

A duality based semismooth Newton framework for solving variational inequalities of the second kind

JUAN CARLOS DE LOS REYES

*Departamento de Matemática, EPN Quito, Ecuador and
Institute of Mathematics, Humboldt-Universität zu Berlin, Germany
E-mail: juan.delosreyes@epn.edu.ec*

MICHAEL HINTERMÜLLER

*Institute of Mathematics, Humboldt-Universität zu Berlin, Germany
E-mail: hint@math.hu-berlin.de*

[Received 18 January 2010 and in revised form 7 September 2011]

In an appropriate function space setting, semismooth Newton methods are proposed for iteratively computing the solution of a rather general class of variational inequalities (VIs) of the second kind. The Newton scheme is based on the Fenchel dual of the original VI problem which is regularized if necessary. In the latter case, consistency of the regularization with respect to the original problem is studied. The application of the general framework to specific model problems including Bingham flows, simplified friction, or total variation regularization in mathematical imaging is described in detail. Finally, numerical experiments are presented in order to verify the theoretical results.

Keywords: Variational inequalities of the second kind; semismooth Newton methods; Moreau–Yosida regularization.

1. Introduction

Variational inequalities (VIs) are a widely used paradigm when modeling physical phenomena involving free boundaries such as, for instance, contact of rigid bodies with or without friction, visco-plastic fluid flow, or flow through porous media. The associated variational formulation allows one to gain insight into the problem structure and to investigate both theoretical and numerical aspects of a pertinent solution algorithm in a function space setting. A rather general form of an elliptic VI may be written as follows:

$$a(y, v - y) + \phi(v) - \phi(y) \geq \langle \ell, v - y \rangle_{Y^*, Y} \quad \text{for all } v \in Y, \quad (1.1)$$

where Y is a Hilbert space, $a : Y \times Y \rightarrow \mathbb{R}$ a bi-continuous bilinear form, $\phi : Y \rightarrow \bar{\mathbb{R}}$ a proper convex function, and $\ell \in Y^*$. Here and below, Y^* denotes the topological dual space of Y with $\langle \cdot, \cdot \rangle_{Y^*, Y}$ the associated duality pairing.

In the special case where the convex function ϕ is the indicator function of a convex set, the problem (1.1) characterizes a so-called VI of the first kind. A frequently studied instance is the obstacle problem, where ϕ is the indicator function of $\{y \in Y : y \leq \psi \text{ a.e. in } \Omega\}$ with ψ representing some suitably given obstacle and “a.e.” standing for “almost everywhere”. The design and analysis of numerical methods for solving such VIs have been intensively carried out in the recent past. Here, we only refer to [21] as well as the monograph [12] and the many references therein.

In the case where such an indicator function representation of ϕ is not available, the variational inequality (1.1) is of the second kind and constitutes a problem of a somewhat different nature. In contrast to VIs of the first kind, the presence of the potentially (or rather typically) non-differentiable function ϕ turns out to be problematic. In fact, while VIs of the first kind may be handled via projection techniques, a more subtle approach appropriately dealing with generalized derivatives is necessary for VIs of the second kind. From a practical point of view, applications of VIs of the second kind involve many important phenomena such as friction, visco-plastic fluid flow or problems in elasto/visco-plasticity. Some analytical results for such problems can be found, e.g., in [9, 12, 28, 32, 37].

Besides frequently used, but rather slow fixed point type techniques, several numerical solution schemes exist for VIs of the second kind. An excellent source for relaxation, penalization and augmented Lagrangian techniques is [12]. Inexact Uzawa techniques were studied in [2] and extensions of multigrid methods in [26, 27, 29, 30], partly with an emphasis on VIs of the first kind. While these methods are typically first order or converge at a linear rate, in [23] and [6, 7] generalized second order methods were proposed for the numerical solution of VIs of the second kind modelling applications in image restoration or Bingham fluid flows, respectively. Due to a certain lack of regularity of the function space problems, only a finite-dimensional convergence analysis of the algorithm was carried out, yielding (at least) a locally superlinear rate of convergence.

In this paper, we aim at bridging the gap between the finite-dimensional convergence analysis in [6, 23] and the function space context of the original continuous problem. Moreover, the layout of a unifying framework for solving variational inequalities of the second kind is an issue. The key ingredient for these steps is a dual regularization strategy which allows us, on one hand, to get uniqueness of the dual variable (in some cases even without an additional dual regularization) and, on the other hand, to work in an adequate function space setting where the superlinear convergence of the infinite-dimensional version of a generalized Newton method can be obtained. An immediate consequence of this result is the mesh independent convergence behavior in the sense of [24].

The plan of this paper is the following. In Section 2 the rather general model problem under consideration is stated in an abstract form and its corresponding Fenchel dual is derived. A family of regularized problems is introduced and studied in Section 3. The emphasis will be on function space regularity properties of the regularized duals and on the convergence of the solutions of the regularized problems towards the solution of the original problem in the case of increasing regularization parameter. In Section 4 a semismooth Newton algorithm for solving the regularized problems is described and, based on general results, its locally superlinear convergence rate is verified. Some model applications are highlighted in Section 5, where their realization within our general framework is outlined as well. Finally, in Section 6 some numerical results are presented in order to experimentally verify our theoretical findings.

2. Problem statement and duality

In this paper we confine ourselves to studying the following model problem:

$$\text{minimize } F(y) + \alpha \int_S |By|_2 \, ds \quad \text{over } y \in Y, \quad (P)$$

where $F : Y \rightarrow \mathbb{R}$ is uniformly convex and lower semicontinuous with the gradient mapping $F'(\cdot) \in Y^*$ satisfying

$$\|F'(y_1) - F'(y_2)\|_{Y^*} \leq \frac{1}{\kappa} \|y_1 - y_2\|_Y \quad \forall y_1, y_2 \in Y, \quad (2.1)$$

for some constant $\kappa > 0$. Here Y denotes a Hilbert space. Moreover, $\alpha > 0$, $\Omega \subset \mathbb{R}^d$, with $d \in \mathbb{N}$, is a bounded domain, and $S \subset \bar{\Omega}$. In fact, we are particularly interested in either $S \subseteq \Omega$ or $S \subseteq \partial\Omega$. Further, we assume that $B \in \mathcal{L}(Y, L^2(S, \mathbb{R}^{l \times m}))$ for $1 \leq l, m \leq d$, with $l, m \in \mathbb{N}$, i.e., B is a continuous linear operator from Y to $L^2(S, \mathbb{R}^{l \times m})$. For $K, L \in \mathbb{R}^{l \times m}$ we use $K \cdot L = \text{trace}(K^\top L)$ and $|K|_2^2 := \text{trace}(K^\top K)$. Note that if $l = m$, then $|\cdot|_2$ becomes the Frobenius norm. For the ease of notation, in what follows we write $\mathbb{L}^2(S) := L^2(S, \mathbb{R}^{l \times m})$.

For $p \in Y^*$, the polar or conjugate function of F is defined by

$$F^*(p) = \sup_{y \in Y} \{ \langle p, y \rangle_{Y^*, Y} - F(y) \}.$$

Under our assumptions on F , the conjugate F^* is uniformly convex with modulus κ ; see [25]. Next we focus on the Fenchel dual of the second term in the objective of (P). For this purpose, we define $G : \mathbb{L}^2(S) \rightarrow \mathbb{R}$ by

$$G(z) := \alpha \int_S |z|_2 \, ds. \tag{2.2}$$

Its convex conjugate is given by

$$G^*(q) = I_{\mathcal{F}}(q), \tag{2.3}$$

with the indicator function

$$I_{\mathcal{M}}(q) = \begin{cases} 0 & \text{if } q \in \mathcal{M}, \\ +\infty & \text{else,} \end{cases}$$

for some set \mathcal{M} , and $\mathcal{F} = \{q \in \mathbb{L}^2(S) : |q|_2 \leq \alpha \text{ a.e. in } S\}$. Choosing $V = \mathbb{L}^2(S)$ and noting that $B \in \mathcal{L}(Y, V)$ by definition, the Fenchel duality theorem [10] yields

$$\inf_{y \in Y} \{ F(y) + G(By) \} = \sup_{p \in V^*} \{ -F^*(B^*p) - G^*(-p) \}. \tag{2.4}$$

Observe that the problem on the left in (2.4) coincides with (P), which we call the *primal problem*, and the one on the right is the associated Fenchel dual of (P). Upon identifying V^* with V , the latter is equivalent to

$$\begin{cases} \text{minimize} & F^*(B^*p) & \text{over } p \in \mathbb{L}^2(S) \\ \text{subject to} & |p|_2 \leq \alpha \text{ a.e. in } S. \end{cases} \tag{DP}$$

The solutions \bar{y} of (P) and \bar{p} of (DP), respectively, are linked by the optimality system

$$\begin{aligned} B^*\bar{p} &= F'(\bar{y}), \\ -\bar{p} &\in \partial G(B\bar{y}), \end{aligned} \tag{2.5}$$

where ∂ denotes the subdifferential from convex analysis [10]. Moreover, (2.4) indicates that there is no duality gap between (P) and (DP). As a consequence, one may solve (DP) and use (2.5) to compute the solution of (P).

