$$S_{\text{ad}}(X(\Omega)) := \{\sigma \in X(\Omega) : \mathrm{dev}\,\sigma(x) \in \mathbb{K}_0 \text{ a.e. in } \Omega\},$$

and if $X(\Omega) = Q$ we write $S_{\text{ad}} := S_{\text{ad}}(Q)$.

2.3. **Johnson's weak formulation.** Following the seminal work of Johnson [34], a suitable weak formulation of Prandtl–Reuss plasticity is given by the following time-dependent variational inequality problem in the velocity and the stress.

Problem 2.1 (Johnson's weak formulation)**.** Let $f \in C([0, T]; L^N(\Omega)^N)$, and $g \in C([0, T]; L^\infty(\Gamma_1)^N)$ with $f(0) = 0$, $g(0) = 0$. Find

$$[\dot{u}, \sigma] : [0, T] \to BD(\Omega) \times Q \quad \text{with } \sigma(0) = 0$$

such that $\sigma \in S_{\text{ad}}(\Sigma(\mathrm{Div}; \Omega))$ and

$$(\sigma, \varepsilon(\tilde{u})) = \langle l(t), \tilde{u}\rangle, \quad \forall\, \tilde{u} \in V, \tag{2.13}$$

$$\langle \dot{u}, \mathrm{Div}\,\tilde{\sigma} - \mathrm{Div}\,\sigma\rangle + (\mathbb{C}^{-1}\dot{\sigma}, \tilde{\sigma} - \sigma) \geq 0 \tag{2.14}$$
$$\forall\, \tilde{\sigma} \in S_{\text{ad}}(\Sigma(\mathrm{Div}; \Omega)),\ \tilde{\sigma}\nu = g(t) \text{ on } \Gamma_1,$$

for a.e. $t \in (0, T)$.

Here, the functional $l(t) \in BD(\Omega)^*$ is defined by

$$l(t) := \int_\Omega f(t)u \, dx + \int_{\Gamma_1} g(t)u \, d\mathcal{H}^{N-1}.$$

Combining a discretization in time with a regularization of the constraint $\sigma \in S_{\text{ad}}$, existence of a solution to this problem is shown in [34, 53] under a non-degenerateness assumption on the loading procedure; cf. Assumption 2.4. It is further obvious that for any solution $[\dot{u}, \sigma]$ of Problem 2.1, the optimal stress σ also solves the following problem.

Problem 2.2 (Johnson's stress problem)**.** Let $f \in C([0,T]; L^N(\Omega)^N)$, and $g \in C([0,T]; L^\infty(\Gamma_1)^N)$ with $f(0) = 0$, $g(0) = 0$. Find

$$\sigma : (0,T) \to Q \quad \text{with } \sigma(0) = 0$$

such that $\sigma \in S_{\text{ad}}(\Sigma(\mathrm{Div};\Omega))$ and

$$\begin{aligned} (\sigma, \varepsilon(\tilde{u})) &= \langle l(t), \tilde{u} \rangle \quad \forall\, \tilde{u} \in V, \\ (\mathbb{C}^{-1}\dot{\sigma}, \tilde{\sigma} - \sigma) \geq 0 \quad &\forall\, \tilde{\sigma} \in S_{\text{ad}}(\Sigma(\mathrm{Div};\Omega)),\ \tilde{\sigma}\nu = g(t) \text{ on } \Gamma_1, -\mathrm{Div}\,\tilde{\sigma} = f(t), \end{aligned}$$

for a.e. $t \in (0,T)$.

In particular, any solution $\sigma : [0,T] \to Q$ pertaining to Problem 2.1 is uniquely determined by the initial condition.

2.4. **Quasi-static evolution.** The problem of Prandtl–Reuss plasticity may also be studied within the context of energetic formulations for a general class of rate-independent systems that are defined by the axioms of energy stability and energy balance [37]. This ultimately leads to a primal problem in u, p and e, which has been derived and analyzed in [15]. Following the latter reference, we make the following assumptions:

$$\Omega \text{ is a bounded } C^2\text{-domain}; \tag{2.15}$$

$$\partial\Omega = \overline{\Gamma}_0 \cup \overline{\Gamma}_1; \quad \Gamma_0 \neq \emptyset, \quad \Gamma_0, \Gamma_1 \subset \partial\Omega \text{ open}; \quad \partial\Gamma_0 = \partial\Gamma_1 \in C^2, \tag{2.16}$$

i.e., the C^2-boundary $\partial\Omega$ is split into two disjoint relatively open parts Γ_0 and Γ_1, with a joint C^2-regular interface $\partial\Gamma_0 = \partial\Gamma_1$ in the sense of [35, p. 20]. These classical geometric conditions may be alleviated at the cost of some nontrivial modifications [21]. We also assume that the elasticity tensor is invariant with respect to the orthogonal subspaces $\mathbb{M}_0^{N\times N}$ and $\{cI_N : c \in \mathbb{R}\}$. Consequently, there exists a positive definite tensor $\mathbb{C}_{\text{dev}} \in (\mathbb{M}_0^{N\times N})^2$ and a scalar $\lambda_0 > 0$ such that

$$\mathbb{C}\sigma = \mathbb{C}_{\text{dev}}\,\mathrm{dev}\,\sigma + \lambda_0\,\mathrm{tr}\,\sigma I_N, \quad \forall\, \sigma \in \mathbb{M}^{N\times N}. \tag{2.17}$$

For $p \in M(\Omega \cup \Gamma_0; \mathbb{M}_0^{N\times N})$, one may further define the functional

$$D(p) := \int_{\Omega\cup\Gamma_0} i^*_{\mathbb{K}_0}(p/|p|)\, d|p| = i^*_{\mathbb{K}_0}(p)(\Omega \cup \Gamma_0),$$

on the basis of the theory of convex functions of measures; see [54, Chapter II(5.)] for an introduction. In fact, denoting the Radon-Nikodým derivative p with respect to its variation $|p|$ by $p/|p|$, we have

$$p/|p| \in L^1_{|p|}(\Omega \cup \Gamma_0; \mathbb{M}_0^{N\times N}),$$

i.e., $p/|p|$ is Lebesgue integrable on $\Omega \cup \Gamma_0$ with respect to the measure $|p|$, such that

$$D : M(\Omega \cup \Gamma_0; \mathbb{M}_0^{N\times N}) \to \mathbb{R}$$

is well-defined, nonnegative and finite. The generalized total variation with respect to D,

$$\mathcal{D}(p; 0, t) := \sup\{\sum_{n=1}^{K} D(p(t_n) - p(t_{n-1})) : K \in \mathbb{N}, 0 = t_0 \leq t_1 \leq \ldots \leq t_K = t\},$$

then accounts for the dissipation in the time interval $[0, t]$, $t \le T$. We further denote the space of absolutely continuous functions on $[0, T]$ with values in a Banach space X by $AC([0, T]; X)$. The space $BV([0, T]; X)$ consists of all X-valued functions on $[0, T]$ with bounded variation. We are now ready to state the notion of quasi-static evolution in perfect plasticity from [15].

Problem 2.3 (Quasi-static evolution)**.** Given

$$f \in AC([0, T]; L^N(\Omega)^N), \quad g \in AC([0, T]; L^\infty(\Gamma_1)^N) \tag{2.18}$$

with $f(0) = 0$ and $g(0) = 0$, find

$$[u, e, p] : [0, T] \to BD(\Omega) \times Q \times M(\Omega \cup \Gamma_0; \mathbb{M}_0^{N\times N})$$

with $[u, e, p](0) = 0$ such that $t \mapsto [u(t), e(t), p(t)]$ is a quasi-static evolution, i.e., the following conditions are fulfilled.

(i) Stability: For every $t \in [0, T]$, it holds that $[u(t), e(t), p(t)] \in W_{\text{ad}}$ and

$$\tfrac{1}{2}(\mathbb{C}e(t), e(t)) - \langle l(t), u(t)\rangle \le \tfrac{1}{2}(\mathbb{C}\tilde{e}, \tilde{e}) + D(\tilde{p} - p(t)) - \langle l(t), \tilde{u}\rangle$$

for all $[\tilde{u}, \tilde{e}, \tilde{p}] \in W_{\text{ad}}$.

(ii) Energy equality: It holds that $p \in BV([0, T], M(\Omega \cup \Gamma_0; \mathbb{M}_0^{N\times N}))$, and for every $t \in [0, T]$ the equation

$$\tfrac{1}{2}(\mathbb{C}e(t), e(t)) - \langle l(t), u(t)\rangle + \mathcal{D}(p; 0, t) = -\int_0^t \langle \dot{l}(s), u(s)\rangle ds$$

is valid.

Under the above assumptions, the existence of a quasi-static evolution $[u, e, p] \in AC([0, T]; BD(\Omega)\times Q \times M(\Omega \cup \Gamma_0; \mathbb{M}_0^{N\times N}))$ can be shown provided a safe-load condition holds uniformly in time.

Assumption 2.4 (Safe-load condition)**.** *There exists $\hat{\sigma} \in AC([0, T]; Q)$ and $\rho > 0$ such that*

(i) $\operatorname{dev}\hat{\sigma} \in AC([0, T]; L^\infty(\Omega; \mathbb{M}_0^{N\times N}))$*,*

(ii) for every $t \in [0, T]$ it holds that

$$\operatorname{Div}\hat{\sigma}(t) = -f(t) \text{ in } \Omega, \quad \hat{\sigma}(t)\nu = g(t) \text{ on } \Gamma_1,$$
$$\operatorname{dev}\hat{\sigma}(t) + B_\rho(0) \subset \mathbb{K}_0 \text{ a.e. in } \Omega,$$

where $B_\rho(0) := \{\tau \in \mathbb{M}_0^{N\times N} : |\tau|_F \le \rho\}$.

In this case, solutions of Problem 2.3 can be obtained as appropriate limits of a sequence of solutions of time-incremental problems defined in section 3; see [15, Theorem 4.5]. Under mild assumptions on the regularity in time, quasi-static evolutions correspond to solutions of the classical weak formulation from Problem 2.1 [15, Theorem 6.1]. In this sense, Problem 2.1 and Problem 2.3 are essentially equivalent. The equivalence of the two solution notions as well as the formal equivalence to the system (2.1)-(2.7) relies on a suitable extension of the meaning of the flow law (2.7) to linearized strains $\varepsilon(u)$ that are only Borel measures, and which reduces to the conventional meaning if $\dot{p} \in Q$. For that reason a duality pairing between admissible stresses and strains is defined in [15], which extends earlier approaches within the context of Hencky plasticity set forth in [35, 54]. Since the stress is in general not continuous this is by no means a trivial issue, and the particular problem structure has to be exploited. In fact, for $[u, e, p] \in W_{\text{ad}}$ and $\sigma \in S_{\text{ad}}(\Sigma(\operatorname{Div}; \Omega))$ a suitable pairing is given by

$$[\operatorname{dev}\sigma, p] := \begin{cases} [\operatorname{dev}\sigma, \operatorname{dev}\varepsilon(u)] - \operatorname{dev}\sigma : \operatorname{dev} e, & \text{in } \Omega, \\ -(\sigma\nu)_T \cdot u\,\mathcal{H}^{N-1}, & \text{on } \Gamma_0, \end{cases} \tag{2.19}$$

where $[\operatorname{dev}\sigma, \operatorname{dev}\varepsilon(u)] \in M(\Omega)$ denotes the measure defined in [35, Theorem 3.2], and $(\sigma\nu)_T := \sigma\nu - (\sigma\nu)_\nu \nu$ is the tangential component of $\sigma\nu$. Accounting for [35, Lemma 2.4], it holds $(\sigma\nu)_T \in$

$L^\infty(\partial\Omega)$, such that $[\operatorname{dev}\sigma, p]$ is well-defined. Moreover it holds $[\operatorname{dev}\sigma, p] \in M(\Omega\cup\Gamma_0)$. The following integration by parts formula from [15] provides a useful characterization of this generalized duality; it holds

$$[\operatorname{dev}\sigma : p](\Omega \cup \Gamma_0) = -(\sigma, e) - \langle \operatorname{Div}\sigma, u\rangle + \langle \sigma\nu, u\rangle_{\Gamma_1}, \tag{2.20}$$

where $\sigma \in S_{\text{ad}}(\Sigma(\operatorname{Div};\Omega))$ with $\sigma\nu \in L^\infty(\Gamma_1;\mathbb{R}^N)$ and $(u, e, p) \in W_{\text{ad}}$. Here, the duality product on the right hand side of (2.20) designates the pairing of $L^\infty(\Gamma_1;\mathbb{R}^N)$ with $L^1(\Gamma_1;\mathbb{R}^N)$. Note that, a priori, we have $\sigma\nu|_{\Gamma_1} \in H_{00}^{-1/2}(\Gamma_1)^N$ since

$$\sigma \in \Sigma(\operatorname{Div};\Omega) \subset H(\operatorname{Div};\Omega).$$

For quasi-static evolution problems in perfect plasticity, important extensions, for example to pressure-sensitive yield criteria [38] and heterogeneous plasticity [51, 21], are available. For an overview of classical approaches via the so-called stress problem we also refer to [19] and the references therein.

3. The time-incremental problem

In this section we formulate the incremental problem of quasi-static evolution of perfect plasticity in weak form. For this purpose we adopt the Assumption 2.4 as well as (2.15), (2.16) and (2.17).

3.1. **Problem statement.**

We assume that the time interval is partitioned into K subdivisions,

$$0 = t_0 < t_1 < \ldots < t_K = T.$$

At a fixed time point t_n, $n = 1, \ldots, K$, we are given the state of the system

$$[u^{n-1}, e^{n-1}, p^{n-1}] \in W_{\text{ad}}$$

from the preceding time instance as well as the current applied forces

$$f^n = f(t_n) \in L^N(\Omega;\mathbb{R}^N), \quad g^n = g(t_n) \in L^\infty(\Gamma_1;\mathbb{R}^N),$$

which define the total load

$$l^n(u) := \int_\Omega f^n \cdot u \, dx + \int_{\Gamma_1} g^n \cdot u \, d\mathcal{H}^{N-1}, \quad u \in BD(\Omega).$$

The time-discretized problem of perfect plasticity at a fixed time instance can be stated as follows [15].