Our goal is to develop efficient (that is, at least locally superlinearly convergent) solution algorithms for (P). For this purpose, we take (DP) as the starting point, introduce a suitable regularization scheme, whenever necessary, which allows us to define a generalized Newton method in a function space, solve (DP) with the help of this Newton solver and utilize (2.5) to compute \bar{y} . To

proceed in this direction, we next state the first order necessary and, due to convexity, also sufficient optimality conditions for (DP):

$$|\bar{p}|_2 \leq \alpha \quad \text{a.e.}, \tag{2.6}$$

$$BF^{*'}(B^* \bar{p}) + \bar{\lambda} = 0, \tag{2.7}$$

$$(\bar{\lambda}, v - \bar{p})_{\mathbb{L}^2(S)} \leq 0 \quad \text{for all } v \in \mathbb{L}^2(S) \text{ with } |v|_2 \leq \alpha \text{ a.e.}, \tag{2.8}$$

or, equivalently,

$$BF^{*'}(B^* \bar{p}) + \bar{\lambda}_s \bar{q} = 0, \text{ with } \bar{q}(x) \in \begin{cases} \left\{ \frac{\bar{p}}{|\bar{p}|_2}(x) \right\} & \text{if } |\bar{p}(x)|_2 > 0, \\ \mathbb{B}(0, 1) & \text{else,} \end{cases} \tag{2.9}$$

$$\bar{\lambda}_s = \max(0, \bar{\lambda}_s + c(|\bar{p}|_2 - \alpha)), \tag{2.10}$$

for some $c > 0$. Here, $\mathbb{B}(0, 1) = \{q \in \mathbb{R}^{l \times m} : |q|_2 \leq 1\}$, and $(\cdot, \cdot)_{\mathbb{L}^2(S)}$ denotes the standard inner product in $\mathbb{L}^2(S)$. Note in addition that (2.10) is equivalent to the complementarity system

$$\bar{\lambda}_s \geq 0, \quad |\bar{p}|_2 \leq \alpha, \quad \bar{\lambda}_s(|\bar{p}|_2 - \alpha) = 0 \text{ a.e. in } S. \tag{2.11}$$

3. Regularized problems

Unless B^* and F^* satisfy certain regularity properties (which we discuss in detail below), the system (2.9)–(2.10) is not suitable for generalized or Newton differentiation and for an application of a generalized (or, in our case, semismooth) version of Newton’s method for its iterative solution; see [18] for both. For this reason, whenever the regularity properties of B^* and F^* are not sufficient, we consider the following regularized form of (DP):

$$\text{minimize } J_\gamma^*(p) := F^*(B^* p) + \frac{\gamma}{2} \|(|p|_2 - \alpha)^+ \|_{\mathbb{L}^2(S)}^2 + \frac{1}{2\gamma} \|Cp\|_{\mathbb{L}^2(S)^n}^2 \quad \text{over } p \in \mathbb{H}, \tag{DP_\gamma}$$

where $\gamma > 0$, $(\cdot)^+ = \max(0, \cdot)$, and $C \in \mathcal{L}(\mathbb{H}, \mathbb{L}^2(S)^n)$, $n \in \mathbb{N}$, for some suitable Hilbert space $\mathbb{H} \subset \mathbb{L}^2(S)$ with dense embedding. We further assume that

$$\|Cp\|_{\mathbb{L}^2(S)^n} \geq \beta_0 \|p\|_{\mathbb{H}} \quad \text{for some } \beta_0 > 0. \tag{3.1}$$

Both regularization terms are necessary whenever $\bar{p} \notin L^{2+\epsilon_p}(S, \mathbb{R}^{l \times m})$ for some $\epsilon_p > 0$; otherwise one may dismiss the term $\frac{1}{2\gamma} \|Cp\|_{\mathbb{L}^2(S)^n}^2$, which lifts p into \mathbb{H} by (3.1), but keep the first regularization term. The benefit from the latter is related to Newton differentiability of the resulting first order optimality condition; for details we refer to Section 4.

The existence of a unique solution $p_\gamma \in \mathbb{H}$ of (DP_γ) can be verified by standard arguments. Moreover, the solution satisfies the following first order necessary and sufficient optimality condition:

$$BF^{*'}(B^* p_\gamma) + \lambda_\gamma + \frac{1}{\gamma} C^* Cp_\gamma = 0, \tag{3.2}$$

$$\lambda_\gamma = \gamma (|p_\gamma|_2 - \alpha)^+ \frac{p_\gamma}{\max(\epsilon, |p_\gamma|_2)}. \tag{3.3}$$

Here and below we use $0 < \epsilon < \alpha$, and the max-operation in the denominator above is understood in the pointwise almost everywhere sense.

From now on we shall assume that $F^* : Y^* \rightarrow \mathbb{R}$ is bounded from below and has the form

$$F^*(v) = \frac{1}{2}a^*(v - z, v - z) + b, \tag{3.4}$$

where $a^*(\cdot, \cdot) : Y^* \times Y^* \rightarrow \mathbb{R}$ is a continuous and coercive bilinear form, $z \in Y^*$, and $b \in \mathbb{R}$. Note that this assumption covers situations where F is associated with a quadratic energy. Particular instances include Bingham flows, frictional contact in elasticity, or applications in elasto/viscoplasticity and imaging science, as we shall see in Section 5.

We start our investigation of (DP_γ) by studying its consistency properties, i.e., the convergence behavior of $\{(p_\gamma, \lambda_\gamma)\}$ as $\gamma \rightarrow \infty$. From now on, we denote weak convergence by \rightharpoonup and strong convergence by \rightarrow .

THEOREM 3.1 Let $\gamma \geq \bar{\gamma} > 0$. Then the sequence $\{(p_\gamma, \lambda_\gamma)\}$ of solutions to (3.2)–(3.3) converges to the solution $(\bar{p}, \bar{\lambda})$ of (2.7)–(2.8) weakly in $\mathbb{L}^2(S) \times \mathbb{L}^2(S)$ as $\gamma \rightarrow \infty$. Moreover,

$$B^* p_\gamma \rightarrow B^* \bar{p} \quad \text{in } Y^* \text{ as } \gamma \rightarrow \infty. \tag{3.5}$$

Proof. Multiplying equation (3.2) by p_γ and using the adjoint operator we get

$$\frac{1}{\gamma} \|Cp_\gamma\|_{\mathbb{L}^2(S)^n}^2 + \langle F^{*'}(B^* p_\gamma), B^* p_\gamma \rangle_{Y, Y^*} + (\lambda_\gamma, p_\gamma)_{\mathbb{L}^2(S)} = 0. \tag{3.6}$$

Using the Cauchy–Schwarz inequality and the form of F^* given in (3.4), (3.6) implies

$$\frac{1}{\gamma} \|Cp_\gamma\|_{\mathbb{L}^2(S)^n}^2 + \kappa \|B^* p_\gamma\|_{Y^*}^2 + (\lambda_\gamma, p_\gamma)_{\mathbb{L}^2(S)} \leq \beta \|z\|_{Y^*} \|B^* p_\gamma\|_{Y^*}, \tag{3.7}$$

with some positive constant β which may take different values on different occasions. From the structure of λ_γ and (3.7) it follows that

$$(\lambda_\gamma, p_\gamma)_{\mathbb{L}^2(S)} = \int_{\mathcal{A}_\gamma} \gamma (|p_\gamma|_2 - \alpha) |p_\gamma|_2 \, dx \leq \beta \|z\|_{Y^*} \|B^*\|_{\mathcal{L}(\mathbb{L}^2(S), Y^*)} \|p_\gamma\|_{\mathbb{L}^2(S)},$$

where $\mathcal{A}_\gamma = \{x \in S : |p_\gamma(x)|_2 \geq \alpha\}$. This implies

$$\int_{\mathcal{A}_\gamma} |p_\gamma|_2^2 \, dx \leq \frac{\beta \|z\|_{Y^*}}{\gamma} \|B^*\|_{\mathcal{L}(\mathbb{L}^2(S), Y^*)} \|p_\gamma\|_{\mathbb{L}^2(S)} + \alpha \int_{\mathcal{A}_\gamma} |p_\gamma|_2 \, dx.$$

On the other hand, $|p_\gamma(x)|_2 \leq \alpha$ on $\mathcal{I}_\gamma = S \setminus \mathcal{A}_\gamma$, and consequently

$$\int_{\mathcal{I}_\gamma} |p_\gamma|_2^2 \, dx \leq \alpha \int_{\mathcal{I}_\gamma} |p_\gamma|_2 \, dx.$$

From the last two inequalities it follows that

$$\|p_\gamma\|_{\mathbb{L}^2(S)}^2 \leq \frac{\beta \|z\|_{Y^*}}{\gamma} \|B^*\|_{\mathcal{L}(\mathbb{L}^2(S), Y^*)} \|p_\gamma\|_{\mathbb{L}^2(S)} + \alpha \int_S |p_\gamma|_2 \, dx, \tag{3.8}$$

which, by Hölder's inequality and $\gamma \geq \bar{\gamma}$, implies that the sequence $\{p_\gamma\}$ is uniformly bounded in $\mathbb{L}^2(S)$. Moreover, from (3.2) and (3.4) it follows that $\{\lambda_\gamma\}$ is uniformly bounded in $\mathbb{L}^2(S)$. Therefore, there exists a point (p^*, λ^*) in $\mathbb{L}^2(S) \times \mathbb{L}^2(S)$ and a subsequence, also denoted by $\{(p_\gamma, \lambda_\gamma)\}$, such that

$$(p_\gamma, \lambda_\gamma) \rightharpoonup (p^*, \lambda^*) \quad \text{in } \mathbb{L}^2(S) \times \mathbb{L}^2(S). \quad (3.9)$$

Next we verify that the weak limit point (p^*, λ^*) satisfies the optimality system for the original problem (P). Since $\|\lambda_\gamma\|_{\mathbb{L}^2(S)} = \|\gamma(|p_\gamma|_2 - \alpha)^+\|_{\mathbb{L}^2(S)} \leq \beta$ for all $\gamma \geq \bar{\gamma}$, it follows that $\|(|p_\gamma|_2 - \alpha)^+\|_{\mathbb{L}^2(S)} \rightarrow 0$ as $\gamma \rightarrow \infty$. As the functional $p_\gamma \mapsto \|(|p_\gamma|_2 - \alpha)^+\|_{\mathbb{L}^2(S)}$ is convex and continuous, it is weakly lower semicontinuous, and consequently

$$0 \leq \int_S \max(0, |p^*|_2 - \alpha)^2 dx \leq \liminf_{\gamma \rightarrow \infty} \int_S \max(0, |p_\gamma|_2 - \alpha)^2 dx = 0. \quad (3.10)$$

Therefore, $|p^*|_2 \leq \alpha$ a.e. in S , and condition (2.6) is verified. From the optimality of p_γ we obtain

$$F^*(B^* p_\gamma) + \frac{1}{2\gamma} \|C p_\gamma\|_{\mathbb{L}^2(S)^n}^2 \leq F^*(B^* \tilde{p}) + \frac{1}{2\gamma} \|C \tilde{p}\|_{\mathbb{L}^2(S)^n}^2 \quad (3.11)$$

for all $\tilde{p} \in \{v \in \mathbb{H} : |v|_2 \leq \alpha \text{ a.e. in } S\}$. Taking the upper limit in the last inequality yields

$$\limsup_{\gamma \rightarrow \infty} F^*(B^* p_\gamma) \leq F^*(B^* \tilde{p}) + \limsup_{\gamma \rightarrow \infty} \frac{1}{2\gamma} \|C \tilde{p}\|_{\mathbb{L}^2(S)^n}^2 \leq F^*(B^* \tilde{p}) \quad (3.12)$$

for all $\tilde{p} \in \{v \in \mathbb{H} : |v|_2 \leq \alpha \text{ a.e. in } S\}$. From the density of $\{v \in \mathbb{H} : |v|_2 \leq \alpha \text{ a.e. in } S\}$ in $\{v \in \mathbb{L}^2(S) : |v|_2 \leq \alpha \text{ a.e. in } S\}$ and the properties of F^* , p_γ and p^* it follows that

$$\limsup_{\gamma \rightarrow \infty} F^*(B^* p_\gamma) \leq F^*(B^* p^*), \quad (3.13)$$

which, by the weak lower semicontinuity of $F^*(B^* \cdot)$, implies that

$$\limsup_{\gamma \rightarrow \infty} F^*(B^* p_\gamma) \leq F^*(B^* p^*) \leq \liminf_{\gamma \rightarrow \infty} F^*(B^* p_\gamma),$$

and consequently together with (3.9) we obtain (3.5).