Problem (P).

$$\begin{cases} \inf \quad J(u,e,p) \quad \text{over } [u,e,p] \in BD(\Omega) \times Q \times M_0(\Omega \cup \Gamma_0; \mathbb{M}_0^{N\times N}) \\ \text{s.t.} \quad [u,e,p] \in W_{\text{ad}}, \end{cases}$$

where W_{ad} is given by (2.12) and the objective functional J is defined by

$$J(u,e,p) := \tfrac{1}{2}(\mathbb{C}e, e) + D(p - p^{n-1}) - \langle l^n, u\rangle.$$

Under the safe-load condition (Assumption 2.4), an equivalent characterization of D is given by

$$D(p) = \sup\{[\operatorname{dev}\sigma : p](\Omega \cup \Gamma_0) : \sigma \in S_{\text{ad}}(\Sigma(\operatorname{Div};\Omega)), \sigma\nu = g^n \text{ on } \Gamma_1\}, \tag{3.1}$$

where the measure $[\operatorname{dev}\sigma : p] \in M(\Omega \cup \Gamma_0)$ is defined in (2.19); for details see [15, Prop. 2.4].

Provided the applied forces fulfill Assumption 2.4, it can be shown that problem (P) has a solution $[u^n, e^n, p^n]$, which is in general only unique in the elastic strain. Below we provide an alternative to the existence proof from [15, Theorem 3.3]. Furthermore, one may construct iteratively a piecewise constant time interpolate from the solutions $[u^n, e^n, p^n]$ of (P) for $n = 1, \ldots K$. For a sequence of subdivisions with vanishing time step, the resulting sequence of time interpolates converges to a quasi-static evolution [15, Theorem 4.5]. Since the time step is kept fixed in the remainder of this paper, we write $[\bar{u}, \bar{e}, \bar{p}] = [u^n, e^n, p^n]$ for a solution to (P).

3.2. **Inf-sup problem formulation.** From a computational point of view, problem (P) poses a variety of complexities; the problem is posed in a nonreflexive Banach space, the objective function is nonsmooth and the constraints are posed in a measure space.

This section is dedicated to a suitable problem reduction, which yields an unconstrained equivalent reformulation posed in a conventional reflexive Lebesgue space. Based on this reformulation, we establish a Fenchel duality result that relates the primal formulation (P) to the (classical) incremental version of the stress problem (Problem 2.2).

Using the constraints in (P), we first eliminate the dependence on p from the optimization problem defining

$$p\lfloor_\Omega = \varepsilon(u) - e, \quad p\lfloor_{\Gamma_0} = -u \odot \nu\, \mathcal{H}^{N-1}. \tag{3.2}$$

Lemma 3.1. *Define $\hat{J} : BD(\Omega) \times Q \to \mathbb{R}$ by*

$$\hat{J}(u,e) := \tfrac{1}{2}(\mathbb{C}e, e) + \sup_{\substack{\sigma \in S_{\mathrm{ad}}(\Sigma(\mathrm{Div};\Omega)) \\ \sigma\nu = g^n \text{ on } \Gamma_1}} \{-\langle \hat{p}^{n-1}, \sigma\rangle - (\sigma, e) - \langle \mathrm{Div}\,\sigma, u\rangle\} - \langle f^n, u\rangle, \tag{3.3}$$

where $\hat{p}^{n-1}$ is understood as an element of $\Sigma(\mathrm{Div};\Omega)^*$ defined by

$$\langle \hat{p}^{n-1}, \sigma\rangle := -(\sigma, e^{n-1}) - \langle \mathrm{Div}\,\sigma, u^{n-1}\rangle, \quad \sigma \in \Sigma(\mathrm{Div};\Omega). \tag{3.4}$$

Then (P) is equivalent to the problem

$$\begin{cases} \inf & \hat{J}(u,e) \quad \text{over } [u,e] \in BD(\Omega) \times Q, \\ \text{s.t.} & \mathrm{div}\, u = \mathrm{tr}\, e \quad \text{in } \Omega, \\ & u \cdot \nu = 0 \quad \text{on } \Gamma_0, \end{cases} \tag{3.5}$$

in the following sense.

- *(i) If $[\bar{u}, \bar{e}, \bar{p}]$ is a solution of (P) then $[\bar{u}, \bar{e}]$ solves (3.5).*
- *(ii) For each solution $[\bar{u}, \bar{e}]$ to (3.5), it holds that $[\bar{u}, \bar{e}, p(\bar{u}, \bar{e})]$ is a solution to (P), where $p(\bar{u}, \bar{e})$ is defined by (3.2).*

Proof. Let $[u, e, p] \in W_{\mathrm{ad}}$. As the safe-load condition is assumed to hold, we have

$$D(p) = \sup\{[\mathrm{dev}\,\sigma, p](\Omega \cup \Gamma_0) : \sigma \in S_{\mathrm{ad}}(\Sigma(\mathrm{Div};\Omega)), \sigma\nu = g^n \text{ on } \Gamma_1\},$$

by (3.1). By (2.20) we further obtain for all $\sigma \in S_{\mathrm{ad}}(\Sigma(\mathrm{Div};\Omega))$ with $\sigma\nu = g^n$ on Γ_1,

$$\begin{aligned}[\mathrm{dev}\,\sigma, p - p^{n-1}](\Omega \cup \Gamma_0) = &- (\sigma, e) - \langle \mathrm{Div}\,\sigma, u\rangle + \langle g^n, u\rangle_{\Gamma_1} \\ &- \langle \hat{p}^{n-1}, \sigma\rangle - \langle g^n, u^{n-1}\rangle,\end{aligned}$$

where $\hat{p}^{n-1}$ is defined in (3.4). Note that $\langle g^n, u^{n-1}\rangle$ is a constant, and, without loss of generality, we assume $\langle g^n, u^{n-1}\rangle = 0$. Hence, we may remove the dependence on p of the objective functional;

$$J(u,e,p) = \hat{J}(u,e), \quad \forall\, [u,e,p] \in W_{\mathrm{ad}}. \tag{3.6}$$

Now let $[\bar{u}, \bar{e}, \bar{p}] \in W_{\mathrm{ad}}$ be a solution of (P) and $[u, e] \in \tilde{W}_{\mathrm{ad}}$, where

$$\tilde{W}_{\mathrm{ad}} := \{[u,e] \in BD(\Omega) \times Q : \mathrm{div}\, u = \mathrm{tr}\, e \text{ in } L^2(\Omega),\ u \cdot \nu = 0 \text{ a.e. on } \Gamma_0\}.$$

By taking the trace in the two conditions (2.12) of the definition of W_{ad}, one may observe that $p = p(u,e)$, cf. (3.2), defines an element $p \in M(\Omega \cup \Gamma_0; \mathbb{M}_0^{N\times N})$ such that $[u, e, p] \in W_{\mathrm{ad}}$ if and only if $[u, e] \in \tilde{W}_{\mathrm{ad}}$. Using (3.6), one deduces that

$$\hat{J}(\bar{u}, \bar{e}) = J(\bar{u}, \bar{e}, \bar{p}) \le J(u, e, p(u,e)) = \hat{J}(u,e),$$

for all $[u, e] \in \tilde{W}_{\mathrm{ad}}$. This proves assertion (i).

Let $[\bar{u},\bar{e}] \in \tilde{W}_{\text{ad}}$ be a solution of (3.5). Following the above discussion, we find that for any $[u,e,p] \in W_{\text{ad}}$ it holds that $[u,e] \in \tilde{W}_{\text{ad}}$. Hence, (3.6) implies that

$$J(\bar{u},\bar{e},p(\bar{u},\bar{e})) = \hat{J}(\bar{u},\bar{e}) \le \hat{J}(u,e) = J(u,e,p),$$

for all $[u,e,p] \in W_{\text{ad}}$, which accomplishes the proof of assertion (ii). □

Since the yield criterion is pressure-insensitive, it can be expected that there is no need to explicitly take account of the plastic incompressibility constraint $\operatorname{tr} p = 0$. In fact, the following lemma shows that the constraints in (3.5) are redundant.

Lemma 3.2. *Let $\hat{J}$ be given by (3.3). The problem*

$$\inf \quad \hat{J}(u,e) \quad \textit{over } [u,e] \in BD(\Omega) \times Q \tag{3.7}$$

is equivalent to (P) in the sense of Lemma 3.1.

Proof. Let $[u,e] \in BD(\Omega) \times Q$. For arbitrary $\varphi \in C^1(\overline{\Omega})$ with $\varphi = 0$ on Γ_1 we define

$$\sigma_\varphi := \hat{\sigma}^n + \varphi I_N,$$

for $\hat{\sigma}^n := \hat{\sigma}(t_n)$ where $\hat{\sigma}$ is the admissible stress evolution according to (2.4). Thus, it holds that $\sigma_\varphi \in S_{\text{ad}}(\Sigma(\operatorname{Div};\Omega))$ with $\sigma_\varphi \nu = g^n$ on Γ_1 and one may derive the following estimate;

$$\begin{aligned}
&\sup_{\substack{\sigma \in S_{\text{ad}}(\Sigma(\operatorname{Div};\Omega)),\\ \sigma\nu = g^n \text{ on } \Gamma_1}} \{-\langle \hat{p}^{n-1},\sigma\rangle - (\sigma,e) - \langle \operatorname{Div}\sigma, u\rangle\} - \langle f^n, u\rangle \\
&\ge \sup_{\varphi \in C^1(\overline{\Omega}),\varphi=0 \text{ on } \Gamma_1} \{-\langle \hat{p}^{n-1},\sigma_\varphi\rangle - (\sigma_\varphi,e) - \langle \operatorname{Div}\varphi I_N, u\rangle\} \\
&= -\langle \hat{p}^{n-1},\hat{\sigma}^n\rangle - (\hat{\sigma}^n,e) + \sup_{\substack{\varphi \in C^1(\overline{\Omega}),\\ \varphi=0 \text{ on } \Gamma_1}} \left\{-\langle \hat{p}^{n-1},\varphi I_N\rangle - (\varphi, \operatorname{tr} e) - \langle \nabla\varphi, u\rangle\right\}.
\end{aligned}$$

Taking the trace in the Green's formula (2.10) implies that

$$\int_\Omega u \cdot \nabla\varphi \, dx = -\int_\Omega \varphi \; d(\operatorname{div} u) + \int_{\partial\Omega} u\nu\,\varphi \, d\mathcal{H}^{N-1}, \tag{3.8}$$

for all $\varphi \in C^1(\overline{\Omega})$, such that

$$-\langle \hat{p}^{n-1}, \varphi I_N\rangle = (\varphi, \operatorname{tr} e^{n-1}) + \langle \nabla\varphi, u^{n-1}\rangle = 0. \tag{3.9}$$

The latter term vanishes since $[u^{n-1},e^{n-1},p^{n-1}] \in W_{\text{ad}}$ implies that

$$\operatorname{div} u^{n-1} = \operatorname{tr} e^{n-1}, \quad u^{n-1}\cdot\nu = 0 \text{ a.e. on } \Gamma_0.$$

By (3.9) and (3.8), one obtains

$$\begin{aligned}
&\sup_{\substack{\sigma \in S_{\text{ad}}(\Sigma(\operatorname{Div};\Omega)),\\ \sigma\nu = g^n \text{ on } \Gamma_1}} \{-\langle \hat{p}^{n-1},\sigma\rangle - (\sigma,e) - \langle \operatorname{Div}\sigma, u\rangle\} - \langle f^n, u\rangle \\
&\ge -\langle \hat{p}^{n-1},\hat{\sigma}^n\rangle - (\hat{\sigma}^n,e) \\
&\quad + \sup_{\substack{\varphi \in C^1(\overline{\Omega}),\\ \varphi=0 \text{ on } \Gamma_1}} \left\{\int_\Omega \varphi\,(d(\operatorname{div} u) - \operatorname{tr} e\, dx) - \int_{\Gamma_0} u\cdot\nu\,\varphi\, d\mathcal{H}^{N-1}\right\},
\end{aligned} \tag{3.10}$$

which implies that $\hat{J}(u,e) = +\infty$ unless

$$\operatorname{div} u - \operatorname{tr} e = 0 \text{ in } \Omega. \tag{3.11}$$

The redundancy of the boundary condition can be derived as follows. It can be verified that the density property

$$\overline{\{\varphi\big|_{\Gamma_0} : \varphi \in C^1(\overline{\Omega}), \varphi = 0 \text{ on } \Gamma_1\}}^{C_0(\Gamma_0)} = C_0(\Gamma_0) \tag{3.12}$$

is fulfilled; in fact, let $w \in C_c(\Gamma_0)$ and choose an extension $\tilde{w} \in C_c(\omega)$ of w to a nonempty open set $\omega \subset \mathbb{R}^N$ with

$$\omega \cap \overline{\Gamma}_1 = \emptyset, \quad \operatorname{supp} w \subset \omega, \quad \tilde{w}|_{\omega\cap\Gamma_0} = w.$$

Let (w_n) be a standard sequence of mollifications of $\tilde{w}$ induced by a smooth kernel $\theta \in C_c(\mathbb{R}^N)$ with

$$\theta \geq 0, \quad \theta|_{B_1(0)^c} = 0, \quad \int_{\mathbb{R}^N} \theta \, dx = 1,$$

i.e.,

$$w_n(x) := (\theta_n * \tilde{w})(x) = \int_{\mathbb{R}^N} \tilde{w}(y)\theta_n(x-y)\, dy, \quad \theta_n(x) := n^N \theta(nx), \qquad \forall\, x \in \mathbb{R}^N.$$

As $\tilde{w} \in C_c(\omega)$, standard properties of mollifications yield that (w_n) converges uniformly to $\tilde{w}$ in ω; cf. [1]. For sufficiently large n, it further holds that $\operatorname{supp} w_n \subset \omega$, and in particular,

$$(w_n|_{\Gamma_0}) \subset \{\varphi\big|_{\Gamma_0} : \varphi \in C^1(\overline{\Omega}), \varphi = 0 \text{ on } \Gamma_1\}.$$

Taking account of the fact that $(w_n|_{\Gamma_0})$ converges uniformly to w on Γ_0, the density property (3.12) is verified.

Exploiting the density property (3.12), one may infer that it holds that

$$\int_{\Gamma_0} u \cdot \nu \, \varphi \;\; d\mathcal{H}^{N-1} = 0, \qquad \forall\, \varphi \in C^1(\overline{\Omega}), \varphi|_{\Gamma_1} = 0,$$

if and only if

$$\|u \cdot \nu \, \mathcal{H}^{N-1}\|_{M(\Gamma_0)} = \|u \cdot \nu\|_{L^1(\Gamma_0)} = 0. \tag{3.13}$$

Finally, (3.10) together with (3.11) and (3.13) imply that $\hat{J}(u, e) < +\infty$ requires that $u \cdot \nu$ vanishes on Γ_0. As a conclusion, the constraints in problem (3.5) are redundant and the assertion follows from Lemma 3.1. □

In comparison to the original problem formulation (P), the elimination of the plastic incompressibility constraints comes at the loss of the finiteness of the objective function. We now prove the coercivity of the objective function pertaining to the equivalent problem (3.5) on $BD(\Omega) \times Q$.