Since $\{\gamma^{-1} \|C p_\gamma\|_{\mathbb{L}^2(S)^n}^2\}$ is uniformly bounded, it follows that $\gamma^{-1} C p_\gamma \rightarrow 0$ in $\mathbb{L}^2(S)^n$ as $\gamma \rightarrow \infty$. Then passing to the limit in (3.2) shows that (p^*, λ^*) is a solution of

$$(\lambda^*, v)_{\mathbb{L}^2(S)} = -(BF^{*'}(B^* p^*), v)_{\mathbb{L}^2(S)} \quad \text{for all } v \in \mathbb{H}, \quad (3.14)$$

which, by the density of \mathbb{H} in $\mathbb{L}^2(S)$, yields (2.7).

From (3.14) and (3.5) we get

$$(\lambda^*, v - p^*)_{\mathbb{L}^2(S)} = -(BF^{*'}(B^* p^*), v - p^*)_{\mathbb{L}^2(S)} = - \lim_{\gamma \rightarrow \infty} (BF^{*'}(B^* p_\gamma), v - p_\gamma)_{\mathbb{L}^2(S)},$$

which, by (3.2), implies that

$$(\lambda^*, v - p^*)_{\mathbb{L}^2(S)} \leq \lim_{\gamma \rightarrow \infty} \left(\frac{1}{\gamma} (C p_\gamma, C v) + (\lambda_\gamma, v - p_\gamma)_{\mathbb{L}^2(S)} \right).$$

Since

$$(\lambda_\gamma, v - p_\gamma)_{\mathbb{L}^2(S)} \leq \gamma \int_S (|p_\gamma|_2 - \alpha)^+ (|v|_2 - |p_\gamma|_2) \, dx \leq \gamma \int_S (|p_\gamma|_2 - \alpha)^+ (\alpha - |p_\gamma|_2) \, dx \leq 0,$$

for all $v \in \mathbb{L}^2(S)$ with $|v|_2 \leq \alpha$ a.e. in S , and $\lim_{\gamma \rightarrow \infty} \gamma^{-1} C p_\gamma \rightarrow 0$, it follows that

$$(\lambda^*, v - p^*)_{\mathbb{L}^2(S)} \leq 0 \quad \text{for all } v \in \mathbb{L}^2(S) \text{ with } |v|_2 \leq \alpha \text{ a.e. in } S.$$

Therefore, from the strict convexity of the problem we conclude that $(p^*, \lambda^*) = (\bar{p}, \bar{\lambda})$, which completes the proof. \square

REMARK 3.2 From $B^* p_\gamma \rightarrow B^* p$ in Y^* , as $\gamma \rightarrow \infty$, also the strong convergence (in Y) of the regularized primal solutions y_γ to the primal solution \bar{y} is obtained.

4. Semismooth Newton method

4.1 General case

The first order optimality condition (3.2)–(3.3) is the starting point for our generalized Newton algorithm for solving (DP_γ) , and hence (P) . The latter is achieved via a continuation technique with respect to (increasing) γ . In this section, however, we keep γ fixed. For the subsequent development we invoke the following assumptions:

$$\begin{cases} \text{(a)} & \hat{\mathbb{H}} := \mathbb{H} \hookrightarrow L^r(S, \mathbb{R}^{l \times m}) \text{ for some } r > 2 \text{ if } l + m = 2, \text{ or} \\ \text{(b)} & C^* C \in \mathcal{L}(\hat{\mathbb{H}}, \mathbb{L}^2(S)) \text{ is a homeomorphism, where} \\ & \hat{\mathbb{H}} \hookrightarrow \mathbb{H} \text{ with } \hat{\mathbb{H}} \hookrightarrow L^\infty(S, \mathbb{R}^{l \times m}) \text{ if } l + m > 2, \end{cases} \quad (4.1)$$

where \hookrightarrow denotes dense and continuous injection.

In view of (3.3), the applicability of a generalized version of Newton’s method for solving (3.2)–(3.3) depends on whether, for fixed $\epsilon \in (0, \alpha)$,

$$p \mapsto (|p|_2 - \alpha)^+ q(p), \quad \text{with } q(p)(x) := \frac{p(x)}{\max(\epsilon, |p(x)|_2)} \text{ for almost all } x \in \Omega,$$

admits a generalized derivative. For this purpose, we recall the notion of a Newton (or slant) derivative (see [18]) of a mapping $f : X \rightarrow Y$ from a Banach space X to a Banach space Y . It resembles the concept of semismoothness, which is well-known in finite dimensions [31, 35]. In its definition we use $o(\cdot)$ with the property $o(\vartheta)/\vartheta \rightarrow 0$ as $\vartheta \rightarrow 0$.

DEFINITION 4.1 The mapping $f : U \subseteq X \rightarrow Y$ is called *Newton differentiable* in the open set U if there exists a family of mappings $g : U \rightarrow \mathcal{L}(X, Y)$ such that

$$\|f(x + \delta_x) - f(x) - g(x + \delta_x)\delta_x\|_Y = o(\|\delta_x\|_X) \quad (4.2)$$

for all $x \in U$ as $\|\delta_x\|_X \rightarrow 0$.

Further, it is known that $\max(0, \cdot) : L^t(S) \rightarrow L^s(S)$ is Newton differentiable if and only if $1 \leq s < t \leq +\infty$; see [18, Proposition 4.1].

Utilizing the same technique as in the proof of [22, Lemma 3.1] we find that for $l + m > 2$,

$$m : p \mapsto (|p|_2 - \alpha)^+ q(p) \quad (4.3)$$

is Newton differentiable as a mapping from $L^\infty(S, \mathbb{R}^{l \times m})$ to $L^s(S, \mathbb{R}^{l \times m})$ for any $s \in [1, \infty)$. A particular Newton derivative is given by

$$\mathfrak{M}(p)(x) = \begin{cases} \left[\frac{pp^\top}{|p|_2^2} + (|p|_2 - \alpha)^+ \frac{1}{|p|_2} \left(\text{id} - \frac{pp^\top}{|p|_2^2} \right) \right](x) & \text{for } x \in \mathcal{A}(p), \\ 0 & \text{for } x \in \mathcal{I}(p), \end{cases} \quad (4.4)$$

where $\mathcal{A}(p) := \{x \in S : |p(x)|_2 > \alpha\}$ and $\mathcal{I}(p) := S \setminus \mathcal{A}(p)$. For $l + m = 2$, (4.3) may be written as

$$m(p) = \max(0, p - \alpha) + \min(0, p + \alpha),$$

which is Newton differentiable as a mapping from $L^t(S)$ to $L^s(S)$ for any $t > s \geq 1$; see, e.g., [18]. In this case, $\mathfrak{M}(p)$ of (4.4) becomes

$$\mathfrak{M}(p) = \chi_{\mathcal{A}^+(p)} + \chi_{\mathcal{A}^-(p)},$$

where $\chi_{\mathcal{A}^+(p)}$ and $\chi_{\mathcal{A}^-(p)}$ are the characteristic functions of $\mathcal{A}^+(p) := \{x \in S : p(x) > \alpha\}$ and $\mathcal{A}^-(p) := \{x \in S : p(x) < -\alpha\}$, respectively.

The above generalized differentiability results are applicable upon performing a reduction step in (3.2)–(3.3). In fact, let

$$D_\gamma := BA^*B^* + \frac{1}{\gamma}C^*C \in \mathcal{L}(\mathbb{H}^*, \mathbb{H}), \quad (4.5)$$

where $A^* \in \mathcal{L}(Y^*, Y)$ satisfies $a^*(w_1, w_2) = \langle A^*w_1, w_2 \rangle_{Y, Y^*}$ for $w_1, w_2 \in Y^*$. Then (3.2) is equivalent to

$$D_\gamma p_\gamma + \lambda_\gamma = BA^*z. \quad (4.6)$$

Setting $z_\gamma := D_\gamma^{-1}BA^*z$, we thus have $p_\gamma = z_\gamma - D_\gamma^{-1}\lambda_\gamma \in L^t(S, \mathbb{R}^{l \times m})$ with $t = r > 2$ in the case of $l + m = 2$ and $t = \infty$ for $l + m > 2$ due to assumption (4.1). The proof of this result uses the facts that $BA^*z, \lambda_\gamma \in \mathbb{L}^2(S)$ and $B \in \mathcal{L}(Y, \mathbb{L}^2(S))$. Using this representation of p_γ in (3.3) yields

$$\lambda_\gamma + \gamma(|D_\gamma^{-1}\lambda_\gamma - z_\gamma|_2 - \alpha)^+ q(D_\gamma^{-1}\lambda_\gamma - z_\gamma) = 0 \quad \text{in } \mathbb{L}^2(S). \quad (4.7)$$

Note that D_γ^{-1} maps $\mathbb{L}^2(S)$ -perturbations of λ_γ into $L^t(S, \mathbb{R}^{l \times m})$, and as a consequence the above Newton differentiability result applies. The generalized linearization of (4.7) at some approximation $\lambda^{(k)} \in \mathbb{L}^2(S)$ of λ_γ is given by

$$\begin{aligned} & \left(\text{id} + \gamma \chi_{\mathcal{A}(p^{(k)})} \left[\frac{(D_\gamma^{-1}\lambda^{(k)} - z_\gamma)(D_\gamma^{-1}\lambda^{(k)} - z_\gamma)^\top}{|D_\gamma^{-1}\lambda^{(k)} - z_\gamma|_2^2} D_\gamma^{-1} + (|D_\gamma^{-1}\lambda^{(k)} - z_\gamma| - \alpha)^+ \right. \right. \\ & \quad \left. \left. \cdot \left(\frac{1}{|D_\gamma^{-1}\lambda^{(k)} - z_\gamma|_2} - \frac{(D_\gamma^{-1}\lambda^{(k)} - z_\gamma)(D_\gamma^{-1}\lambda^{(k)} - z_\gamma)^\top}{|D_\gamma^{-1}\lambda^{(k)} - z_\gamma|_2^3} \right) D_\gamma^{-1} \right] \right) \delta_\lambda \\ & = -\lambda^{(k)} - \gamma(|D_\gamma^{-1}\lambda^{(k)} - z_\gamma|_2 - \alpha)^+ q(D_\gamma^{-1}\lambda^{(k)} - z_\gamma), \end{aligned} \quad (4.8)$$

where $p^{(k)} = D_\gamma^{-1}\lambda^{(k)} - z_\gamma$ (note the difference in sign between the relations p_γ/λ_γ and $p^{(k)}/\lambda^{(k)}$, respectively). Set $\delta_p := D_\gamma^{-1}\delta_\lambda$. Then (4.8) becomes

$$(D_\gamma + \gamma \mathfrak{M}(p^{(k)}))\delta_p = -D_\gamma p^{(k)} - D_\gamma z_\gamma - \gamma m(p^{(k)}). \quad (4.9)$$

Now we have all the ingredients to define the semismooth Newton method for solving (3.2)–(3.3).

SEMISMOOTH NEWTON ALGORITHM.

- (i) Choose $\lambda^{(0)} \in \mathbb{L}^2(S)$, set $k := 0$, and compute $p^{(0)} = D_\gamma^{-1}\lambda^{(0)} - z_\gamma$.
- (ii) Unless some stopping rule is satisfied, compute the solution $\delta_p^{(k)} \in L^t(S, \mathbb{R}^{l \times m})$ of (4.9), and set $\delta_\lambda^{(k)} := D_\gamma \delta_p^{(k)} \in \mathbb{L}^2(S)$.
- (iii) Set $\lambda^{(k+1)} := \lambda^{(k)} + \delta_\lambda^{(k)}$, $p^{(k+1)} := p^{(k)} + \delta_p^{(k)}$, and $k := k + 1$. Return to (ii).

We start our investigation of this algorithm by establishing the uniform invertibility of the operator associated with the left hand side of system (4.9) in step (ii).

LEMMA 4.2 The operator

$$W(p^{(k)}) := D_\gamma + \gamma \mathfrak{M}(p^{(k)}) \in \mathcal{L}(\mathbb{H}, \mathbb{H}^*)$$

is uniformly invertible (i.e., invertible independently of $p^{(k)}$). Moreover, if $W(p^{(k)})\delta_p^{(k)} = r^{(k)}$ for some $r^{(k)} \in \mathbb{L}^2(S)$, then $\|\delta_\lambda^{(k)}\|_{\mathbb{L}^2(S)} \leq \beta(\gamma)\|r^{(k)}\|_{\mathbb{L}^2(S)}$ where $\beta(\gamma) > 0$ is a constant depending on γ .

Proof. Let $v \in \mathbb{H}$. We start by noting that

$$\langle \mathfrak{M}(p^{(k)})v, v \rangle_{\mathbb{H}^*, \mathbb{H}} \geq 0,$$

since

$$\left\langle \chi_{\mathcal{A}_0(p^{(k)})} \left(\text{id} - \frac{p^{(k)}(p^{(k)})^\top}{|p^{(k)}|_2^2} \right) v, v \right\rangle_{\mathbb{H}^*, \mathbb{H}} = \int_S \chi_{\mathcal{A}_0(p^{(k)})} \left(\text{id} - \frac{p^{(k)}(p^{(k)})^\top}{|p^{(k)}|_2^2} \right) v^2 \, dx \geq 0,$$

with $\chi_{\mathcal{A}_0(p^{(k)})}$ the characteristic function of $\mathcal{A}_0(p^{(k)}) := \{x \in S : |p^{(k)}(x)|_2 > 0\}$. Thus,

$$\begin{aligned} \langle W(p^{(k)})v, v \rangle_{\mathbb{H}^*, \mathbb{H}} &= \langle A^* B^* v, B^* v \rangle_{Y, Y^*} + \gamma \langle \mathfrak{M}(p^{(k)})v, v \rangle_{\mathbb{H}^*, \mathbb{H}} + \frac{1}{\gamma} \|Cv\|_{\mathbb{L}^2(S)^n}^2 \\ &\geq \frac{1}{\gamma} \|Cv\|_{\mathbb{L}^2(S)^n}^2 \geq \frac{1}{\gamma} \beta_0^2 \|v\|_{\mathbb{H}}^2, \end{aligned}$$

by assumption (3.1), which proves the first assertion.

For the proof of the second statement we first assume that $\hat{\mathbb{H}} \subset \mathbb{H}$ in the strict sense. Let $r^{(k)} \in \mathbb{L}^2(S)$. Then, by the first part and the properties of \mathbb{H} , we get

$$\|\delta_p^{(k)}\|_{\mathbb{L}^2(S)} \leq \beta \|\delta_p^{(k)}\|_{\mathbb{H}} \leq \beta(\gamma) \|r^{(k)}\|_{\mathbb{L}^2(S)} \tag{4.10}$$

with some constant $\beta(\gamma) > 0$ depending on γ . This constant may take different values at different occurrences. We set $\beta_C := \|(C^*C)^{-1}\|_{\mathcal{L}(\mathbb{L}^2(S), \hat{\mathbb{H}})} < \infty$ by assumption (4.1)(b). We conclude from (4.5) that

$$\begin{aligned} \|\delta_p^{(k)}\|_{\hat{\mathbb{H}}} &\leq \gamma \beta_C (\|BA^* B^*\|_{\mathcal{L}(Y^*, \mathbb{L}^2(S))} + \gamma \|\mathfrak{M}(p^{(k)})\|_{\mathcal{L}(\mathbb{L}^2(S), \mathbb{L}^2(S))}) \|\delta_p^{(k)}\|_{\mathbb{L}^2(S)} \\ &\quad + \gamma \beta_C \|r^{(k)}\|_{\mathbb{L}^2(S)}. \end{aligned} \tag{4.11}$$

Now, (4.10) yields

$$\|\delta_p^{(k)}\|_{\hat{\mathbb{H}}} \leq \beta(\gamma) \|r^{(k)}\|_{\mathbb{L}^2(S)}. \quad (4.12)$$

Finally, the above estimate and assumption (4.1)(b) imply

$$\|\delta_\lambda^{(k)}\|_{\mathbb{L}^2(S)} \leq \|D_\gamma\|_{\mathcal{L}(\hat{\mathbb{H}}, \mathbb{L}^2(S))} \|\delta_p^{(k)}\|_{\hat{\mathbb{H}}} \leq \beta(\gamma) \|r^{(k)}\|_{\mathbb{L}^2(S)}.$$

The assertion in the case $\hat{\mathbb{H}} = \mathbb{H}$ follows readily by utilizing $\mathbb{H} \subset L^r(S, \mathbb{R}^{l \times m})$ for some $r > 2$. \square

As a consequence of the above lemma we have $p^{(k+1)} \in \hat{\mathbb{H}}$ provided that $p^{(k)} \in \hat{\mathbb{H}}$. Hence, \mathfrak{m} admits a Newton derivative at $\lambda^{(k+1)}$ and step (ii) of our algorithm is well-defined in this respect.

The locally q -superlinear convergence of the semismooth Newton algorithm is established next.

THEOREM 4.3 If $\lambda^{(0)} \in \mathbb{L}^2(S)$ is sufficiently close to λ_γ , then the semismooth Newton iteration is well-defined and satisfies

$$\|\lambda^{(k+1)} - \lambda_\gamma\|_{\mathbb{L}^2(S)} = \mathcal{O}(\|\lambda^{(k)} - \lambda_\gamma\|_{\mathbb{L}^2(S)}) \quad \text{as } k \rightarrow \infty.$$

Proof. We define $F_\gamma(\lambda) = \lambda + \gamma \mathfrak{m}(D_\gamma^{-1} \lambda - z_\gamma)$, which corresponds to the left hand side of the reduced equation (4.7) with $\lambda = \lambda_\gamma$. From the generalized Newton step (i.e., step (ii) of our algorithm) we infer

$$\begin{aligned} \|\lambda^{(k+1)} - \lambda_\gamma\|_{\mathbb{L}^2(S)} &= \|D_\gamma(D_\gamma + \gamma \mathfrak{M}(D_\gamma^{-1} \lambda^{(k)} - z_\gamma))^{-1}(F_\gamma(\lambda_\gamma) - F_\gamma(\lambda^{(k)})) \\ &\quad - (D_\gamma + \gamma \mathfrak{M}(D_\gamma^{-1} \lambda^{(k)} - z_\gamma)) D_\gamma^{-1}(\lambda_\gamma - \lambda^{(k)})\|_{\mathbb{L}^2(S)} \\ &\leq \beta(\gamma) \|F_\gamma(\lambda_\gamma) - F_\gamma(\lambda^{(k)}) - (D_\gamma + \gamma \mathfrak{M}(D_\gamma^{-1} \lambda^{(k)} - z_\gamma)) D_\gamma^{-1}(\lambda_\gamma - \lambda^{(k)})\|_{\mathbb{L}^2(S)} \\ &= \beta(\gamma) \|\mathfrak{m}(D_\gamma^{-1} \lambda_\gamma - z_\gamma) - \mathfrak{m}(D_\gamma^{-1} \lambda^{(k)} - z_\gamma) - \mathfrak{M}(D_\gamma^{-1} \lambda^{(k)} - z_\gamma) D_\gamma^{-1}(\lambda_\gamma - \lambda^{(k)})\|_{\mathbb{L}^2(S)} \\ &= \beta(\gamma) \|\mathfrak{m}(p_\gamma) - \mathfrak{m}(p^{(k)}) - \mathfrak{M}(p^{(k)})(p_\gamma - p^{(k)})\|_{\mathbb{L}^2(S)} \end{aligned} \quad (4.13)$$

by using $p_\gamma = D_\gamma^{-1} \lambda_\gamma - z_\gamma$ in contrast to the sign convention used earlier. Setting $p := p_\gamma$ and $h = p^{(k)} - p_\gamma$, we further study the right hand side above. In fact, we get

$$\|\mathfrak{m}(p_\gamma) - \mathfrak{m}(p^{(k)}) - \mathfrak{M}(p^{(k)})(p_\gamma - p^{(k)})\|_{\mathbb{L}^2(S)} = \|\mathfrak{m}(p+h) - \mathfrak{m}(p) - \mathfrak{M}(p+h)h\|_{\mathbb{L}^2(S)}.$$

Since $p, h \in L^t(S, \mathbb{R}^{l \times m})$ with $t = r$ for $l+m = 2$ and $t = \infty$ if $l+m > 2$ due to assumption (4.1) and by the Newton differentiability of $\mathfrak{m} : L^t(S, \mathbb{R}^{l \times m}) \rightarrow \mathbb{L}^2(S)$ (with \mathfrak{M} as an associated Newton derivative), we have

$$\|\mathfrak{m}(p+h) - \mathfrak{m}(p) - \mathfrak{M}(p+h)h\|_{\mathbb{L}^2(S)} = \mathcal{O}(\|h\|_{L^t(S, \mathbb{R}^{l \times m})}) \quad \text{as } \|h\|_{L^t(S, \mathbb{R}^{l \times m})} \rightarrow 0. \quad (4.14)$$