Lemma 3.3. *The reduced objective function*

$$\hat{J} \colon BD(\Omega) \times Q \to \mathbb{R} \cup \{+\infty\}$$

from (3.3) *is coercive. More precisely, there exist constants $c_0 \in \mathbb{R}, c_1 > 0$, such that*

$$\begin{aligned} \sup_{\substack{\sigma \in S_{ad}(\Sigma(\mathrm{Div};\Omega)) \\ \sigma\nu = g^n \text{ on } \Gamma_1}} & \{-\langle \hat{p}^{n-1}, \sigma\rangle - (\sigma, e) - \langle \operatorname{Div}\sigma, u\rangle\} - \langle f^n, u\rangle \\ & \geq c_0 - c_1\|e\|_Q \\ & \quad + \rho \max(\|\varepsilon(u)\|_{M(\Omega;\mathbb{M}^{N\times N})}, -\|\varepsilon(u)\|_{M(\Omega;\mathbb{M}^{N\times N})} + \tfrac{1}{\sqrt{2}}\|u\|_{L^1(\Gamma_0;\mathbb{R}^N)}), \end{aligned} \tag{3.14}$$

for all $[u, e] \in BD(\Omega) \times Q$, where $\rho > 0$ is the constant from Assumption 2.4.

Proof. First, we state the elementary result

$$|\operatorname{dev}\tau|_F \le |\tau|_F \quad \text{for all } \tau \in \mathbb{M}^{N\times N}. \tag{3.15}$$

Making use of Assumption 2.4 and (3.15), it holds that

$$\begin{aligned}
&\sup_{\substack{\sigma\in S_{\mathrm{ad}}(\Sigma(\mathrm{Div};\Omega)),\\ \sigma\nu=g^n \text{ on } \Gamma_1}} \{-\langle \hat{p}^{n-1},\sigma\rangle - (\sigma,e) - \langle \operatorname{Div}\sigma, u\rangle\} - \langle f^n, u\rangle \\
&\ge \sup_{\substack{\tau\in C^1(\overline{\Omega},\mathbb{M}^{N\times N}),\tau=0 \text{ on } \Gamma_1,\\ \|\tau\|_{C(\overline{\Omega};\mathbb{M}^{N\times N})}\le\rho}} \{-\langle \hat{p}^{n-1},\hat{\sigma}^n+\tau\rangle - (\hat{\sigma}^n+\tau,e) - \langle \operatorname{Div}\tau, u\rangle\} \\
&\ge c + \sup_{\substack{\tau\in C^1(\overline{\Omega};\mathbb{M}^{N\times N}),\tau=0 \text{ on } \Gamma_1,\\ \|\tau\|_{C(\overline{\Omega};\mathbb{M}^{N\times N})}\le\rho}} \{-\langle \hat{p}^{n-1},\tau\rangle - (\hat{\sigma}^n+\tau,e) - \langle \operatorname{Div}\tau, u\rangle\},
\end{aligned}$$

for all $e\in Q$ and $u\in BD(\Omega)$, where $c\in\mathbb{R}$ denotes a constant which may take different values on different occasions. Using Green's formula for $BD(\Omega)$-functions (2.10), one obtains

$$\begin{aligned}
-\langle \hat{p}^{n-1},\tau\rangle &= (e^{n-1},\tau) + \langle u^{n-1}, \operatorname{Div}\tau\rangle \\
&\ge -c\|e^{n-1}\|_Q - \int_\Omega \tau : \varepsilon(u^{n-1}) + \int_{\Gamma_0} (u^{n-1}\odot\nu) : \tau \, d\mathcal{H}^{N-1} \\
&\ge -c\|e^{n-1}\|_Q - \rho\left(|\varepsilon(u^{n-1})|_F(\Omega) + \|u^{n-1}\odot\nu\|_{L^1(\Gamma_0;\mathbb{M}^{N\times N})}\right)
\end{aligned}$$

and

$$-(\hat{\sigma}^n+\tau,e) \ge -(\|\hat{\sigma}^n\|_Q + \rho|\Omega|^{1/2})\|e\|_Q$$

for all $\tau\in C^1(\overline{\Omega},\mathbb{M}^{N\times N})$ with $\|\tau\|_{C(\overline{\Omega},\mathbb{M}^{N\times N})}\le\rho$ and $\tau|_{\Gamma_1}=0$. This implies that

$$\begin{aligned}
&\sup_{\substack{\sigma\in S_{\mathrm{ad}}(\Sigma(\mathrm{Div};\Omega)),\\ \sigma\nu=g^n \text{ on } \Gamma_1}} \{-\langle \hat{p}^{n-1},\sigma\rangle - (\sigma,e) - \langle \operatorname{Div}\sigma, u\rangle\} - \langle f^n, u\rangle \\
&\qquad \ge c_0 - c_1\|e\|_Q + \sup_{\substack{\tau\in C^1(\overline{\Omega};\mathbb{M}^{N\times N}),\tau=0 \text{ on } \Gamma_1,\\ \|\tau\|_{C(\overline{\Omega};\mathbb{M}^{N\times N})}\le\rho}} \{-\langle \operatorname{Div}\tau, u\rangle\},
\end{aligned} \tag{3.16}$$

where

$$\begin{aligned}
\sup_{\substack{\tau\in C^1(\overline{\Omega},\mathbb{M}^{N\times N}),\tau=0 \text{ on } \Gamma_1,\\ \|\tau\|_{C(\overline{\Omega};\mathbb{M}^{N\times N})}\le\rho}} \{-\langle \operatorname{Div}\tau, u\rangle\} &\ge \sup_{\substack{\tau\in C_0^1(\Omega;\mathbb{M}^{N\times N}),\\ \|\tau\|_{C_0(\Omega;\mathbb{M}^{N\times N})}\le\rho}} \{-\langle \operatorname{Div}\tau, u\rangle\} \\
&= \rho\|\varepsilon(u)\|_{M(\Omega;\mathbb{M}^{N\times N})}.
\end{aligned} \tag{3.17}$$

Furthermore, it is well known that for $\partial\Omega\in C^2$ each $\tau\in C^1(\partial\Omega)$ may be extended to a function $T_\tau\in C^1(\overline{\Omega})$ given by

$$T_\tau(x) := \varphi(r\operatorname{dist}(x,\partial\Omega))\tau(\pi(x));$$

see [22, Lemma 6.38]. Here, π denotes the locally uniquely determined projection of x onto the boundary $\partial\Omega$, $r\in\mathbb{R}$ is sufficiently large, and $\varphi\in C^\infty(\mathbb{R})$ denotes a smooth function with

$$\varphi(t)\in[0,1]\;\;\forall\, t\in\mathbb{R}, \quad \varphi(t)=0 \text{ for } t\ge 2, \quad \varphi(t)=1 \text{ for } t\le 1.$$

Again using (2.10), one obtains

$$\begin{aligned}
&\sup_{\substack{\tau\in C^1(\overline{\Omega};\mathbb{M}^{N\times N}),\tau=0 \text{ on } \Gamma_1,\\ \|\tau\|_{C(\overline{\Omega};\mathbb{M}^{N\times N})}\leq\rho}} \{-\langle \operatorname{Div}\tau, u\rangle\} \\
&\qquad\geq \sup_{\substack{\tau\in C^1_0(\Gamma_0;\mathbb{M}^{N\times N}),\tau=0 \text{ on } \Gamma_1\\ \|\tau\|_{C_0(\Gamma_0;\mathbb{M}^{N\times N})}\leq\rho}} \{-\langle \operatorname{Div} T_\tau, u\rangle\} \\
&\qquad= \sup_{\substack{\tau\in C^1_0(\Gamma_0;\mathbb{M}^{N\times N}),\\ \|\tau\|_{C_0(\Gamma_0;\mathbb{M}^{N\times N})}\leq\rho}} \left(\int_\Omega T_\tau : \varepsilon(u) - \int_{\Gamma_0} (u\odot\nu) : \tau \, d\mathcal{H}^{N-1}\right) \\
&\qquad\geq \rho\left(-\|\varepsilon(u)\|_{M(\Omega;\mathbb{M}^{N\times N})} + \|u\odot\nu\|_{L^1(\Gamma_0;\mathbb{M}^{N\times N})}\right) \\
&\qquad\geq \rho\left(-\|\varepsilon(u)\|_{M(\Omega;\mathbb{M}^{N\times N})} + \tfrac{1}{\sqrt{2}}\|u\|_{L^1(\Gamma_0;\mathbb{R}^N)}\right).
\end{aligned}$$

In the latter estimate we use the elementary property

$$|a\odot b|_F \geq \tfrac{1}{\sqrt{2}}|a|_2|b|_2,$$

and together with (3.16), (3.17), the proof of (3.14) is accomplished. The coercivity of the objective function $\hat{J}$ in $BD(\Omega)\times Q$ now follows from (3.14), the ellipticity property

$$(\mathbb{C}e, e)_Q \geq \kappa_{\mathbb{C}}\|e\|_Q^2,$$

and the fact that

$$u \mapsto \|u\|_{L^1(\Gamma_0;\mathbb{R}^N)} + \|u\|_{M(\Omega;\mathbb{M}^{N\times N})}$$

defines an equivalent norm on $BD(\Omega)$; see [54]. □

The significance of the preceding lemma is twofold. To begin with, the objective function $\hat{J}$ is also well-defined as an extended real-valued function on $L^{N/(N-1)}(\Omega)^N\times Q$; cf. (3.3). The estimates (3.16) and (3.17) further imply that the implication

$$u\in L^{N/(N-1)}(\Omega)^N \setminus BD(\Omega) \Longrightarrow \hat{J}(u,e) = +\infty, \tag{3.18}$$

is valid for all $e\in Q$, i.e., the regularity constraint $\varepsilon(u)\in M(\Omega)$ is implicitly fulfilled as a result of the minimization of the objective function. Consequently, we obtain an equivalent Lebesgue space setting for problem (P).

Problem (Pred).

$$\inf \quad \hat{J}(u,e) \qquad \text{over } [u,e]\in L^{N/(N-1)}(\Omega)^N\times Q,$$

where $\hat{J}: L^{N/(N-1)}(\Omega)^N\times Q \to \mathbb{R}\cup\{+\infty\}$ is given by (3.3).

For the implicit regularity constraint we also refer to a similar situation from image restoration problems, where a suitable Fenchel (pre-)dualization of the problem of total bounded variation regularization relies on a similar argument [27]. In the context of perfect plasticity however, the argument additionally hinges on the validity of the safe-load condition.

A second immediate consequence of Lemma 3.3 is that, under the standing assumptions, the existence of solutions to (P) or, equivalently (Pred), follows from standard arguments. The results of this section, including the alternative existence proof for (P) to [15, Thereom 3.3], are summarized in the following theorem.

Theorem 3.4. *The incremental problem (P) of quasi-static evolution in perfect plasticity is equivalent to problem (Pred) in the sense of Lemma 3.1, and (P) has a solution $[\bar{u},\bar{e},\bar{p}]$, which is unique in $\bar{e}$.*

Proof. The equivalence of the problems (P) and (Pred) is a result of Lemma 3.1, Lemma 3.2 and (3.18). For the existence proof, we use the problem formulation (3.7). As a pointwise limit of affine continuous functions, the mapping

$$[u,e] \mapsto \sup_{\substack{\sigma\in S_{\text{ad}}(\Sigma(\mathrm{Div};\Omega)),\\ \sigma\nu=g^n \text{ on } \Gamma_1}} \{-\langle \hat{p}^{n-1},\sigma\rangle - (\sigma,e) - \langle \mathrm{Div}\,\sigma, u\rangle\} - \langle f^n, u\rangle \tag{3.19}$$

is sequentially l.s.c. (lower semicontinuous) in $L^{N/(N-1)}(\Omega)^N \times Q$ equipped with the weak×weak topology. If $u_k \overset{*}{\rightharpoonup} u$ in $BD(\Omega)$ then (u_k) is bounded in $BD(\Omega)$ and fulfills $u_k \to u \in L^1(\Omega)^N$. By the continuous embedding (2.9), each subsequence of (u_k) has a subsequence converging weakly in $L^{N/(N-1)}(\Omega)^N$ to u. Hence, the entire sequence (u_k) weakly converges to u in $L^{N/(N-1)}(\Omega)^N$. Consequently, the mapping from (3.19) is also sequentially l.s.c. in $BD(\Omega) \times Q$ endowed with the weak*×weak topology. Together with the coercivity property in $BD(\Omega)\times Q$ given by Lemma 3.3, the direct method can be applied to prove the existence of a solution $[\bar{u},\bar{e}]$ to (3.7). The existence of a solution to (P) follows by Lemma 3.2, and the uniqueness of $\bar{e}$ is an immediate consequence of the convexity of $\hat{J}$ and the strict convexity of the mapping $e \mapsto (\mathbb{C}e, e)_Q$. □

4. The incremental stress problem as a Fenchel dual problem

In contrast to the original problem, (Pred) defines an unconstrained convex minimization problem in a reflexive Banach space. Therefore, this alternative formulation seems more attractive from a computational point of view. As a nonsmooth convex minimization problem, it is natural to analyze (P) via (Pred) within Fenchel duality theory, for which we refer to [5, 20]. In fact, for the simpler *Hencky plasticity* model, (Lagrangian) duality results linking the stress problem to the strain problem and its relaxation are known; cf. [54, p.251 ff.]. The goal of this paragraph is to demonstrate that the classical incremental stress problem of perfect plasticity, i.e., the time-incremental variant of (2.1), can be derived from the (incremental) primal problem (P) of quasi-static evolution within the theory of Fenchel duality. The result is based on the alternative functional analytic setting provided by the reduced problem formulation (Pred).

4.1. Fenchel duality set-up.

For further reference, we introduce the set of admissible stresses with a given normal component $\tilde{g}$ on Γ_1;

$$S_{\text{ad}}(\tilde{g}) := \{\sigma \in S_{\text{ad}}(\Sigma(\mathrm{Div};\Omega)) : \sigma\nu = \tilde{g} \text{ in } [H_{00}^{-1/2}(\Gamma_1)]^N\}, \tag{4.1}$$

where $\tilde{g} \in [H_{00}^{-1/2}(\Gamma_1)]^N$ is fixed. Note that the regularity of the normal component is ensured by the property $S_{\text{ad}}(\tilde{g}) \subset H(\mathrm{div};\Omega)$.