Note that from assumption (4.1) we infer

$$\|h\|_{L^t(S, \mathbb{R}^{l \times m})} = \|D_\gamma^{-1}(\lambda^{(k)} - \lambda_\gamma)\|_{L^t(S, \mathbb{R}^{l \times m})} \leq \beta(\gamma) \|\lambda^{(k)} - \lambda_\gamma\|_{\mathbb{L}^2(S)}. \quad (4.15)$$

Now, due to (4.14), similar arguments to those in [18, Thm. 1.1] yield the well-definedness of the semismooth Newton solver, i.e., if the method is initialized within a sufficiently small $\mathbb{L}^2(S)$ -neighborhood of λ_γ , then all iterates stay within this neighborhood and satisfy

$$\|\mathfrak{m}(p+h) - \mathfrak{m}(p) - \mathfrak{M}(p+h)h\|_{\mathbb{L}^2(S)} \leq \frac{\eta}{\hat{\beta}(\gamma)} \|h\|_{\mathbb{L}^2(S)} \quad (4.16)$$

for some fixed $\eta \in (0, 1)$ and $\hat{\beta}(\gamma)$ depending on γ . Moreover, increasing $\hat{\beta}(\gamma)$ if necessary, due to (4.13) it follows that

$$\|\lambda^{(k+1)} - \lambda_\gamma\|_{\mathbb{L}^2(S)} \leq \eta \|\lambda^{(k)} - \lambda_\gamma\|_{\mathbb{L}^2(S)}.$$

Thus, (4.13)–(4.16) yield

$$\|\lambda^{(k+1)} - \lambda_\gamma\|_{\mathbb{L}^2(S)} = \mathcal{O}(\|\lambda^{(k)} - \lambda_\gamma\|_{\mathbb{L}^2(S)})$$

as $k \rightarrow \infty$, which ends the proof. \square

4.2 Global convergence

In order to guarantee global convergence, i.e., convergence regardless of the choice of the initial point $\lambda^{(0)} \in \mathbb{L}^2(S)$, of our semismooth Newton method we equip the algorithm with an Armijo line search step. For this purpose, note that the right hand side in (4.9) is equal to $-J_\gamma^{*\prime}(p^{(k)})$ and $W(p^{(k)})$ is a specific Newton derivative of $J_\gamma^{*\prime}(p^{(k)})$.

PROPOSITION 4.4 The Newton direction $\delta_p^{(k)} := p^{(k+1)} - p^{(k)}$ is a descent direction for $J_\gamma^*(\cdot)$ at $p^{(k)}$, i.e., there exists a constant $\beta_{J_\gamma^*} > 0$ independent of $p^{(k)}$ and $\delta_p^{(k)}$ such that

$$\langle J_\gamma^{*\prime}(p^{(k)}), \delta_p^{(k)} \rangle_{\mathbb{H}^*, \mathbb{H}} \leq -\beta_{J_\gamma^*} \|J_\gamma^{*\prime}(p^{(k)})\|_{\mathbb{H}^*}^2.$$

Proof. We have

$$-\langle J_\gamma^{*\prime}(p^{(k)}), W(p^{(k)})^{-1} J_\gamma^{*\prime}(p^{(k)}) \rangle_{\mathbb{H}^*, \mathbb{H}} \leq -\frac{\beta_0^2}{\gamma \beta_W^2(\gamma)} \|J_\gamma^{*\prime}(p^{(k)})\|_{\mathbb{H}^*}^2$$

with a constant $\beta_W(\gamma)$ satisfying $\|W(p^{(k)})\|_{\mathcal{L}(\mathbb{H}, \mathbb{H}^*)} \leq \beta_W(\gamma)$ for all $k \in \mathbb{N}$ and depending on γ .

Then the assertion holds with $\beta_{J_\gamma^*} = \frac{\beta_0^2}{\gamma \beta_W^2(\gamma)}$. \square

From Proposition 4.4 we immediately infer that an Armijo line search may be used for the globalization of our generalized Newton method. For this reason, we replace step (iii) of the semismooth Newton algorithm by

(iii') Find the smallest index $l^{(k)} \in \mathbb{N}_0$ such that

$$J_\gamma^*(p^{(k)} + \omega^{l^{(k)}} \delta_p^{(k)}) - J_\gamma^*(p^{(k)}) \leq -\theta \omega^{l^{(k)}} \langle J_\gamma^{*\prime}(p^{(k)}), \delta_p^{(k)} \rangle_{\mathbb{H}^*, \mathbb{H}}.$$

Then set $\tau^k := \omega^{l^{(k)}}$, and $p^{(k+1)} := p^{(k)} + \tau^k \delta_p^{(k)}$, $\lambda^{(k+1)} := \lambda^{(k)} + \tau^k \delta_\lambda^{(k)}$, as well as $k := k + 1$; return to (ii).

Here, $\omega, \theta \in (0, 1)$ denote fixed parameters. The new step (iii') implements a backtracking line search to obtain a step size $\tau^{(k)}$. More sophisticated techniques, e.g., relying on interpolation, may be utilized as well. Standard arguments [34] guarantee the global convergence of the generalized Newton iteration to p_γ resp. λ_γ .

4.3 Special case

In applications, sometimes $BF^{*''}(B^*p)B^*$ satisfies

$$BF^{*''}(B^*p)B^* \in \mathcal{L}(\mathbb{L}^2(S), L^r(S, \mathbb{R}^{l \times m})) \quad \text{for some } r > 2. \quad (4.17)$$

A particular example is the simplified friction problem, as we shall see in the subsequent section. In such a situation, one may still use the penalization/regularization framework of the previous section. However, alternatively, the following direct approach is possible (and preferable). Instead of (DP_γ) consider

$$\begin{aligned} \text{minimize} \quad & F^*(B^*p) + \frac{\mu}{2} \|p\|_{\mathbb{L}^2(S)}^2 \quad \text{over } p \in \mathbb{L}^2(S) \\ \text{subject to} \quad & |p|_2 \leq \alpha \text{ a.e. in } S \end{aligned} \quad (DP_\mu)$$

for some $\mu > 0$. Again, it is standard to deduce the existence of a unique solution $p_\mu \in \mathbb{L}^2(S)$ of (DP_μ) , which satisfies the first order optimality condition

$$BF^{*'}(B^*p_\mu) + \lambda_\mu + \mu p_\mu = 0, \quad (4.18)$$

$$\lambda_\mu = (\lambda_\mu + c(|p_\mu|_2 - \alpha))^+ \quad (4.19)$$

for some $c > 0$. Note that (4.19) is equivalent to

$$\lambda_\mu \geq 0, \quad |p_\mu|_2 - \alpha \leq 0, \quad \lambda_\mu(|p_\mu|_2 - \alpha) = 0 \quad \text{a.e. in } S,$$

which constitutes the complementarity system associated with the pointwise constraint in (DP_μ) .

For $l = m = 1$ we obtain, upon choosing $c = \mu > 0$,

$$BF^{*'}(B^*p_\mu) + \mu p_\mu + (-BF^{*'}(B^*p_\mu) - \mu\alpha)^+ + (-BF^{*'}(B^*p_\mu) + \mu\alpha)^- = 0,$$

where $(\cdot)^- = \min(\cdot, 0)$ in the pointwise a.e. sense. The corresponding semismooth Newton step computation (with solution $\delta_p^{(k)}$) reads

$$\begin{aligned} & (\text{id} - \chi_{\mathcal{A}_+^{(k)}} - \chi_{\mathcal{A}_-^{(k)}})BF^{*''}(B^*p^{(k)})B^* + \mu \text{id} \delta_p \\ & = -BF^{*'}(B^*p^{(k)}) - \mu p^{(k)} - (-BF^{*'}(B^*p^{(k)}) - \mu\alpha)^+ - (-BF^{*'}(B^*p^{(k)}) + \mu\alpha)^-, \end{aligned}$$

where

$$\mathcal{A}_+^{(k)} := \{-BF^{*'}(B^*p^{(k)}) - \mu\alpha > 0\} \quad \text{and} \quad \mathcal{A}_-^{(k)} := \{-BF^{*'}(B^*p^{(k)}) + \mu\alpha < 0\}.$$

Moreover, $\chi_{\mathcal{A}_+^{(k)}}$ and $\chi_{\mathcal{A}_-^{(k)}}$ denote the characteristic functions of $\mathcal{A}_+^{(k)}$ and $\mathcal{A}_-^{(k)}$, respectively. The locally superlinear convergence of the associated Newton iteration follows from the chain rule for a Newton differentiable function post-composed with a C^1 -function in [19, App. B] and the convergence result [18, Thm. 1.1]. From the semismooth Newton step it becomes apparent that $\mu > 0$ guarantees the solvability of the Newton equation for computing $\delta_p^{(k)}$.

5. Applications

One of the main purposes of this paper is to show that our approach is indeed a unifying framework for solving VIs of the second kind. In order to illustrate this, we study several applications by means of the above technique.

5.1 Simplified friction

A simplified friction problem was introduced in [9] and studied from a numerical point of view in [12]. For its formulation we set $Y := \{y \in H^1(\Omega) : \tau_0 y = 0 \text{ a.e. on } \Gamma_0\}$, where Ω is an open bounded domain with a Lipschitz boundary $\Gamma = \partial\Omega$, $\Gamma_f \subset \Gamma$, and $\Gamma_0 = \Gamma \setminus \Gamma_f$, $\ell \in L^2(\Omega)$, and $\tau_0 : H^1(\Omega) \rightarrow H^{1/2}(\Gamma_0) \subset L^2(\Gamma_0)$ is the zero-order trace operator. The problem fits into the framework (P) with

$$F(y) = \frac{1}{2} \int_{\Omega} |\nabla y|_2^2 dx - (\ell, y)_{L^2(\Omega)} =: \frac{1}{2} a(y, y) - (\ell, y)_{L^2(\Omega)}.$$

For $y_1, y_2 \in Y$ one readily checks that

$$\|F'(y_1) - F'(y_2)\|_{Y^*} = \sup_{v \in Y, \|v\|_Y=1} \int_{\Omega} \nabla(y_1 - y_2) \cdot \nabla v dx \leq \|y_1 - y_2\|_Y.$$

Hence, (2.1) holds true with $\kappa = 1$. The convex conjugate of $F : Y \rightarrow \mathbb{R}$ is given by

$$F^* : Y^* \rightarrow \mathbb{R}, \quad F^*(w) = \frac{1}{2} \int_{\Omega} |\nabla u(w)|_2^2 dx =: \frac{1}{2} a(u(w), u(w)),$$

where $u(w)$ is the unique solution of

$$a(u, v) - (\ell, v)_{L^2(\Omega)} - \langle w, v \rangle_{Y^*, Y} = 0 \quad \text{for all } v \in Y.$$

We further have $S = \Gamma_f$, $B = \tau_f$, where $\tau_f : Y \rightarrow \mathbb{L}^2(S)$ denotes the zero-order trace operator, and thus $l = m = 1$. Hence, the Fenchel dual reads

$$\begin{aligned} & \text{minimize} \quad \int_{\Omega} |\nabla u(\tau_f^* p)|_2^2 dx \quad \text{over } p \in \mathbb{L}^2(S) = L^2(\Gamma_f) \\ & \text{subject to} \quad |p|_2 \leq \alpha \text{ a.e. on } \Gamma_f, \end{aligned}$$

where τ_f^* denotes the adjoint of τ_f . Here, we identify $L^2(\Gamma_f)$ with its dual. By the Sobolev embedding theorem we have $H^{1/2}(\Gamma_f) \subset L^r(\Gamma_f)$ compactly for $r \in [1, \infty)$. Hence, by the properties of the trace operator [40], $BF^{*''}(B^* p)B^*$ satisfies (4.17). Here we use the differentiability of $u(\cdot) : Y^* \rightarrow Y$ with $u'(y^*)\varphi$, with $\varphi \in Y^*$, satisfying

$$a(u'(y^*)\varphi, v) = (\ell, v)_{L^2(\Omega)} \quad \text{for all } v \in Y$$

and $\langle u''(y^*)\psi, \varphi \rangle = 0$ for $\varphi, \psi \in Y^*$.

Moreover, $l + m = 2$ so that the first case of assumption (4.1) is relevant and $\hat{\mathbb{H}} = \mathbb{H} = H_0^1(\Omega)$ may be chosen as the latter embeds continuously into $L^{r(d)}(\Omega)$, with $r(d) = \infty$ for $d = 1$, $1 \leq r(d) < \infty$ for $d = 2$ and $1 \leq r(d) \leq 2d/(d-2)$ for dimensions $d \geq 3$. In particular, we have $r(d) > 2$ for all $d \in \mathbb{N}$. Then, $C := \nabla$ satisfies (3.1) by the Poincaré–Friedrichs inequality [40].

As a consequence, for the numerical solution of the simplified friction problem the semismooth Newton approaches of Section 4.1 or Section 4.3 are both applicable.

For a related investigation of a semismooth Newton approach to the simplified friction problem we refer to [39]. We also note that these considerations may be generalized to contact problems in elasticity; see [28] for the latter.

5.2 Bingham fluids (scalar case)

The flow of a visco-plastic fluid in a pipe was studied theoretically in [13, 32] and numerically in [12]. It fits into our framework by choosing $Y = H_0^1(\Omega)$, where $\Omega \subset \mathbb{R}^2$ denotes a bounded domain with a Lipschitz boundary. Further we have $\ell \in Y^* = H^{-1}(\Omega)$ and

$$F(y) = \frac{\nu}{2} \int_{\Omega} |\nabla y|_2^2 \, dx - \langle \ell, y \rangle_{Y^*, Y},$$

where $\nu > 0$ stands for the viscosity coefficient of the fluid. Moreover, $S = \Omega$ and $B = \nabla : Y \rightarrow \mathbb{L}^2(\Omega)$ with $l = 2$ and $m = 1$. One easily checks that F satisfies (2.1) with $\kappa = \nu^{-1}$. Its convex conjugate is given by

$$F^* : H^{-1}(\Omega) \rightarrow \mathbb{R}, \quad F^*(w) = \frac{1}{2\nu} \|w + \ell\|_{H^{-1}(\Omega)}^2,$$

where $\|w\|_{H^{-1}(\Omega)}^2 = \langle w, (-\Delta)^{-1}w \rangle_{H^{-1}(\Omega), H_0^1(\Omega)}$ for $w \in H^{-1}(\Omega)$, with $(-\Delta)^{-1} \in \mathcal{L}(H^{-1}(\Omega), H_0^1(\Omega))$ the solution operator of $-\Delta u = w$ in $H^{-1}(\Omega)$. In this case $a^*(w, w) := \frac{1}{\nu} \|w\|_{H^{-1}(\Omega)}^2$ and $z = \ell, b = 0$. Hence, the Fenchel dual of the Bingham flow problem is given by

$$\begin{aligned} &\text{minimize} && \frac{1}{2\nu} \|\operatorname{div} p + \ell\|_{H^{-1}(\Omega)}^2 && \text{over } p \in \mathbb{L}^2(S) = L^2(\Omega, \mathbb{R}^2) \\ &\text{subject to} && |p|_2 \leq \alpha \text{ a.e. on } \Omega. \end{aligned}$$

In general, $BF^{**}(B^*p)B^* \in \mathcal{L}(\mathbb{L}^2(\Omega), \mathbb{L}^2(\Omega))$, but it no longer satisfies (4.17) in contrast to the simplified friction case. Concerning the regularization of p , under sufficient regularity of Ω we may choose $\mathbb{H} = H_0^1(\Omega, \mathbb{R}^2), \hat{\mathbb{H}} = H^2(\Omega, \mathbb{R}^2) \cap H_0^1(\Omega, \mathbb{R}^2)$, and $C = \nabla$. The requirement (3.1) follows from the Poincaré–Friedrichs inequality and the compact embedding of $\hat{\mathbb{H}}$ into $L^\infty(S, \mathbb{R}^2)$ from the Sobolev embedding theorem; see, e.g., [40]. Further, for a given $g \in \mathbb{L}^2(\Omega)$, finding $\delta_p \in \hat{\mathbb{H}}$ such that $C^*C\delta_p = g$ in $\mathbb{L}^2(\Omega)$ is equivalent to finding the solution of $-\Delta\delta_p = g$ in $\mathbb{L}^2(\Omega)$ with $\delta_p = 0$ on $\partial\Omega$. We also note that for $\ell \in L^2(\Omega)$ and additional regularity of Ω , the solution \bar{y} of the Bingham flow problem satisfies $\bar{y} \in H^2(\Omega) \cap H_0^1(\Omega)$ (cf. [3, Thm. 15]).

Summarizing, for the numerical solution of the Bingham flow problem, our general semismooth Newton approach of Section 4 is applicable.

For a different semismooth Newton approach relying on finite-dimensional arguments we refer to [6].

5.3 Total variation regularization in image processing

In [38] total variation regularization was proposed as an efficient technique for edge-preserving image restoration (ROF-model). In [23] a dual formulation of a slightly modified version of the associated variational problem was derived. The resulting ROF-type model reads as follows:

$$\text{minimize} \quad \frac{1}{2} \|Ky - y_d\|_{L^2(\Omega)}^2 + \alpha \int_{\Omega} |\nabla y| \, dx + \frac{\beta}{2} \|\nabla y\|_{L^2(\Omega)}^2 \quad \text{over } y \in H_0^1(\Omega), \quad (5.1)$$

where Ω is a Lipschitz (image) domain, $K \in \mathcal{L}(L^2(\Omega), L^2(\Omega))$ a blurring operator, $y_d \in L^2(\Omega)$ the given image data, and $0 < \beta \ll \alpha$ regularization parameters. This problem fits into our

framework (P) by choosing $Y = H_0^1(\Omega)$,

$$F(y) = \frac{1}{2} \|Ky - z\|_{L^2(\Omega)}^2 + \frac{\beta}{2} \|\nabla y\|_{L^2(\Omega)}^2,$$

which satisfies (2.1) with $\kappa = (\beta + \|K\|_{\mathcal{L}(L^2(\Omega), L^2(\Omega))})^{-1}$, $S := \Omega$ and $B = \nabla : H_0^1(\Omega) \rightarrow \mathbb{L}^2(\Omega)$ with $l = d$ and $m = 1$. Then, the convex conjugate of F is

$$F^* : H^{-1}(\Omega) \rightarrow \mathbb{R}, \quad F^*(w) = \frac{1}{2} \|w + K^*y_d\|_{H^{-1}(\Omega)}^2 - \frac{1}{2} \|y_d\|_{L^2(\Omega)}^2$$

with $a^*(w, w) := \|w\|_{H^{-1}(\Omega)}^2 = \langle (K^*K - \beta\Delta)^{-1}w, w \rangle_{H_0^1(\Omega), H^{-1}(\Omega)}$ for $w \in H^{-1}(\Omega)$ with $\Delta : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$, where K^* denotes the adjoint of K . Thus, $z := -K^*y_d \in L^2(\Omega)$ and $b = -\frac{1}{2} \|y_d\|_{L^2(\Omega)}^2$. This yields the Fenchel dual problem

$$\begin{aligned} &\text{minimize} && \frac{1}{2} \|\operatorname{div} p + K^*y_d\|_{H^{-1}(\Omega)}^2 - \frac{1}{2} \|y_d\|_{L^2(\Omega)}^2 && \text{over } p \in \mathbb{L}^2(\Omega) = L^2(\Omega, \mathbb{R}^d) \\ &\text{subject to} && |p|_2 \leq \alpha \text{ a.e. in } \Omega. \end{aligned}$$

As in our previous example, $BF^{*''}(B^*u)B^* \in \mathcal{L}(\mathbb{L}^2(\Omega), \mathbb{L}^2(\Omega))$, but it does not satisfy (4.17), in general. In imaging, typically $d = 2 = l$, $m = 1$, and hence $l + m > 2$. We choose $C := \nabla$, $\mathbb{H} = H_0^1(\Omega, \mathbb{R}^d)$, and $\hat{\mathbb{H}} := H^2(\Omega, \mathbb{R}^d) \cap H_0^1(\Omega, \mathbb{R}^d)$. Then (3.1) holds true due to the Poincaré–Friedrichs inequality. Moreover, as the image domain is usually convex and polygonal, elliptic regularity theory [15] implies that the solution of (3.2)–(3.3) satisfies $p_\gamma \in H^2(\Omega, \mathbb{R}^d) \cap H_0^1(\Omega, \mathbb{R}^d)$. From the Sobolev embedding theorem we see that $H^2(\Omega, \mathbb{R}^d)$ is compactly embedded in $L^\infty(\Omega, \mathbb{R}^d)$ for $d \leq 3$. Similarly, one finds that $\delta_p^{(k)}$ of step (ii) in our semismooth Newton algorithm satisfies $\delta_p^{(k)} \in L^\infty(\Omega, \mathbb{R}^d)$. Hence, assumption (4.1)(b) is fulfilled and the map associated with (3.3) is Newton differentiable.