Under Assumption 2.4, $S_{\text{ad}}(g^n)$ is nonempty, such that the indicator function

$$\begin{gathered} i_{S_{\text{ad}}(g^n)} : \Sigma(\mathrm{Div};\Omega) \to \mathbb{R}\cup\{+\infty\}, \\ i_{S_{\text{ad}}(g^n)}(\sigma) = 0, \text{ if } \sigma \in S_{\text{ad}}(g^n), \quad i_{S_{\text{ad}}(g^n)}(\sigma) = +\infty, \text{ if } \sigma \notin S_{\text{ad}}(g^n), \end{gathered}$$

is proper. We also define the bounded linear operator

$$\Lambda \in \mathcal{L}(L^{N/(N-1)}(\Omega)^N \times Q, \Sigma(\mathrm{Div};\Omega)^*), \quad \Lambda(u,e) := -\mathrm{Div}^* u - e, \tag{4.2}$$

and we set

$$F(u,e) := -\langle f^n, u\rangle + \tfrac{1}{2}(\mathbb{C}e, e), \quad G(\sigma^*) := \sup_{\sigma\in S_{\text{ad}}(g^n)} \langle \sigma^*, \sigma\rangle, \tag{4.3}$$

for $[u,e] \in L^{N/(N-1)}(\Omega)^N \times Q$ and $\sigma^* \in \Sigma(\mathrm{Div};\Omega)^*$. With these definitions, (Pred) takes the equivalent compact form

$$\begin{cases} \min & F(u,e) + G(\Lambda[u,e] - \langle \hat{p}^{n-1}, \,.\, \rangle) \\ \text{over} & [u,e] \in L^{N/(N-1)}(\Omega)^N \times Q. \end{cases} \tag{4.4}$$

Following [5, Chapter 4], the Fenchel dual problem of (4.4) is given by

$$-\inf \quad F^*(-\Lambda^*\sigma) + G^*(\sigma) + \langle \hat{p}^{n-1}, \sigma\rangle \quad \text{over } \sigma \in \Sigma(\mathrm{Div};\Omega), \tag{4.5}$$

where F^* and G^* are the *Fenchel conjugates* pertaining to F and G, respectively.

4.2. **Computation of the Fenchel conjugates.** The *Fenchel* or *convex conjugate* $j^* : X^* \to \mathbb{R} \cup \{+\infty\}$ of a convex function $j : X \to \mathbb{R} \cup \{+\infty\}$ on a Banach space X is defined as the functional

$$j^*(w^*) := \sup_{w \in X} \{\langle w^*, w\rangle - j(w)\}.$$

Observe that $G = i^*_{S_{\text{ad}}(g^n)}$, and a straightforward computation leads to

$$F^*(u^*, e^*) = i_{\{-f^n\}}(u^*) + \tfrac{1}{2}(\mathbb{C}^{-1}e^*, e^*), \quad G^*(\sigma) = i^{**}_{S_{\text{ad}}(g^n)}(\sigma),$$

for $[u^*, e^*] \in [L^N(\Omega)]^N \times Q$ and $\sigma \in \Sigma(\mathrm{Div};\Omega)$. The adjoint of Λ is given by

$$\Lambda^*\sigma = [-\operatorname{Div}\sigma, -\sigma] \ \in [L^N(\Omega)]^N \times Q. \tag{4.6}$$

Since $S_{\text{ad}}(g^n) \subset \Sigma(\mathrm{Div};\Omega)$ is nonempty, convex and closed, it holds that

$$G^* = i^{**}_{S_{\text{ad}}(g^n)} = i_{S_{\text{ad}}(g^n)},$$

such that (4.5) amounts to the following problem.

Problem (D).

$$\begin{cases} \inf & \tfrac{1}{2}(\mathbb{C}^{-1}\sigma, \sigma) + \langle \hat{p}^{n-1}, \sigma\rangle \\ \text{s.t.} & -\operatorname{Div}\sigma = f^n, \ \sigma\nu = g^n \text{ on } \Gamma_1, \ \sigma \in S_{\text{ad}} \\ \text{over} & \sigma \in \Sigma(\mathrm{Div};\Omega). \end{cases}$$

The definition (3.4) of $\hat{p}^{n-1}$ allows to reformulate (D) as a problem in the larger Hilbert space Q with the help of the adjoint operator to $\varepsilon \in \mathcal{L}(V,Q)$.

$$\begin{cases} \inf & \tfrac{1}{2}(\mathbb{C}^{-1}\sigma, \sigma) - (\mathbb{C}^{-1}\sigma_{n-1}, \sigma) \\ \text{s.t.} & \varepsilon^*\sigma = l^n \ \text{ in } V^*, \ \sigma \in S_{\text{ad}}, \\ \text{over} & \sigma \in Q \end{cases} \tag{4.7}$$

We note that problem (D), or (4.7), is exactly the stress problem (Problem 2.2) of perfect plasticity in incremental form resulting from an implicit Euler time discretization;

$$\dot{\sigma}(t_n) \approx \tfrac{\sigma(t_n) - \sigma(t_{n-1})}{t_n - t_{n-1}}.$$

We summarize the result in the following theorem.

Theorem 4.1. *Let the applied forces f and g fulfill Assumption 2.4. A Fenchel dual problem of the time-incremental problem of quasi-static evolution in perfect plasticity in reduced form (problem (Pred)) is given by (D), which is the stress problem in incremental form. There is no duality gap between primal and dual problem, i.e., it holds that*

$$\inf \textit{(Pred)} = -\inf \textit{(D)}. \tag{4.8}$$

Proof. In order to prove (4.8), it suffices that the following constraint qualification is fulfilled;

$$-\hat{p}^{n-1} \in \operatorname{int}(\operatorname{dom} G - \Lambda \operatorname{dom} F); \tag{4.9}$$

cf. [5, Theorem 1, p.221]. The validity of (4.9) can be seen as follows: From the definition of the adjoint (4.6), it follows directly that Λ^* is injective, such that the range of Λ is dense in $\Sigma(\mathrm{Div};\Omega)^*$. Since the range of Λ^* is closed, the surjectivity of Λ follows from the closed range theorem. Together with $\operatorname{dom} G \neq \emptyset$, $\operatorname{dom} F = L^{N/(N-1)}(\Omega)^N \times Q$ and the surjectivity of Λ, one obtains

$$\operatorname{dom} G - \Lambda \operatorname{dom} F = \Sigma(\mathrm{Div};\Omega)^*,$$

such that the constraint qualification (4.9) is satisfied. □

We stress that the proof of Theorem 4.1 requires the correct choice of the topologies for the domain and image space of the operator Λ. In fact, using the reduced formulation (P$_{\text{red}}$) in the Lebesgue space setting we avoid explicit incorporation of the space $BD(\Omega)$ for the displacement.

4.3. **Primal-dual optimality conditions.** Under Assumption 2.4, the admissible set of (D) is nonempty and it follows from standard arguments that (D) has a unique solution $\bar{\sigma} \in \Sigma(\mathrm{Div};\Omega)$. By virtue of (4.8), saddle points $[\bar{u},\bar{e};\bar{\sigma}]$, where $[\bar{u},\bar{e}]$ solves (P$_{\text{red}}$), are characterized by the following primal-dual optimality conditions (see [20, III, Remark 4.2], for instance);

$$\bar{\sigma} \in S_{\text{ad}}(g^n), \quad \mathrm{Div}\,\bar{\sigma} = -f^n, \quad \mathbb{C}\bar{e} = \bar{\sigma}, \tag{4.10}$$

$$-\hat{p}^{n-1} - \mathrm{Div}^*\,\bar{u} - \bar{e} \in N_{S_{\text{ad}}(g^n)}(\bar{\sigma}). \tag{4.11}$$

Here, $N_{S_{\text{ad}}(g^n)}(\bar{\sigma})$ denotes the normal cone to the set $S_{\text{ad}}(g^n) \subset \Sigma(\mathrm{Div};\Omega)$ at $\bar{\sigma}$. Note that (4.11) is equivalent to

$$\langle \bar{u} - u^{n-1}, \mathrm{Div}\,\tilde{\sigma} - \mathrm{Div}\,\bar{\sigma}\rangle + (\bar{e} - e^{n-1}, \tilde{\sigma} - \bar{\sigma}) \geq 0, \quad \forall\, \tilde{\sigma} \in S_{\text{ad}}(g^n); \tag{4.12}$$

that is, the optimality system (4.10)-(4.11) represents precisely the time-discretized version of the stress problem (Problem 2.1). Whereas [15, Theorem 3.6(c)] only derives (4.10) as a necessary optimality condition for a solution $[\bar{u}, \mathbb{C}^{-1}\bar{\sigma}, \bar{p}]$ to the primal problem (P), our result shows that by additionally incorporating the normal cone condition (4.12), one obtains necessary and sufficient optimality conditions for the time-discretized primal problem in quasi-static perfect plasticity. A rigorous Fenchel duality result for the time-discrete primal problem of perfect plasticity and the dual stress problem has thus been established.

5. A NEW ALGORITHMIC SCHEME

5.1. **A modified visco-plastic regularization.** In the remainder of this paper we intend to design an infinite-dimensional algorithm to solve the time-incremental problem of perfect plasticity based on a new regularization scheme. A classical approach to the problem of perfect plasticity is the visco-plastic regularization, which is essentially a Moreau-Yosida regularization $i^{\mu}_{\mathbb{K}_0}$ of the indicator function $i_{\mathbb{K}_0}$ associated with the constraint $\operatorname{dev}\sigma(x) \in \mathbb{K}_0$, such that the inclusion in (2.8) is replaced by the smooth equation

$$\dot{p} = {i^{\mu}_{\mathbb{K}_0}}'(\sigma), \quad \text{with } i^{\mu}_{\mathbb{K}_0}(\sigma) := \frac{\mu}{2} \inf_{\substack{\tilde{\sigma} \in \mathbb{M}^{N\times N}:\\ \operatorname{dev}\tilde{\sigma} \in \mathbb{K}_0}} \left\{ |\tilde{\sigma} - \sigma|_F^2 \right\}.$$

The basis for the existence proofs in [34, 53] is that perfect plasticity can be characterized as the limit of visco-plasticity as $\mu \to +\infty$. On the level of the weak formulation in terms of the stress (4.7), this approach essentially corresponds to a Moreau-Yosida regularization of the constraint $\sigma \in S_{\text{ad}}$ in the space Q. In [47, Lemma 3.8], it is shown that the visco-plastic regularization is equivalent to a problem of plasticity with kinematic hardening, where the hardening modulus depends on the regularization parameter γ. As discussed in [30], the problem of hardening plasticity requires further regularization techniques in order to allow for an efficient infinite-dimensional solver that converges mesh-independently upon discretization. For these reasons, it appears to be worthwhile to consider an alternative regularization scheme that differs from a vanishing hardening approach.

In this section we propose a primal modification that combines the usual visco-plastic regularization of the flow law with a Tikhonov regularization of the objective functional in (Pred) that maintains the original function space setting. As it turns out, this approach allows to recover a one-to-one relation between the approximations of the primal variable pair $[u, p]$ and the solution of a suitably modified version of the incremental stress problem (D) in the original infinite-dimensional setting. In particular, the approximations of u are not assumed to be elements of the Sobolev space V.

On the level of the primal problem, consider the following family of regularized problems induced by a sequence of positive parameters $\mu > 0$.

Problem (P$_\mu$).

$$\begin{cases} \inf & \hat{J}_\mu(u,e) \\ \text{over} & [u,e] \in L^{N'}(\Omega)^N \times Q, \end{cases}$$

where

$$\hat{J}_\mu(u,e) := \tfrac{1}{\mu N'}\|u\|^{N'}_{L^{N'}(\Omega)^N} - \langle f^n, u\rangle + \tfrac{1}{2}(\mathbb{C}e, e) + \sup_{\substack{\sigma\in\Sigma(\mathrm{Div};\Omega),\\ \sigma\nu = g^n \text{ on } \Gamma_1}} \{-\langle \hat{p}^{n-1}, \sigma\rangle - (\sigma, e) - \langle \mathrm{Div}\,\sigma, u\rangle - i^\mu_{S_{\mathrm{ad}}}(\sigma)\}.$$

Here, $i^\mu_{S_{\mathrm{ad}}}$ is defined as the Moreau-Yosida regularization of $i_{S_{\mathrm{ad}}}$ as a mapping defined on Q;

$$i^\mu_{S_{\mathrm{ad}}}(\sigma) := \tfrac{\mu}{2} \inf_{\tilde{\sigma}\in S_{\mathrm{ad}}} \|\sigma - \tilde{\sigma}\|^2_Q.$$

Note that, according to (2.9), it holds $BD(\Omega) \hookrightarrow L^{N'}(\Omega)^N$, where $N' := N/(N-1)$. Existence and uniqueness of a solution to Problem (P$_\mu$) then follows by standard arguments from convex analysis as summarized in the following proposition.