As a consequence, the general semismooth Newton approach of Section 3 is applicable for the numerical solution of the modified ROF-model.

5.4 Bingham fluids (vector case)

Finally, we revisit Bingham fluids, but now we consider the multidimensional model. In this case, we have $y \in \{w \in \mathbb{H}_0^1(\Omega, \mathbb{R}^d) : \operatorname{div} w = 0\} =: Y$, with $d > 1$,

$$F(y) = \frac{\nu}{2} \int_{\Omega} |D(y)|_2^2 \, dx - \langle \ell, y \rangle_{Y^*, Y}, \quad \nu > 0,$$

where $|D(y)|_2$ represents the Frobenius norm of the $d \times d$ -matrix $D(y) = \frac{1}{2}(\nabla y + \nabla y^\top)$, and $B = D : Y \rightarrow L^2(\Omega, \mathbb{R}^{d,d})$. Using analogous arguments to those in Section 5.2 one finds that the Fenchel dual is given by

$$\begin{aligned} &\text{minimize} && \frac{1}{2} a^*(\operatorname{Div} p + \ell, \operatorname{Div} p + \ell) && \text{over } p \in \mathbb{L}^2(S) = L^2(\Omega, \mathbb{R}^{d,d}) \\ &\text{subject to} && |p|_2 \leq \alpha \text{ a.e. on } \Omega, \end{aligned}$$

where $a^*(w, w) = \frac{1}{\nu} \langle w, (-\mathcal{S})^{-1}w \rangle_{Y^*, Y}$ for $w \in Y$, with $(-\mathcal{S})^{-1}$ denoting the solution operator associated with the Stokes problem and $\nu > 0$ the viscosity coefficient of the fluid. Here, $\ell \in Y^*$ is given data and Div denotes the divergence operator applied to matrix-valued quantities yielding a vector-valued output. As in the scalar situation, our general semismooth Newton framework of Section 3 is applicable.

We point out that similar to the vector-valued Bingham case, there is a vectorial version of the total variation regularization of Section 5.3. It models the restoration of color images within an rgb-context (red-green-blue), where each component of y corresponds to one color channel.

For more information and the application of a semismooth Newton solver to vector-valued Bingham flows and color images restoration, we refer to [7] and [8], respectively.

6. Numerical examples

We end this paper by validating our theoretical findings by means of numerical tests. Here we focus on the Bingham flow problems as described in Sections 5.2 and 5.4. Correspondingly, our first example below considers the scalar case of the Bingham flow problem and the second one the vector case.

For the verification of the convergence rate of the semismooth Newton algorithm we introduce

$$\varrho^{(k)} = \|p^{(k)} - p^{(k-1)}\|_{\mathbb{L}^2}, \quad \vartheta^{(k)} = \frac{\|p^{(k)} - p^{(k-1)}\|_{\mathbb{L}^2}}{\|p^{(k-1)} - p^{(k-2)}\|_{\mathbb{L}^2}}. \quad (6.1)$$

The first quantity $\varrho^{(k)}$ is used to check q -linear and the second one q -superlinear convergence.

For the numerical evaluation of the cost functional and the \mathbb{L}^2 -norms in (6.1), a discrete \mathbb{L}^2 -norm is used, which is obtained by using the rectangle quadrature rule (see [36, p. 373]). The semismooth Newton algorithm is terminated as soon as $\varrho^{(k)}$ is smaller than a given tolerance. The latter is typically chosen as $\text{tol} = 10^{-5}$. Unless otherwise specified, the solutions of the resulting linear systems in each semismooth Newton iteration, i.e., step (ii), were obtained by a direct solver relying on sparse Cholesky techniques. For the use of inexact Newton techniques employing preconditioned iterative solvers we refer to [8, 23]. All algorithms were implemented in MATLAB 7.6 with a machine precision of $\text{eps} = 2.2204\text{e-}16$.

6.1 Example 1

Our first example models a Bingham fluid in the cross section of a pipe. As the test domain, we consider the unit square $\Omega = (0, 1)^2$. A homogeneous Dirichlet boundary condition is imposed along the whole boundary, and $\ell \equiv 10$ is considered. Let us recall that this choice of ℓ corresponds to a linear decay of the pressure in the pipe.

In this case, the operator A^* is the inverse of the Laplacian operator (see Section 5.2) on Y^* . To avoid this inverse, we rather involve the velocity variable $y = A^*B^*p - A^*z$ explicitly and solve the coupled system

$$\begin{pmatrix} A & B^* \\ B & \frac{1}{\gamma}C^*C + \gamma\mathfrak{M}(p) \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_p \end{pmatrix} = \begin{pmatrix} -Ay - B^*p + \ell \\ -By - \frac{1}{\gamma}C^*Cp - \gamma\mathfrak{m}(p)p \end{pmatrix}.$$

Moreover, in this specific application we have

$$C^*C = -\Delta : H_0^1(\Omega) \rightarrow H^{-1}(\Omega) \quad \text{and} \quad B = \nabla : H_0^1(\Omega) \rightarrow \mathbb{L}^2(\Omega),$$

which yields the following system to be solved in each semismooth Newton step:

$$\begin{pmatrix} -v\Delta & -\operatorname{div} \\ \nabla & -\frac{1}{\gamma}\Delta + \gamma\mathfrak{M}(p) \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_p \end{pmatrix} = \begin{pmatrix} v\Delta y + \operatorname{div} p + \ell \\ -\nabla y + \frac{1}{\gamma}\Delta p - \gamma\mathfrak{m}(p)p \end{pmatrix}.$$

6.1.1 *Discretization.* The domain Ω is discretized by an equidistant mesh with $h = 1/(n + 1)$, $n \in \mathbb{N}$, denoting the associated mesh size. We set $N = n \cdot n$. For the discrete Laplacian the standard five-point finite difference stencil is used, whereas the velocity gradient and the dual divergence operators are obtained by using centered differences. The operator C is discretized by using backward differences. For a vector $v \in \mathbb{R}^m$, we denote by $\operatorname{diag}(v)$ the $m \times m$ diagonal matrix with v in the main diagonal. The $m \times m$ identity matrix is denoted by I_m , and we write \otimes for the direct product (or Kronecker product) of matrices.

Introducing the indicator vector act of the active set, which is given by

$$act_i = \operatorname{sign}(\max(|p|_i - \alpha, 0)),$$

where $|p|_i = (p_{1,i}^2 + p_{2,i}^2)^{1/2}$ for $i = 1, \dots, N$ with $p \in \mathbb{R}^{2 \times N}$, the linear system to be solved in each semismooth Newton iteration is given by

$$\begin{pmatrix} -\Delta_h & -B_h^\top \\ B_h & -\frac{1}{\gamma}I_2 \otimes \Delta_h + H + P \cdot D - P \cdot F \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_p \end{pmatrix} = \begin{pmatrix} \Delta_h y + B_h^\top p + \ell \\ \frac{1}{\gamma}(I_2 \otimes \Delta_h)p - Hp - B_h y \end{pmatrix},$$

where $-\Delta_h$ stands for the discrete Laplacian, B_h for the discrete gradient operator,

$$\begin{aligned} H &:= I_2 \otimes \operatorname{diag}(\max(0, \gamma(|p|_i - \alpha)) \cdot |p|_i^{-1}), \\ D &:= \gamma I_2 \otimes \operatorname{diag}(act_i \cdot |p|_i^{-2}), \\ F &:= I_2 \otimes \operatorname{diag}(\max(0, \gamma(|p|_i - \alpha)) \cdot |p|_i^{-3}), \end{aligned}$$

where the respective diagonal entry is set to zero whenever $|p|_i = 0$, and P is the tri-diagonal matrix

$$P = \begin{pmatrix} p_{1,1}^2 & & & p_{1,1}p_{2,1} & & \mathbf{0} \\ & \ddots & & & \ddots & \\ & & p_{1,n}^2 & & & p_{1,N}p_{2,N} \\ p_{2,1}p_{1,1} & & & p_{2,1}^2 & & \\ & \ddots & & & \ddots & \\ \mathbf{0} & & p_{2,N}p_{1,N} & & & p_{2,n}^2 \end{pmatrix}.$$

6.1.2 *Numerical performance.* Figure 1 shows the velocity profile (left) and the Euclidean norm of the dual variable (right) at the internal discretization points, for the data $v = 0.1$, $\alpha = 1.2$, $\gamma = 10^3$ and $n = 120$.

From the graphics it can be observed that the Bingham fluid behaves like a solid in the center of the cross section. This phenomenon is described in detail in [32], where the so-called nucleus of the Bingham inequality is analyzed, and it is verified numerically in, e.g., [6, 12]. Further, a complementarity relation between the velocity and the Euclidean norm of the dual variable may be observed. In fact, the Euclidean norm of the dual takes a constant value of α in the regions where

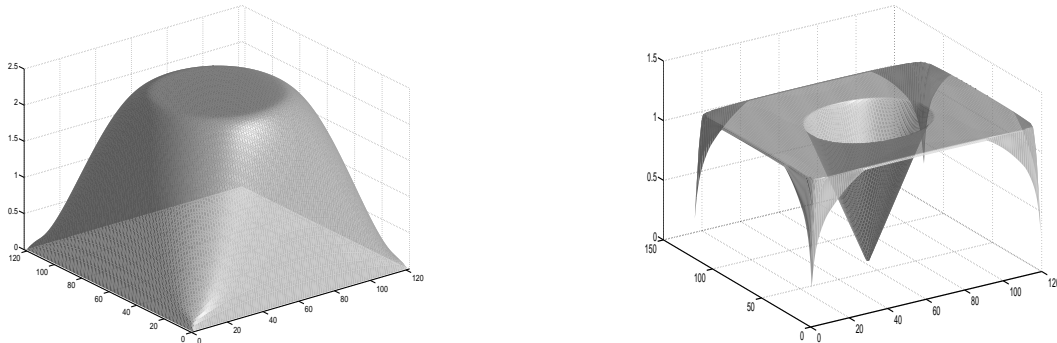


FIG. 1. Example 1: Flow of a Bingham fluid with data $\nu = 0.1$, $\alpha = 1.2$, $\ell = 10$, and regularization parameter $\gamma = 10^3$; velocity profile (left) and Euclidean norm of dual (right).

the material behaves like a fluid. Since the regularized dual variable is in $\mathbb{H}_0^1(\Omega)$, the value of its norm decreases strongly to zero near the boundary of the domain.

The behavior of the semismooth Newton algorithm is depicted in Table 1. The superlinear convergence of the method towards the final iterations can be verified experimentally; compare the last two columns. We also find that the size of the active set decreases monotonically after the third iteration.