Proposition 5.1. *Let the safe-load condition (Assumption 2.4) be fulfilled. Then Problem (P$_\mu$) admits a unique solution $[u_\mu, e_\mu]$, which satisfies $u_\mu \in BD(\Omega)$, $u_\mu\nu = 0$ on Γ_0 and $\mathrm{div}\, u_\mu = \mathrm{tr}\, e_\mu$ in Ω.*

Proof. The function

$$[u,e] \mapsto \sup_{\substack{\sigma\in\Sigma(\mathrm{Div};\Omega),\\ \sigma\nu = g^n \text{ on } \Gamma_1}} \{-\langle \hat{p}^{n-1}, \sigma\rangle - (\sigma, e) - \langle \mathrm{Div}\,\sigma, u\rangle - i^\mu_{S_{\mathrm{ad}}}(\sigma)\}$$

represents the pointwise supremum of a sequence of affine functions on $L^{N'}(\Omega)^N \times Q$ and as such, it is convex and weakly l.s.c. in $L^{N'}(\Omega)^N \times Q$. Under Assumption 2.4 it is also proper. The additional strictly convex term

$$\tfrac{1}{N'\mu}\|u\|^{N'}_{L^{N'}(\Omega)^N} \tag{5.1}$$

yields the coercivity of $\hat{J}_\mu$ on $L^{N'}(\Omega)^N \times Q$. Existence and uniqueness of a solution now follows from the direct method. The regularity statement $\varepsilon(u) \in M(\Omega; \mathbb{M}^{N\times N})$ follows under Assumption 2.4 by

$$\begin{aligned} &\sup_{\substack{\sigma\in\Sigma(\mathrm{Div};\Omega),\\ \sigma\nu = g^n \text{ on } \Gamma_1}} \{-\langle \hat{p}^{n-1}, \sigma\rangle - (\sigma, e) - \langle \mathrm{Div}\,\sigma, u\rangle - i^\mu_{S_{\mathrm{ad}}}(\sigma)\} - \langle f^n, u\rangle \\ &\quad\geq \sup_{\substack{\sigma\in S_{\mathrm{ad}}(\Sigma(\mathrm{Div};\Omega)),\\ \sigma\nu = g^n \text{ on } \Gamma_1}} \{-\langle \hat{p}^{n-1}, \sigma\rangle - (\sigma, e) - \langle \mathrm{Div}\,\sigma, u\rangle\} - \langle f^n, u\rangle, \end{aligned} \tag{5.2}$$

together with the estimate (3.14). Since $u_\mu \in BD(\Omega)$, the validity of the plastic incompressibility conditions $u_\mu \cdot \nu = 0$ on Γ_0 and $\operatorname{div} u_\mu = \operatorname{tr} e_\mu$ can be deduced from (5.2) as in the proof of Lemma 3.2. ☐

Unlike the case of the visco-plastic regularization, we do neither dispose of an explicit problem formulation of (P$_\mu$) in terms of u nor is it possible to prove that the optimal displacement u_μ is an element of the Sobolev space V. Instead, (P$_\mu$) does not impose a higher strain regularity than the initial problem (P) and therefore it does not fall into the realm of hardening plasticity.

It can also be expected that (P$_\mu$) yields a close approximation of (Pred), at least for large μ. Before discussing this issue, we proceed by computing an associated Fenchel dual problem that turns out to be a penalized version of the incremental stress problem.

Problem (D$_\mu$).

$$\begin{cases} \inf & J^*_\mu(\sigma) \\ \text{s.t.} & \sigma\nu = g^n \text{ on } \Gamma_1 \\ \text{over} & \sigma \in \Sigma(\mathrm{Div};\Omega). \end{cases}$$

with

$$J^*_\mu(\sigma) := \tfrac{1}{2}(\mathbb{C}^{-1}\sigma,\sigma) + \langle \hat{p}^{n-1},\sigma\rangle + \tfrac{\mu^{N-1}}{N}\|\operatorname{Div}\sigma + f^n\|^N_{L^N(\Omega)^N} + i^\mu_{S_{\text{ad}}}(\sigma).$$

Proposition 5.2. *Let the safe-load condition (Assumption 2.4) be fulfilled. Then a Fenchel dual problem to (P$_\mu$) is given by the modified stress problem (D$_\mu$). Moreover, (D$_\mu$) has a unique solution σ_μ and there is no duality gap, i.e.,*

$$\min \textit{(P}_\mu\textit{)} = -\min \textit{(D}_\mu\textit{)}. \tag{5.3}$$

Proof. Since

$$\sigma \mapsto \tfrac{1}{2}(\mathbb{C}^{-1}\sigma,\sigma) + \tfrac{\mu^{N-1}}{N}\|\operatorname{Div}\sigma + f^n\|^N_{L^N(\Omega)^N}$$

defines a strictly convex and coercive functional on $\Sigma(\mathrm{Div};\Omega)$, existence and uniqueness of a solution σ_μ to (D$_\mu$) follows from standard arguments.

Using the linear operator Λ from (4.2) we rewrite (P$_\mu$) in compact form as

$$\min \quad F_\mu(u,e) + G_\mu(\Lambda[u,e] - \hat{p}^{n-1}) \quad \text{over } [u,e] \in L^{N'}(\Omega)^N \times Q, \tag{5.4}$$

with

$$F_\mu : L^{N'}(\Omega)^N \times Q \to \mathbb{R}\cup\{\infty\}, \quad F_\mu(u,e) := \tfrac{1}{\mu N'}\|u\|^{N'}_{L^{N'}(\Omega)^N} - \langle f^n, u\rangle + \tfrac{1}{2}(\mathbb{C}e,e),$$
$$G_\mu : \Sigma(\mathrm{Div};\Omega)^* \to \mathbb{R}\cup\{\infty\}, \quad G_\mu(\sigma^*) := \sup_{\substack{\sigma\in\Sigma(\mathrm{Div};\Omega),\\ \sigma\nu = g^n \text{ on } \Gamma_1}} \{\langle\sigma^*,\sigma\rangle - i^\mu_{S_{\text{ad}}}(\sigma)\}.$$

An application of [20, I, Remark 4.1] leads to

$$F^*_\mu(u^*,e^*) = \tfrac{\mu^{N-1}}{N}\ \|u^* + f^n\|^N_{L^N(\Omega)^N} + \tfrac{1}{2}(\mathbb{C}^{-1}e^*,e^*),$$

for all $[u^*,e^*] \in L^N(\Omega)^N \times Q$. Moreover, it holds that $G_\mu(\sigma^*) = \tilde{G}^*_\mu(\sigma^*)$ for

$$\tilde{G}_\mu(\sigma) := i_{\Sigma_{g^n}(\mathrm{Div};\Omega)}(\sigma) + i^\mu_{S_{\text{ad}}}(\sigma), \quad \sigma \in \Sigma(\mathrm{Div};\Omega),$$

where

$$\Sigma_{\tilde{g}}(\mathrm{Div};\Omega) := \{\sigma\in\Sigma(\mathrm{Div};\Omega) : \sigma\nu = \tilde{g} \text{ on } \Gamma_1\}, \quad \tilde{g} \in H^{-1/2}_{00}(\Gamma_1).$$

Since $\tilde{G}_\mu$ is convex, l.s.c. and proper, one obtains

$$G^*_\mu = \tilde{G}_\mu = i_{\Sigma_{g^n}(\mathrm{Div};\Omega)} + i^\mu_{S_{\text{ad}}}.$$

The Fenchel dual problem of (P$_\mu$) corresponding to this setting is given by

$$-\inf \quad F_\mu^*(-\Lambda^*\sigma) + G_\mu^*(\sigma) + \langle \hat{p}^{n-1}, \sigma\rangle, \tag{5.5}$$

which is exactly problem (D$_\mu$). Under the safe-load condition, the validity of (4.9) can be verified as in Theorem 4.1, such that (5.3) holds. ☐

Hence, adding the strictly convex term (5.1) to the objective function $\hat{J}$ in (Pred) results in a penalty approach to the mechanical equilibrium constraint $-\operatorname{Div}\sigma = f^n$ in the space $L^N(\Omega)^N$. Standard properties of the Moreau-Yosida regularization further ensure that the objective function J_μ^* is convex and continuously Fréchet differentiable as a functional on $\Sigma(\operatorname{Div};\Omega)$. Since both problems are uniquely solvable, we retrieve a one-to-one relation between regularized stresses and strains via the primal-dual optimality conditions for the saddle point $[u_\mu, e_\mu; \sigma_\mu] \in BD(\Omega)\times Q\times\Sigma(\operatorname{Div};\Omega)$; see [20, III, Remark 4.2]. In fact, $[u_\mu, e_\mu; \sigma_\mu]$ can be characterized by the existence of $\lambda_\mu \in \Sigma(\operatorname{Div};\Omega)^*$ such that

$$\mathbb{C}e_\mu = \sigma_\mu \text{ in } Q, \quad \sigma_\mu\nu = g^n \text{ on } \Gamma_1 \tag{5.6}$$

$$|u_\mu|^{1/(N-1)} \star \operatorname{sign}(u_\mu) = \mu(f^n + \operatorname{Div}\sigma_\mu) \text{ in } \Omega \tag{5.7}$$

$$-\hat{p}^{n-1} - \operatorname{Div}^* u_\mu - (\mathbb{C}^{-1} + \mu\,\mathrm{id})\sigma_\mu + \mu\pi_{S_{\mathsf{ad}}}(\sigma_\mu) - \lambda_\mu = 0, \tag{5.8}$$

$$\lambda_\mu \in N_{\Sigma_{g^n}(\operatorname{Div};\Omega)}(\sigma_\mu), \tag{5.9}$$

where $\pi_{S_{\mathsf{ad}}}$ denotes the projection on S_{ad} in the space Q, and $N_{\Sigma_{g^n}(\operatorname{Div};\Omega)}(\sigma_\mu)$ is the normal cone at σ_μ to $\Sigma_{g^n}(\operatorname{Div};\Omega) \subset \Sigma(\operatorname{Div};\Omega)$. Here, (5.7) and the application of 'sign' have to be understood componentwise, where

$$|a|^p := [|a_1|^p, \ldots, |a_d|^p], \quad a\star b := [a_1b_1, \ldots, a_db_d]$$

denote Hadamard products for vectors $a, b \in \mathbb{R}^d$ in (5.7).

This shows that the displacement u_μ can be easily computed from the solution σ_μ of (D$_\mu$) using (5.7). In contrast to the primal problem (P$_\mu$), which is only given in inf-sup-form, the dual problem is again given explicitly. This facilitates the analysis of the consistency of the regularization with regard to the limit problems (P) and (D).

Theorem 5.3 (Consistency)**.** *Let the safe-load condition (Assumption 2.4) be satisfied. Then the following assertions hold true.*

(i) The sequence of approximate elastic strains (e_μ) fulfills

$$e_\mu \to \bar{e} \quad \text{in } Q, \quad \text{for } \mu\to\infty.$$

The sequence of approximate displacements (u_μ) is bounded in $BD(\Omega)$ and for any limit $\bar{u} \in BD(\Omega)$ of a weakly*-convergent subsequence of $(u_\mu) \subset BD(\Omega)$, it holds that $[\bar{u}, \bar{e}]$ is a solution of (Pred).

(ii) The sequence of approximate stresses (σ_μ) fulfills

$$\sigma_\mu \to \bar{\sigma} \quad \text{in } \Sigma(\operatorname{Div};\Omega), \quad \text{for } \mu\to\infty.$$

Proof. Step 1 (dual problem).

First observe that the sequence of minimizers (σ_μ), whose existence is guaranteed by Proposition 5.2, is bounded in $\Sigma(\operatorname{Div};\Omega)$. Indeed, we have

$$\begin{aligned}
J_\mu^*(\sigma_\mu) &\geq \tfrac{1}{2}(\mathbb{C}^{-1}\sigma_\mu,\sigma_\mu) + \langle \hat{p}^{n-1},\sigma_\mu\rangle + \tfrac{\mu^{N-1}}{N}\|\operatorname{Div}\sigma_\mu + f^n\|_{L^N(\Omega)^N}^N \\
&\geq \tfrac{\kappa_{\mathbb{C}^{-1}}}{2}\|\sigma_\mu\|_Q^2 - c\|\sigma_\mu\|_Q - c\|\operatorname{Div}\sigma_\mu\|_{L^N(\Omega)^N} + \tfrac{\mu^{N-1}}{N}\|\operatorname{Div}\sigma_\mu + f^n\|_{L^N(\Omega)^N}^N \\
&\geq \tfrac{\kappa_{\mathbb{C}^{-1}}}{2}\|\sigma_\mu\|_Q^2 - c\|\sigma_\mu\|_Q - c\|\operatorname{Div}\sigma_\mu + f^n\|_{L^N(\Omega)^N} - c\|f^n\|_{L^N(\Omega)^N} \\
&\qquad + \tfrac{\mu^{N-1}}{N}\|\operatorname{Div}\sigma_\mu + f^n\|_{L^N(\Omega)^N}^N,
\end{aligned} \tag{5.10}$$

where $c := \|p^{n-1}\|_{\Sigma(\operatorname{Div};\Omega)^*}$ and $\kappa_{\mathbb{C}^{-1}} > 0$ is a constant that fulfills

$$(\mathbb{C}^{-1}\sigma,\sigma) \geq \kappa_{\mathbb{C}^{-1}}\|\sigma\|_Q^2, \quad \forall\, \sigma\in Q. \tag{5.11}$$

On the other hand, it holds that

$$J_\mu^*(\sigma_\mu) \leq J_\mu^*(\bar\sigma) = \tfrac{1}{2}(\mathbb{C}^{-1}\bar\sigma,\bar\sigma) + \langle \hat{p}^{n-1},\bar\sigma\rangle, \quad \forall\, \mu > 0. \tag{5.12}$$

Together with (5.10), this implies that $(\sigma_\mu)\subset\Sigma(\operatorname{Div};\Omega)$ is bounded.

Under the safe-load condition (Assumption 2.4), the objective function J_μ^* of the dual problem (D$_\mu$) is proper, weakly l.s.c. and pointwise monotonically increasing for all $\mu > 0$. The pointwise limit is given by

$$\lim_{\mu\to\infty}\left(J_\mu^*(\sigma) + i_{\Sigma_{g^n}(\operatorname{Div};\Omega)}(\sigma)\right) = \tfrac{1}{2}(\mathbb{C}^{-1}\sigma,\sigma) + \langle \hat{p}^{n-1},\sigma\rangle, \tag{5.13}$$

in the case where $\sigma\in S_{\text{ad}}(g^n)$, $-\operatorname{Div}\sigma = f^n$, and

$$\lim_{\mu\to\infty}\left(J_\mu^*(\sigma) + i_{\Sigma_{g^n}(\operatorname{Div};\Omega)}(\sigma)\right) = +\infty, \tag{5.14}$$

else. An application of [14, Prop. 5.4] yields that (5.13) and (5.14) also hold as Γ-limits in the space $\Sigma(\operatorname{Div};\Omega)$ endowed with the weak topology. It follows that each weak limit point of (σ_μ) in $\Sigma(\operatorname{Div};\Omega)$ is the solution $\bar\sigma$ of (D); see [14, Corollary 7.20]. By uniqueness, this also holds for the entire sequence, i.e., $\sigma_\mu \rightharpoonup \bar\sigma$ in $\Sigma(\operatorname{Div};\Omega)$. The strong convergence of (σ_μ) can be deduced as follows.

$$\begin{aligned}
0 &\leq \liminf_{\mu\to\infty}\left(\tfrac{\mu^{N-1}}{N}\|\operatorname{Div}\sigma_\mu + f^n\|_{L^N(\Omega)^N}^N + i_{S_{\text{ad}}}^\mu(\sigma_\mu)\right) \\
&\leq \limsup_{\mu\to\infty}\left(\tfrac{\mu^{N-1}}{N}\|\operatorname{Div}\sigma_\mu + f^n\|_{L^N(\Omega)^N}^N + i_{S_{\text{ad}}}^\mu(\sigma_\mu)\right) \\
&\leq \limsup_{\mu\to\infty}\left(\tfrac{\kappa_{\mathbb{C}^{-1}}}{2}\|\sigma_\mu-\bar\sigma\|_Q^2 + \tfrac{\mu^{N-1}}{N}\|\operatorname{Div}\sigma_\mu + f^n\|_{L^N(\Omega)^N}^N + i_{S_{\text{ad}}}^\mu(\sigma_\mu)\right) \\
&\leq \limsup_{\mu\to\infty}\left(J_\mu^*(\sigma_\mu) - \langle \hat{p}^{n-1},\sigma_\mu\rangle + \tfrac{1}{2}(\mathbb{C}^{-1}\bar\sigma,\bar\sigma) - (\mathbb{C}^{-1}\sigma_\mu,\bar\sigma)\right) \\
&\leq \limsup_{\mu\to\infty}\left((\mathbb{C}^{-1}\bar\sigma,\bar\sigma) + \langle \hat{p}^{n-1},\bar\sigma-\sigma_\mu\rangle - (\mathbb{C}^{-1}\sigma_\mu,\bar\sigma)\right) = 0,
\end{aligned}$$

where we use (5.11),(5.12) and $\sigma_\mu \rightharpoonup \bar\sigma$ in $\Sigma(\operatorname{Div};\Omega)$. This entails that (σ_μ) converges strongly to $\bar\sigma$ in $\Sigma(\operatorname{Div};\Omega)$ and that

$$J_\mu^*(\sigma_\mu) \to \tfrac{1}{2}(\mathbb{C}^{-1}\bar\sigma,\bar\sigma) + \langle \hat{p}^{n-1},\bar\sigma\rangle. \tag{5.15}$$

We immediately infer that $e_\mu = \mathbb{C}^{-1}\sigma_\mu \to \mathbb{C}^{-1}\bar\sigma = \bar e$ in Q.

Step 2 (primal problem). Owing to the upper bound

$$\begin{aligned}
c &\geq \hat{J}_\mu(u^{n-1},e^{n-1}) \geq \hat{J}_\mu(u_\mu,e_\mu) \geq \hat{J}(u_\mu,e_\mu) \\
&\geq c_0 - c_1\|e_\mu\| + \kappa_{\mathbb{C}}\|e_\mu\|^2 \\
&\qquad + \rho\max(\|\varepsilon(u_\mu)\|_{M(\Omega)}, -\|\varepsilon(u_\mu)\|_{M(\Omega)} + \tfrac{1}{\sqrt{2}}\|u_\mu\|_{L^1(\Gamma_0;\mathbb{R}^N)}),
\end{aligned}$$

where the lower estimate follows from Lemma 3.3, one obtains the uniform boundedness of $(e_\mu) \subset Q$ and $(u_\mu) \subset BD(\Omega)$. Consequently, one may extract a subsequence also denoted by $([u_\mu, e_\mu])$, such that

$$u_\mu \overset{*}{\rightharpoonup} \tilde{u} \text{ in } BD(\Omega), \quad e_\mu \rightharpoonup \tilde{e} \text{ in } Q.$$

Using the sequential weak$^*\times$weak lower semicontinuity of $\hat{J}$, (cf. the proof of Theorem 3.4) and (5.3), one obtains

$$\begin{aligned}\hat{J}(\tilde{u},\tilde{e}) &\le \liminf_{\mu\to\infty} \hat{J}(u_\mu, e_\mu) \le \liminf_{\mu\to\infty} \hat{J}_\mu(u_\mu, e_\mu) \\ &= \liminf_{\mu\to\infty}\ \min\,(\mathsf{P}_\mu) = -\limsup_{\mu\to\infty} \min\,(\mathsf{D}_\mu),\end{aligned} \tag{5.16}$$

Using (5.15), one obtains

$$\limsup_{\mu\to\infty} \min\,(\mathsf{D}_\mu) = \lim_{\mu\to\infty}\ \min\,(\mathsf{D}_\mu) = \min(D),$$

With the help of (4.8), the estimate (5.16) implies that

$$\hat{J}(\tilde{u},\tilde{e}) \le -\min\,(\mathsf{D}) = \min\,(\mathsf{Pred}),$$

i.e., $[\tilde{u},\tilde{e}]$ solves (Pred). □

5.2. **An infinite-dimensional semismooth Newton method.** This section aims to provide a theoretical framework for an efficient infinite-dimensional algorithmic scheme to solve the regularized problems (D_μ) for a fixed parameter $\mu \gg 0$ based on the semismooth Newton method [13, 32]. Owing to Theorem 5.3 it is justified to make the assumption that (D_μ) represents a good approximation of (P). Using the primal-dual optimality condition (5.7), it is further possible to retrieve $[u_\mu, e_\mu]$ by solving (D_μ). In this section, we make the restrictive assumption that $N = 2$, which implies that the incremental stress problem as well as its regularization is posed in the Hilbert space $H(\mathrm{Div};\Omega)$.

For simplicity, we henceforth also assume that $g^n = 0$. From a theoretical viewpoint this does not impose a restriction since Korn's inequality ensures that there exists an element $\xi = \varepsilon(\hat{u}) \in H(\mathrm{Div};\Omega)$, $\hat{u} \in V$, that fulfills

$$-\operatorname{Div}\xi = f^n, \quad \xi\nu = g^n \text{ on } \Gamma_1. \tag{5.17}$$

In the usual way, one may then use ξ to transform (D_μ) into an equivalent problem with a homogeneous normal trace condition. For $g^n = 0$, problem (D_μ) reads

$$(\tilde{D}_\mu) \qquad \begin{cases} \inf & \frac{1}{2}(\mathbb{C}^{-1}\sigma,\sigma) + \langle \hat{p}^{n-1},\sigma\rangle + \frac{\mu}{2}\|\operatorname{Div}\sigma + f^n\|^2_{L^2(\Omega)^2} + i^{\mu}_{S_{\mathrm{ad}}}(\sigma) \\ \text{s.t.} & \sigma\nu = 0 \text{ on } \Gamma_1, \\ \text{over} & \sigma \in H(\mathrm{Div};\Omega). \end{cases}$$

5.2.1. *The semismooth Newton method.* The semismooth Newton method relies on the notion of Newton differentiability, which can be found in [13, 32].

Definition 5.4 (Newton differentiability)**.** Let X, Y be Banach spaces and $U \subset X$ be an open set. A mapping $F : U \to Y$ is called Newton differentiable in U if there exists a family of mappings $G_F : U \to \mathcal{L}(X,Y)$ which satisfy

$$\|F(x+h) - F(x) - G_F(x+h)h\|_Y = o(\|h\|_X), \quad \|h\|_X \to 0,$$

for all $x \in U$.

Provided $G_F(x)$ is invertible for all $x \in U$, the corresponding generalized Newton method to solve $F(x) = 0$ for Newton differentiable F is defined iteratively by

$$x^{(j+1)} = x^{(j)} - G_F(x^{(j)})^{-1}F(x^{(j)}), \quad x^{(0)} \in U. \tag{5.18}$$

Following [32, Theorem 1.1], the sequence $(x^{(j)})$ converges locally at a superlinear rate if $\{G_F(x^{(j)})^{-1} : k \in \mathbb{N}\}$ is uniformly bounded. Moreover, the convergence rates are mesh independent upon discretization, which means that the convergence quotients remain stable for sufficiently small mesh width. In practice, the mesh independence of an algorithm ensures that the iteration numbers stay bounded as the mesh width gets finer. We refer to [26, 31] for detailed mesh independence results for the semismooth Newton method.

At this point, it should be emphasized that mesh-independent convergence requires the Newton differentiability of the operator F with respect to the original (infinite-dimensional) setting, which in turn necessitates a norm gap with respect to domain and image space of F; cf., for instance, Lemma 5.8 below. In the context of (D$_\mu$) however, the problem lacks the necessary norm gap since $H(\mathrm{Div};\Omega)$ does not embed into L^p-spaces for $p > 2$. Therefore, a direct application of the semismooth Newton method to the optimality conditions associated with the discrete formulation of (D$_\mu$) results in a mesh-dependent solver.

5.2.2. *Tikhonov regularization.* In order to overcome this drawback, we suggest to replace problem (D$_\mu$) by a Tikhonov-regularized problem induced by a continuous and elliptic symmetric bilinear form

$$b(\,.\,,\,.\,) : H^1(\Omega;\mathbb{M}^{2\times 2}) \times H^1(\Omega;\mathbb{M}^{2\times 2}) \to \mathbb{R}$$

on the dense Hilbert subspace

$$H^1(\Omega;\mathbb{M}^{2\times 2}) \hookrightarrow H(\mathrm{Div};\Omega)$$

of $\mathbb{M}^{2\times 2}$-valued functions on Ω with distributional partial derivatives in Q, and we denote by

$$B \in \mathcal{L}(H^1(\Omega;\mathbb{M}^{2\times 2}), H^1(\Omega;\mathbb{M}^{2\times 2})^*)$$

the bounded linear operator associated with the bilinear form b. Note that the symmetry condition in the definition of the space $H^1(\Omega;\mathbb{M}^{2\times 2})$ can be easily imposed using a simple parametrization. We now contemplate the following approximation of the regularized incremental stress problem ($\tilde{D}_\mu$), which is induced by a sequence of positive parameters (γ).

$$(\mathrm{D}_{\mu,\gamma}) \qquad \begin{cases} \inf & J^*_{\mu,\gamma}(\sigma) \\ \text{s.t.} & \sigma\nu = 0 \text{ on } \Gamma_1, \\ \text{over} & \sigma \in H^1(\Omega,\mathbb{M}^{2\times 2}), \end{cases}$$

where

$$J^*_{\mu,\gamma}(\sigma) := \tfrac{1}{2}(\mathbb{C}^{-1}\sigma,\sigma) + \langle \hat{p}^{n-1},\sigma\rangle + \tfrac{\mu}{2}\|\operatorname{Div}\sigma + f^n\|^2_{L^2(\Omega)^2} + i^\mu_{S_{\mathrm{ad}}}(\sigma) + \tfrac{1}{2\gamma}b(\sigma,\sigma).$$

The assumptions on b ensure that each problem (D$_{\mu,\gamma}$) has a unique solution, which is henceforth denoted by $\sigma_{\mu,\gamma}$. The problem (D$_{\mu,\gamma}$) further promises a good approximation of (D$_\mu$) at least for large γ. In fact, in order to relate the problems (D$_{\mu,\gamma}$) to (D$_\mu$) it is necessary to extend the density property

$$\overline{C_c^\infty(\Omega;\mathbb{M}^{N\times N})}^{H(\mathrm{Div};\Omega)} = H_0(\mathrm{Div};\Omega), \tag{5.19}$$

from [23, I, Theorem 2.6], to problems with mixed boundary conditions. For this purpose we define the appropriate subspace

$$H_{0,\Gamma_1}(\mathrm{Div};\Omega) := \{\sigma \in H(\mathrm{Div};\Omega) : \sigma\nu = 0 \text{ on } \Gamma_1\}$$

of $H(\mathrm{Div};\Omega)$-functions whose normal component vanishes on Γ_1 in the sense of the space $H_{00}^{-1/2}(\Gamma_1)$. We further make the following technical assumption on the boundary portion Γ_0.

Assumption 5.5. *The splitting of $\partial\Omega = \Gamma_0 \cup \Gamma_1 \cup \partial\Gamma_0$ is regular enough to ensure that the density result*

$$\overline{C^\infty_{0,\Gamma_1}(\overline{\Omega})}^{H^1(\Omega)} = H^1_{0,\Gamma_1}(\Omega) \tag{5.20}$$

for $H^1_{0,\Gamma_1}(\Omega) = \{u \in H^1(\Omega) : u = 0 \text{ on } \Gamma_1\}$ *and*

$$C^\infty_{0,\Gamma_1}(\overline{\Omega}) := \{\varphi \in C^\infty(\overline{\Omega}), \varphi = 0 \text{ on } \Gamma_1\} \tag{5.21}$$

holds true.

According to [17, 9], condition (5.20) may only be violated by some degenerate Γ_0, such that, from a practical point of view, Assumption 5.5 does not represent a restriction.

Lemma 5.6. *Let $N \in \mathbb{N}$ and suppose Assumption 5.5 holds true. Then the density property*

$$\overline{C^\infty_{0,\Gamma_1}(\overline{\Omega}; \mathbb{M}^{N\times N})}^{H(\mathrm{Div};\Omega)} = H_{0,\Gamma_1}(\mathrm{Div}; \Omega)$$

is satisfied, where

$$C^\infty_{0,\Gamma_1}(\overline{\Omega}; \mathbb{M}^{N\times N}) := \{\varphi \in C^\infty(\overline{\Omega}; \mathbb{M}^{N\times N}) : \varphi|_{\Gamma_1} = 0\},$$

Proof. The continuity of the normal trace operator restricted to Γ_1 [6],

$$\tau_\nu^{\Gamma_1} : H(\mathrm{Div}; \Omega) \to [H^{-1/2}_{00}(\Gamma_1)]^N, \quad \tau_\nu^{\Gamma_1}(\sigma) := \tau_\nu(\sigma)\big|_{\Gamma_1},$$

ensures that $H_{0,\Gamma_1}(\mathrm{Div}; \Omega) = \ker \tau_\nu^{\Gamma_1}$ is a closed subspace of $H(\mathrm{Div}; \Omega)$ and consequently, the inclusion

$$\overline{C^\infty_{0,\Gamma_1}(\overline{\Omega}; \mathbb{M}^{N\times N})}^{H(\mathrm{Div};\Omega)} \subset \ker \tau_\nu^{\Gamma_1}$$

is valid. Following the strategy of the proof of [23, I, Theorem 2.6] we show that any linear form on $(\ker \tau_\nu^{\Gamma_1})^*$ that vanishes on $C^\infty_{0,\Gamma_1}(\overline{\Omega}; \mathbb{M}^{N\times N})$ is identical to zero. In fact, let $\sigma^* \in (\ker \tau_\nu^{\Gamma_1})^*$ with

$$\langle \sigma^*, \sigma\rangle = 0, \quad \forall\, \sigma \in C^\infty_{0,\Gamma_1}(\overline{\Omega}; \mathbb{M}^{N\times N}). \tag{5.22}$$

By the Riesz Representation Theorem, there exists $\sigma_0 \in \ker \tau_\nu^{\Gamma_1}$ such that

$$\langle \sigma^*, \sigma\rangle = (\sigma_0, \sigma)_Q + (q_0, \mathrm{Div}\,\sigma)_{L^2(\Omega)^N}, \quad \forall\, \sigma \in \ker \tau_\nu^{\Gamma_1}, \tag{5.23}$$

where $q_0 := \mathrm{Div}\,\sigma_0$. Testing (5.23) with functions $\sigma \in C^\infty_c(\Omega; \mathbb{M}^{N\times N})$, one deduces that $\varepsilon(q_0) = \sigma_0$. Thus, it holds that

$$q_0 \in H^1(\Omega)^N. \tag{5.24}$$

We further prove that $q_0 = 0$ on Γ_0. Using (5.22), Green's formula implies that

$$\begin{aligned} \langle \sigma^*, \sigma\rangle &= (\varepsilon(q_0), \sigma)_Q + (q_0, \mathrm{Div}\,\sigma)_{L^2(\Omega)^N} \\ &= \langle \sigma\nu, q_0\rangle_{(H^{-1/2}(\partial\Omega)^N, H^{1/2}(\partial\Omega)^N)} \\ &= \int_{\Gamma_0} (\sigma\nu) q_0 \, d\mathcal{H}^{N-1} = 0, \end{aligned} \tag{5.25}$$

for all $\sigma \in C^\infty_{0,\Gamma_1}(\overline{\Omega}; \mathbb{M}^{N\times N})$. By the density property (5.20) and the fact that

$$\tau_\nu^{\Gamma_0} : H^1_{0,\Gamma_1}(\Omega, \mathbb{M}^{N\times N}) \to [H^{1/2}_{00}(\Gamma_0)]^N, \quad \tau_\nu^{\Gamma_0}(\sigma) := \tau_\nu(\sigma)|_{\Gamma_0},$$

defines a surjective and continuous linear operator (cf. [43, Chapter 5]), (5.25) implies that

$$\int_{\Gamma_0} z \cdot q_0 \, d\mathcal{H}^{N-1} = 0 \quad \forall\, z \in H^{1/2}_{00}(\Gamma_0)^N.$$

By the density of $H^{1/2}_{00}(\Gamma_0)$ in $L^2(\Gamma_0)$, we have that $q_0 = 0$ on Γ_0. It follows from (5.24) that $q_0 \in H^1_{0,\Gamma_0}(\Omega)^N$ and, by definition, also $q_0\big|_{\Gamma_1} \in H^{1/2}_{00}(\Gamma_1)^N$. Let $\sigma \in \ker \tau_\nu^{\Gamma_1}$. Using $q_0 \in H^1_{0,\Gamma_0}(\Omega)^N$, we infer that

$$\langle \sigma^*, \sigma\rangle = (\varepsilon(q_0), \sigma) + (q_0, \mathrm{Div}\,\sigma)_{L^2(\Omega)^N} = \langle \sigma\nu, q_0\rangle_{(H^{-1/2}_{00}(\Gamma_1)^N, H^{1/2}_{00}(\Gamma_1)^N)} = 0,$$

which shows that σ^* is the zero functional on $\ker \tau_\nu^{\Gamma_1}$. $\square$

With the help of the density property provided by Lemma 5.6, the main consistency result for $\gamma \to \infty$ can be derived.

Theorem 5.7. *Let $\mu > 0$ be fixed and let Assumption 5.5 be fulfilled. For a sequence of positive parameters $(\gamma) \subset \mathbb{R}^+$ with $\gamma \to \infty$, the solutions $\sigma_{\mu,\gamma} \in H^1(\Omega; \mathbb{M}^{2\times 2})$ to ($\mathrm{D}_{\mu,\gamma}$) fulfill*

$$\sigma_{\mu,\gamma} \to \sigma_\mu \ \text{ in } H(\mathrm{Div};\Omega), \quad \text{as } \gamma \to \infty,$$

where σ_μ is the solution of ($\tilde{D}_\mu$).

Proof. By convexity, the solution σ_μ of problem ($\tilde{D}_\mu$) is characterized by the variational inequality

$$a_\mu(\sigma_\mu, \tilde\sigma - \sigma_\mu) + j_\mu(\tilde\sigma) - j_\mu(\sigma_\mu) \geq \langle l_\mu, \sigma - \tilde\sigma\rangle, \quad \forall\, \tilde\sigma \in H(\mathrm{Div};\Omega),$$

where

$$\begin{aligned} a_\mu(\sigma,\tilde\sigma) &:= (\mathbb{C}^{-1}\sigma, \tilde\sigma) + \mu(\mathrm{Div}\,\sigma, \mathrm{Div}\,\tilde\sigma)_{L^2(\Omega)^2},\\ j_\mu(\tilde\sigma) &:= i^\mu_{S_{\mathrm{ad}}}(\tilde\sigma) + i_{H_{0,\Gamma_1}(\mathrm{Div};\Omega)}(\tilde\sigma),\\ l_\mu(\sigma) &:= -\langle \hat p^{n-1}, \sigma\rangle - \mu(\mathrm{Div}\,\sigma_\mu, f^n). \end{aligned}$$

On the other hand, the solution $\sigma_{\mu,\gamma}$ of ($\mathrm{D}_{\mu,\gamma}$) is characterized by the variational inequality

$$a_\mu(\sigma_{\mu,\gamma}, \tilde\sigma - \sigma_{\mu,\gamma}) + j_{\mu,\gamma}(\tilde\sigma) - j_{\mu,\gamma}(\sigma_{\mu,\gamma}) \geq \langle l_\mu, \tilde\sigma - \sigma_{\gamma,\mu}\rangle, \quad \forall\, \tilde\sigma \in H(\mathrm{Div};\Omega),$$

where

$$j_{\mu,\gamma}(\tilde\sigma) := i^\mu_{S_{\mathrm{ad}}}(\tilde\sigma) + i_{H_{0,\Gamma_1}(\mathrm{Div};\Omega)}(\tilde\sigma) + \tfrac{1}{\gamma}\|\tilde\sigma\|_{H^1(\Omega;\mathbb{M}^{2\times2})}.$$

Here, it is understood that $j_{\mu,\gamma}(\tilde\sigma) = +\infty$ for $\tilde\sigma \notin H^1(\Omega;\mathbb{M}^{2\times 2})$. Using [14, Proposition 5.7], it is easy to see that $(j_{\mu,\gamma})$ Mosco-converges in $H(\mathrm{Div};\Omega)$ to

$$i^\mu_{S_{\mathrm{ad}}} + i_{\overline{H_{0,\Gamma_1}(\mathrm{Div};\Omega)\cap H^1(\Omega;\mathbb{M}^{2\times2})}^{H(\mathrm{Div};\Omega)}},$$

as $\gamma \to \infty$; see [41, Definition 1.1]. From Lemma 5.6, it follows that

$$\overline{H_{0,\Gamma_1}(\mathrm{Div};\Omega) \cap H^1(\Omega;\mathbb{M}^{2\times2})}^{H(\mathrm{Div};\Omega)} = H_{0,\Gamma_1}(\mathrm{Div};\Omega),$$

which entails that $(j_{\mu,\gamma})$ Mosco-converges to j_μ in $H(\mathrm{Div};\Omega)$ for $\gamma \to \infty$. Together with the ellipticity of a on $H(\mathrm{Div};\Omega)$, standard arguments from the perturbation of variational inequalities (see, for instance, [24, I, Theorem 6.2]) prove that $(\sigma_{\mu,\gamma})$ converges strongly to σ_μ in $H(\mathrm{Div};\Omega)$, which concludes the proof. ☐

5.2.3. *The von Mises yield criterion.* In this section we onsider the special case, where the set of admissible stresses is determined by the von Mises yield criterion, i.e.,

$$\mathbb{K}_0 := \{q \in \mathbb{M}^{2\times2}_0 : |q|_F \leq \sigma_y\}, \quad \sigma_y > 0 \text{ fixed},$$

which is one of the most frequently used yield criteria in practice. In this case, the projection onto the feasible set $S_{\mathrm{ad}} = \{\sigma \in Q : \mathrm{dev}\,\sigma \in \mathbb{K}_0 \text{ a.e. in } \Omega\}$ in the space Q can be computed pointwise. In fact, one obtains

$$\pi_{S_{\mathrm{ad}}}(\sigma) = \sigma - [|\,\mathrm{dev}\,\sigma|_{\mathsf{F}} - \sigma_y]^+ \frac{\mathrm{dev}\,\sigma}{|\,\mathrm{dev}\,\sigma|_F}. \tag{5.26}$$

Under these premises, the problem ($\mathrm{D}_{\mu,\gamma}$) takes the form

$$\begin{cases} \inf & \frac12(\mathbb{C}^{-1}\sigma,\sigma) + \langle \hat p^{n-1},\sigma\rangle + \frac{\mu}{2}\|\,\mathrm{Div}\,\sigma + f^n\|^2_{L^2(\Omega)^2} \\ & \qquad +\frac{\mu}{2}\|[|\,\mathrm{dev}(\sigma)|_F - \sigma_y]^+\|^2_{L^2(\Omega)} + \frac{1}{2\gamma} b(\sigma,\sigma). \\ \text{over} & \sigma \in H^1_{0,\nu}(\Omega;\mathbb{M}^{2\times2}), \end{cases} \tag{5.27}$$

where

$$H^1_{0,\nu}(\Omega;\mathbb{M}^{2\times2}) := \{\sigma \in H(\Omega;\mathbb{M}^{2\times2}) : \sigma\nu = 0 \text{ on } \Gamma_1\}.$$

It turns out that the optimality conditions associated with (5.27) takes the form of a Newton differentiable operator equation. In fact, since $J^*_{\mu,\gamma}$ is convex and Fréchet differentiable, the necessary and sufficient optimality condition for the solution $\sigma_{\mu,\gamma}$ to (5.27), is characterized by the nonsmooth operator equation

$$\Psi_{\mu,\gamma}(\sigma_{\mu,\gamma}) = 0, \tag{5.28}$$

where $\Psi_{\mu,\gamma} : H^1_{0,\nu}(\Omega; \mathbb{M}^{2\times 2}) \to H^1_{0,\nu}(\Omega; \mathbb{M}^{2\times 2})^*$ is defined by

$$\Psi_{\mu,\gamma}(\sigma) := \mathbb{C}^{-1}\sigma + l_\mu + \mu \operatorname{Div}^* \operatorname{Div} \sigma + \mu \operatorname{dev}^* \mathfrak{m}(\operatorname{dev} \sigma) + \tfrac{1}{\gamma} B\sigma. \tag{5.29}$$

Here,

$$\mathfrak{m}(\sigma) := [(|\sigma|_F - \sigma_y)]^+ \mathfrak{q}(\sigma), \text{where } \mathfrak{q}(\sigma) = \begin{cases} \sigma/|\sigma|_F, & \text{if } \sigma \neq 0, \\ 0, & \text{else;} \end{cases}$$

denotes the nonlinear operator associated with the Fréchet derivative of the Moreau-Yosida regularization. We proceed by showing that this equation can be solved efficiently by a generalized Newton scheme, which requires the operator $\Psi_{\mu,\gamma}$ to be Newton differentiable in the sense of Definition 5.4. In this regard, the only issue is the generalized differentiability of the function $\mathfrak{m}$, as all other terms in (5.29) are Fréchet differentiable. The following result is available; cf. [29].

Lemma 5.8. *Let $\beta \in L^\infty(\Omega)$ with $\beta(x) \geq c > 0$ a.e. in Ω. Then the mapping*

$$\mathfrak{m} : u \mapsto [|u|_2 - \beta]^+ \mathfrak{q}(u)$$

is Newton differentiable as a mapping from $L^p(\Omega)^d \to L^s(\Omega)^d$ for $3 \leq 3s \leq p < +\infty$ and

$$G_{\mathfrak{m}}(u) := i_{\mathcal{A}(u)} \cdot \mathfrak{M}(u)$$

defines a Newton derivative of $\mathfrak{m}$, where

$$\begin{aligned} \rho(u) &:= [|u|_2 - \beta]^+ \tfrac{1}{|u|_2}, \\ \mathfrak{M}(u)(\,.\,) &:= \rho(u)(\,.\,) + (1 - \rho(u)) \tfrac{u u^\top(\,.\,)}{|u|_2^2}, \\ \mathcal{A}(u) &:= \{x \in \Omega : |u|_2(x) > \beta(x)\}. \end{aligned}$$

Corollary 5.9. *A Newton derivative*

$$G_{\Psi_{\mu,\gamma}}(\sigma) \in \mathcal{L}(H^1_{0,\nu}(\Omega; \mathbb{M}^{2\times 2}), H^1_{0,\nu}(\Omega; \mathbb{M}^{2\times 2})^*)$$

of $\Psi_{\mu,\gamma}$ at σ is given by

$$\langle G_{\Psi_{\mu,\gamma}}(\sigma)\tilde{\sigma}, \,.\,\rangle := (\mathbb{C}^{-1}\tilde{\sigma}, \,.\,) + \mu \operatorname{Div}^* \operatorname{Div} \tilde{\sigma} + \mu \operatorname{dev}^* G_{\mathfrak{m}}(\operatorname{dev} \sigma)[\operatorname{dev} \tilde{\sigma}] + \tfrac{1}{\gamma} B\tilde{\sigma},$$

for all $\tilde{\sigma} \in H^1_{0,\nu}(\Omega; \mathbb{M}^{2\times 2})$. Here, $G_{\mathfrak{m}}$ denotes the Newton derivative of $\mathfrak{m}$ according to Lemma 5.8. Moreover $G_{\Psi_{\mu,\gamma}}(\sigma)$ is uniformly invertible, i.e., independent of σ.

Proof. According to Lemma 5.8 and the Sobolev embedding theorem, the mapping

$$\sigma \to \mathfrak{m}(\operatorname{dev} \sigma)$$

is Newton differentiable as a mapping from $H^1_{0,\nu}(\Omega; \mathbb{M}^{2\times 2})$ to $Q_0 := \{q \in Q : \operatorname{tr}(q) = 0 \text{ a.e. in } \Omega\}$. It follows immediately that the mapping $\Psi_{\mu,\gamma}$ is Newton differentiable with Newton derivative $G_{\Psi_{\mu,\gamma}}$. It is further straightforward (see, for instance, [30, Lemma 5.5]) to show that $G_{\Psi_{\mu,\gamma}}(\sigma)$ is uniformly elliptic, i.e., there exists $c = c(\mu, \gamma) > 0$ (independent of σ) such that

$$\langle G_{\Psi_{\mu,\gamma}}(\sigma)\tilde{\sigma}, \tilde{\sigma}\rangle \geq c\|\tilde{\sigma}\|^2_{H^1_{0,\nu}(\Omega;\mathbb{M}^{2\times 2})},$$

which entails that $\|G^{-1}_{\Psi_{\mu,\gamma}}(\sigma)\|$ is uniformly bounded. $\square$

Remark 5.10. If the safe-load condition (Assumption 2.4) holds, one can choose $\xi = \hat{\sigma}^n$ as the shift element in (5.17). Then, according to Lemma 5.8, the Newton differentiability of $\Psi_{\mu,\gamma}$ is still valid since $\operatorname{dev} \hat{\sigma}^n \in L^\infty(\Omega; \mathbb{M}_0^{N\times N})$.

As a result of [32, Theorem 1.1], it can be inferred that the corresponding Newton iteration (5.18) (with $F = \Psi_{\mu,\gamma}$) is well-defined provided the starting point $\sigma^{(0)}$ is sufficiently close to $\sigma_{\mu,\gamma}$. Moreover, the iterates $(\sigma^{(j)})$ converge locally at a superlinear rate, which is mesh-independent upon discretization. To enforce global convergence, one may equip the search directions

$$\delta^{(j)} := -G_{\Psi_{\mu,\gamma}}(\sigma^{(j)})\Psi_{\mu,\gamma}(\sigma^{(j)})$$

with a step size determined by the Armijo line search procedure. The resulting method is summarized in Algorithm (SSN(μ,γ)).

Algorithm SSN(μ,γ): Globalized SSN algorithm

---

input: $\sigma^{(0)} \in H^1_{0,\nu}(\Omega; \mathbb{M}^{2\times 2})$

set $j := 0$;
while some stopping rule is not satisfied **do**
  compute the solution $\delta^{(j)} \in$ of $G_{\Psi_{\mu,\gamma}}(\sigma^{(j)})\delta^{(j)} = -\Psi_{\mu,\gamma}(\sigma^{(j)})$;
  determine $\alpha^{(j)} > 0$ by an Armijo line search based on $\alpha \mapsto J^*_{\mu,\gamma}(\sigma^{(j)} + \alpha\delta^{(j)})$;
  set $\sigma^{(j+1)} := \sigma^{(j)} + \alpha^{(j)}\delta^{(j)}$ and $j := j + 1$;

With the help of the gradient-relatedness of the search direction and the strong convexity of the objective function $J^*_{\mu,\gamma}$, it is standard to infer that the sequence $(\sigma^{(j)})$ generated by SSN(μ,γ) equipped with an Armijo line search is globally convergent in the sense that $(\sigma^{(j)})$ converges strongly to the solution of (5.27) in the norm of $H^1_{0,\nu}(\Omega; \mathbb{M}^{2\times 2})$;

$$\sigma^{(j)} \to \sigma_{\mu,\gamma} \quad \text{in } H^1(\Omega; \mathbb{M}^{2\times 2}).$$

We refer, e.g., to [10] for details.

5.3. **Outlook.** In order to approximate the solution σ of (D) one needs to pass to the limit in (D$_{\mu,\gamma}$) with $\mu, \gamma \to +\infty$. This can be achieved by a path-following strategy in the spirit of [30, 28]. While the semismooth reformulation of problem (D$_\mu$) based on a Tikhonov regularization resembles the approach for hardening plasticity from [30], the construction of a consistent regularization in perfect plasticity, which does not rely on a vanishing hardening (or visco-plastic) approach, necessitates a more involved inspection. In fact, for the study of the limiting case as $\mu, \gamma \to +\infty$, one may resign to the stability analysis from [30]. However, this approach is complicated by the presence of the additional equality constraints

$$-\operatorname{Div}\sigma = f^n, \quad \sigma\nu = g^n \text{ on } \Gamma_1, \tag{5.30}$$

defining the feasible set of the limit problem (D), since the extension of the required density property (cf. [30, Assumption 4.1]) to incorporate the equality constraints (5.30) is not possible. As a consequence, a special coupling of the penalization-regularization parameters μ and γ is necessary to be consistent with the limit problem (D). For this issue, we refer, e.g., to [40, Proposition 2.4.6.]. Moreover, the effect on the primal problem (P$_\mu$) of the Tikhonov regularization of (D$_\mu$) remains to be investigated.

Another aspect concerns the convergence of finite element discretizations for (D$_{\mu,\gamma}$) based on a sequence of meshes with decreasing mesh width h. The stability of the corresponding discretized-regularized scheme requires a convergence result as $h \to 0$ and $\mu, \gamma \to \infty$, which necessitates a suitable coupling between the three parameters. Even if the convergence of the discretized-regularized stresses can be shown, it is still necessary to pass to the limit (as the mesh width tends to 0) in the resulting discretized version of the primal-dual optimality conditions (5.6)-(5.9) (or their regularized version) in order to study the convergence of the approximative displacements and plastic strains.

In the case of elasticity, which formally corresponds to problem (D) with $\mathbb{K}_0 = \mathbb{M}_0^{N\times N}$, the convergence of the discrete stress-displacement pair in mixed finite element methods hinges on the validity of the LBB condition for saddle point problems [11]. Moreover, a conformal discretization of the stresses requires the incorporation of the symmetry and divergence constraints from the definition of the space $H(\mathrm{Div};\Omega)$. As a result, rather sophisticated finite elements, for example those of Arnold and Winther [4] are proposed in the literature. The resulting finite-dimensional approximation usually involves a very large number of (local) degrees of freedom. It is expected that these aspects also need to be taken into account in order to construct a stable discrete primal-dual path-following strategy in order to solve the plasticity problem (P) via (D$_\mu$) or (D$_{\mu,\gamma}$).

References

[1] R.A. Adams and J.J.F. Fournier. *Sobolev Spaces*. Pure and Applied Mathematics. Elsevier Science, 2003.

[2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Clarendon Press Oxford, 2000.

[3] G. Anzellotti and M. Giaquinta. Existence of the displacements field for an elasto-plastic body subject to Hencky's law and Von Mises yield condition. *manuscripta mathematica*, 32(1):101–136, 1980.

[4] D.N. Arnold and R. Winther. Mixed finite elements for elasticity. *Numerische Mathematik*, 92(3): 401–419, 2002.

[5] J.P. Aubin and I. Ekeland. *Applied Nonlinear Analysis*. Dover Books on Mathematics Series. Dover Publications, 1984.

[6] C. Baiocchi and A. Capelo. *Variational and Quasivariational Inequalities*. John Wiley and Sons, 1984.

[7] S. Bartels and T. Roubíček. Numerical approaches to thermally coupled perfect plasticity. *Numerical Methods for Partial Differential Equations*, 29(6):1837–1863, 2013.

[8] S. Bartels, A. Mielke, and T. Roubíček. Quasi-static small-strain plasticity in the limit of vanishing hardening and its numerical approximation. *SIAM J. Numer. Anal.*, 50(2):951–976, 2012.

[9] J.-M. E. Bernard. Density results in Sobolev spaces whose elements vanish on a part of the boundary. *Chinese Annals of Mathematics, Series B*, 32(6):823–846, 2011.

[10] J.-F. Bonnans, J.C. Gilbert, C. Lemaréchal, and C.A. Sagastizábal. *Numerical Optimization*. Springer, 2nd edition, 2006.

[11] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Spinger Series in Computational Mathematics, 1991.

[12] C. Carstensen. Convergence of adaptive finite element methods in computational mechanics. *Appl. Numer. Math.*, 59(9):2119–2130, 2009.

[13] X. Chen, Z. Nashed, and L. Qi. Smoothing methods and semismooth methods for nondifferentiable operator equations. *SIAM Journal on Numerical Analysis*, 38(4):pp. 1200–1216, 2001.

[14] G. Dal Maso. *An Introduction to Γ-convergence*. Birkhäuser, Boston, 1993.

[15] G. Dal Maso, A. DeSimone, and M.G. Mora. Quasistatic evolution problems for linearly elastic-perfectly plastic materials. *Archive for Rational Mechanics and Analysis*, 180(2):237–291, 2006.

[16] A. Demyanov. Regularity of stresses in Prandtl-Reuss perfect plasticity. *Calculus of Variations and Partial Differential Equations*, 34(1):23–72, 2009.

[17] P. Doktor and A. Ženíšek. The density of infinitely differentiable functions in Sobolev spaces with mixed boundary conditions. *Appl. Math., Praha*, 51(5):517–547, 2006.

[18] G. Duvaut and J.-L. Lions. *Les inéquations en mécanique et en phyisque*. Dunod, Paris, 1972.

[19] F. Ebobisse and B.D. Reddy. Some mathematical problems in perfect plasticity. *Computer Methods in Applied Mechanics and Engineering*, 193(48):5071 – 5094, 2004.

[20] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1987.

[21] G.A. Francfort and A. Giacomini. Small-strain heterogeneous elastoplasticity revisited. *Communications on Pure and Applied Mathematics*, 65(9):1185–1241, 2012.

[22] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, 1998.

[23] V. Girault and P.-A. Raviart. *Finite Element Methods for Navier-Stokes equations*. Springer-Verlag, 1986.

[24] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Springer, 1984.

[25] W. Han and B.D. Reddy. *Plasticity: Mathematical Theory and Numerical Analysis*. Springer, New York, 2nd edition, 2013.

[26] M. Hintermüller. Mesh independence and fast local convergence of a primal-dual active-set method for mixed control-state constrained elliptic control problems. *ANZIAM J.*, 49(1):1–38, 2007.

[27] M. Hintermüller and K. Kunisch. Total bounded variation regularization as bilaterally constrained optimization problem. *SIAM J. Applied Mathematics*, 64:1311–1333, 2004.

[28] M. Hintermüller and K. Kunisch. Feasible and non-interior path-following in constrained minimization with low multiplier regularity. *SIAM J. Control and Optimization*, 45(4):1198–1221, 2006.

[29] M. Hintermüller and C.N. Rautenberg. A sequential minimization technique for elliptic quasi-variational inequalities with gradient constraints. *SIAM Journal on Optimization*, 22(4):1224–1257, 2012.

[30] M. Hintermüller and S. Rösel. A duality-based path-following semismooth Newton method for elasto-plastic contact problems. *Journal of Computational and Applied Mathematics*, 292:150 – 173, 2016.

[31] M. Hintermüller and M. Ulbrich. A mesh-independence result for semismooth Newton methods. *Math. Program., Ser. B*, 171:151–184, 2004.

[32] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semi-smooth Newton method. *SIAM J. Optimization*, 13(3):865–888, 2003.

[33] K. Ito and K. Kunisch. Augmented Lagrangian methods for nonsmooth, convex optimization in Hilbert spaces. *Nonlinear Anal., Theory Methods Appl., Ser. A, Theory Methods*, 41(5-6):591–616, 2000.

[34] C. Johnson. Existence theorems for plasticity problems. *J. Math. pures et appl.*, 55:431–444, 1976.

[35] R. Kohn and R. Temam. Dual spaces of stresses and strains, with applications to Hencky plasticity. *Applied Mathematics and Optimization*, 10(1):1–35, 1983. ISSN 0095-4616.

[36] S. Li and W.K. Liu. Numerical simulations of strain localization in inelastic solids using mesh-free methods. *International Journal for Numerical Methods in Engineering*, 48(9):1285–1309, 2000.

[37] A. Mainik and A. Mielke. Existence results for energetic models for rate-independent systems. *Calculus of Variations and Partial Differential Equations*, 22(1):73–99, 2005.

[38] Gianni Dal Maso, Alexey Demyanov, and Antonio DeSimone. Quasistatic evolution problems for pressure-sensitive plastic materials. *Milan Journal of Mathematics*, 75(1):117–134, 2007.

[39] H. Matthies. Finite element approximations in thermo-plasticity. *Numerical Functional Analysis and Optimization*, 1(2):145–160, 1979.

[40] A. Mielke and T. Roubíček. *Rate-independent Systems*. Springer, 2015.

[41] U. Mosco. Convergence of convex sets and of solutions of variational inequalities. *Advances in Mathematics*, 3(4):510 – 585, 1969.

[42] A. Needleman and V. Tvergaard. Analyses of plastic flow localization in metals. *Applied Mechanics Reviews*, 45(3S):S3–S18, 1992.

[43] J. T. Oden and N. Kikuchi. *Contact Problems in Elasticity*. Society for Industrial and Applied Mathematics, 1988.

[44] R. Rannacher and F.-T. Suttmeier. A posteriori error control in finite element methods via duality techniques: Application to perfect plasticity. *Computational mechanics*, 21(2):123–133, 1998.

[45] S. Repin. Errors of finite element method for perfectly elasto-plastic problems. *Mathematical Models and Methods in Applied Sciences*, 06(05):587–604, 1996.

[46] T. Roubíček and J. Valdman. Perfect Plasticity with Damage and Healing at Small Strains, Its Modeling, Analysis, and Computer Implementation. *SIAM Journal on Applied Mathematics*, 76(1): 314–340, 2016.

[47] M. Sauter. *Numerical Analysis of Algorithms for Infinitesimal Associated and Non-Associated Elasto-Plasticity*. PhD thesis, Karlsruher Institut für Technologie, 2010.

[48] M. Sauter and C. Wieners. On the superlinear convergence in computational elasto-plasticity. *Computer Methods in Applied Mechanics and Engineering*, 200(49-52):3646 – 3658, 2011.

[49] A. Schwarz, J. Schröder, and G. Starke. Least-squares mixed finite elements for small strain elasto-viscoplasticity. *International journal for numerical methods in engineering*, 77(10):1351–1370, 2009.

[50] J.C. Simo and T.J.R. Hughes. *Computational Inelasticity*. Springer-Verlag Berlin, 1998.

[51] F. Solombrino. Quasistatic Evolution in Perfect Plasticity for General Heterogeneous Materials. *Archive for Rational Mechanics and Analysis*, 212(1):283–330, 2014.

[52] G. Strang and R. Temam. Functions of Bounded Deformation. *Archive for Rational Mechanics and Analysis*, 75(1):7–21, 1980.

[53] P.-M. Suquet. Sur les équations de la plasticité: existence et régularité des solutions. *J. Mec., Paris*, 20:3–39, 1981.

[54] R. Temam. *Mathematical Problems in Plasticity*. Gauthier-Villars, 1985.

[55] C. Wieners. Multigrid methods for Prandtl-Reuss plasticity. *Numerical Linear Algebra with Applications*, 6(6):457–478, 1999.

[56] C. Wieners. Nonlinear solution methods for infinitesimal perfect plasticity. *Zeitschrift für Angewandte Mathematik und Mechanik*, 87(8-9):643–660, 2007.