TABLE 1

Convergence behavior for Example 1 with $\nu = 0.1$, $\alpha = 1.2$, $\ell = 10$, and regularization parameter $\gamma = 5 \times 10^2$ for $n = 80$.

Iteration	$ \mathcal{A}_k $	$\varrho^{(k)}$	$\vartheta^{(k)}$
1	12800	0.3551	—
2	6518	1.3968	3.934147
3	11456	0.4949	0.354309
4	11196	0.2356	0.475990
5	10796	0.0807	0.342702
6	10354	0.07146	0.885218
7	10328	0.01656	0.231809
8	10088	0.04015	2.424103
9	10072	0.00483	0.120374
10	9970	0.01334	2.759938
11	9932	0.001176	0.088149
12	9928	2.2032e-5	0.018734
13	9928	1.8242e-9	0.000083

6.1.3 Dependence on the mesh. In Table 2 the number of iterations of the semismooth Newton algorithm until successful termination is provided for different mesh sizes h . One clearly finds that the algorithm exhibits a mesh independent convergence behavior as the number of iterations differs at most by ± 1 for the different values of h and fixed $\gamma = 10^3$. This behavior persists for finer mesh sizes. For a theoretical verification of the mesh independence of semismooth Newton methods see [24].

TABLE 2

Example 1: number of iterations vs. mesh size; $\alpha = 1.2$, $\ell = 10$, and $\nu = 0.1$.

n	10	20	30	40	50	80
# it.	15	14	13	15	14	15

TABLE 3

Nested iterations for Example 1: number of iterations vs. mesh size and regularization parameter; $\nu = 1$, $\alpha = 2$, $\ell = 10$.

n	50	100	200	
γ	10^2	10^3	10^5	
# it.	9	8	13	$\Sigma = 30$

As the regularization parameter γ increases, the required number of iterations until successful termination typically increases as well. A coarse-to-fine grid refinement with appropriate initializations on the respective grids helps to reduce the computational effort when solving problems on fine meshes. We study this effect on a mesh with $n = 200$. For $\nu = 1$, $\alpha = 2$ and $\gamma = 10^5$ the numerical solution is obtained after 35 iterations on a grid with mesh step size $n = 200$. The computational effort, and hence the number of semismooth Newton steps on the desired fine mesh, is reduced significantly by considering a nested iteration approach which uses the solution obtained for a smaller regularization parameter on a coarser grid as the initial guess for the semismooth Newton algorithm for the next larger parameter value and next finer mesh size. The solution on the coarser grid is prolonged to the finer grid by using a nine-point prolongation scheme (cf. [17]). According to the results in Table 3, this scheme leads to a reduction of the total number of semismooth Newton iterations, in particular on the fine mesh (with 35 iterations on the fine mesh originally). Additionally, a more efficient computational behavior results from the fact that most of the iterations take place on coarse grids and only the last ones are computed on the finest grid. We point out that our approach represents only the first step towards a fully integrated mesh refinement/regularization parameter update scheme. For the latter we refer to the path-following technique in [21], where the γ -update is guided by a fully automatic extrapolation scheme.

We mention that the techniques of [12] were recently revisited and extended in the context of Bingham fluids [5]. In particular, a smooth penalty approach to the dual problem allowing for classical Newton steps is considered. A numerical comparison of the latter with the primal-dual semismooth Newton method proposed in [23] was exhaustively carried out in [6] showing the advantages of the nonsmooth approach and motivating the further development of semismooth Newton methods for the numerical solution of Bingham fluids.

Also, a solution technique based on introducing a new variable p which replaces ∇y in the energy associated with the Bingham fluid and then penalizing the condition $p = \nabla y$ in an augmented Lagrangian fashion is analysed in [5]. Such an approach has recently become popular in total variation based image processing.

In the case of Bingham fluids, our methodology proposed in this paper relies on dualization and the Moreau–Yosida regularization of the indicator function of the dual feasible set. It therefore contains features of a penalty method, but, in contrast to [5], the resulting Euler–Lagrange system is

not necessarily Fréchet differentiable. Rather it is semismooth only, and thus reduces degeneracies, i.e., very flat transitions into the active set.

In [1] a mixed formulation for numerically solving the minimization problem associated with Bingham fluids was considered. The authors use the Bercovier–Engelman regularization of the norm, i.e., $\sqrt{\cdot} \approx \sqrt{\cdot + \tau^2}$, yielding a C^∞ -regularization. It is observed that a smooth version of Newton’s method solving the regularized problem deteriorates as τ^2 is decreased towards zero. Thus, the authors rather rely on a Picard iteration. Moreover, the Bercovier–Engelman regularization has to be handled carefully in order to be able to detect arrested states. In our approach arrested states can be identified clearly, which is due to the non-smoothness in the Euler–Lagrange system and the application of a semismooth Newton solver. Further, our method is shown to be stable under mesh refinements and increasing γ .

6.2 Example 2

We consider the flow of a visco-plastic two-dimensional Bingham fluid in the unit square $\Omega = (0, 1)^2$. Homogeneous Dirichlet boundary conditions are imposed on the whole boundary, and the volumetric force $\ell = 300(x_2 - 0.5, 0.5 - x_1)^\top$ is considered. With this data, the associated VI models a visco-plastic fluid in a reservoir (see [9] for further details).

As in the scalar variant, the proposed semismooth Newton algorithm involves the solution of a Stokes-like problem in every iteration. For its numerical implementation, both the velocity vector field y and the pressure q are considered explicitly. The resulting system to be solved in each iteration has the following form:

$$\begin{pmatrix} -v\vec{\Delta} & \nabla & -\text{Div} \\ -\text{div} & 0 & 0 \\ -\mathcal{T} & 0 & -\frac{1}{\gamma}\vec{\vec{\Delta}} + \gamma\mathfrak{M}(p) \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_q \\ \delta_p \end{pmatrix} = \begin{pmatrix} v\vec{\Delta}y - \nabla q + \text{Div } p + f \\ -\text{div } y \\ \frac{1}{\gamma}\vec{\vec{\Delta}}p - \gamma\mathfrak{m}(p)p + \mathcal{T}y \end{pmatrix}$$

where $\vec{\Delta} := I_2 \otimes \Delta$, $\vec{\vec{\Delta}} := I_4 \otimes \Delta$ and \mathcal{T} is the partial derivative operator given by

$$\mathcal{T} := \begin{pmatrix} \partial_1 & 0 \\ \frac{1}{2}\partial_2 & \frac{1}{2}\partial_1 \\ \frac{1}{2}\partial_2 & \frac{1}{2}\partial_1 \\ 0 & \partial_2 \end{pmatrix}.$$

6.2.1 Discretization. The system was discretized using a homogeneous MAC scheme (see [4, 33]). In this scheme, a semi-staggered grid is utilized, i.e., the values of the two velocity components are taken on the grid points, while the pressure is considered at the center of each square cell; see Figure 2. To obtain compatibility between the variables, the discrete dual quantity is also considered on the grid points.

Following [33, Section 3.2], for the discretization of the Laplacian we use the nine-point approximation:

$$\Delta_h v = \frac{v_{i-1,j-1} + v_{i+1,j-1} + v_{i-1,j+1} + v_{i+1,j+1} - 4v_{i,j}}{2h^2},$$

yielding the matrix A . For the gradient of the pressure, the approximation

$$\nabla_h q = \left(\frac{q_{i,j} - q_{i-1,j} + q_{i,j-1} - q_{i-1,j-1}}{2h}, \frac{q_{i,j} - q_{i,j-1} + q_{i-1,j} - q_{i-1,j-1}}{2h} \right)$$

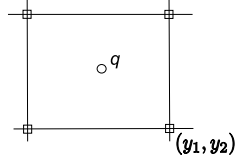


FIG. 2. Non-staggered grid.

is utilized, yielding the matrix G . For the divergence, the vector divergence of the dual and the rate of deformation tensor, we consider the approximations proposed in [33, p. 626]. The resulting approximations of the partial derivatives, to be applied to the velocity, are then denoted by E_1 and E_2 . In addition, since the dual variable has the matrix form

$$p(x) = \begin{pmatrix} p_1(x) & p_2(x) \\ p_3(x) & p_4(x) \end{pmatrix},$$

the discretization of each component $p_i(x)$, $i = 1, \dots, 4$, yields a vector $\vec{p}_{i,k}$, $k = 1, \dots, N$. By vectorizing the matrix, we obtain the vector $\vec{p} = (\vec{p}_1 \vec{p}_2 \vec{p}_3 \vec{p}_4)^\top$. We denote the Hadamard product of vectors by \star and introduce the matrices

$$P_{i,j} := \text{diag}(\vec{p}_i \star \vec{p}_j),$$

and the block matrices

$$P_i = (P_{i,1} \ P_{i,2} \ P_{i,3} \ P_{i,4}).$$

As before, we introduce the indicator vector of the active set by

$$act_i = \text{sign}(\max(|p|_i - \alpha, 0)),$$

with the appropriate modification of $|\cdot|_i$.

Since the MAC discretization scheme considered is not stable, we use an additional stabilization through a penalty term in the incompressibility condition (see [16, p. 125]). An alternative stabilization procedure is considered, for instance, in [33, p. 628].

The resulting system to be solved in each iteration is then

$$\begin{pmatrix} A_S & G & B \\ C & \varepsilon I & 0 \\ -T & 0 & R \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_q \\ \delta_p \end{pmatrix} = \begin{pmatrix} -A_S y - Gq - Bp + f \\ -Cy - \varepsilon q \\ -\frac{1}{\gamma}(I_4 \otimes A)p - Hp + Ty \end{pmatrix}$$

where $0 < \varepsilon \ll 1$, $A_S = I_2 \otimes A$, $E = (E_1 \ E_2)^\top$, $B = -I_2 \otimes E$,

$$T = \begin{pmatrix} E_1 & 0 \\ \frac{1}{2}E_2 & \frac{1}{2}E_1 \\ \frac{1}{2}E_2 & \frac{1}{2}E_1 \\ 0 & E_2 \end{pmatrix}$$

and

$$R = I_4 \otimes \left(\frac{1}{\gamma} A + H \right) + \begin{pmatrix} P_1(D - F) & 0 & 0 & 0 \\ 0 & P_2(D - F) & 0 & 0 \\ 0 & 0 & P_3(D - F) & 0 \\ 0 & 0 & 0 & P_4(D - F) \end{pmatrix},$$

with

$$H := \text{diag}(\max(0, \gamma(|p|_i - \alpha)) \cdot |p|_i^{-1}),$$

$$D := \gamma \text{diag}(act_i \cdot |p|_i^{-2}),$$

$$F := \text{diag}(\max(0, \gamma(|p|_i - \alpha)) \cdot |p|_i^{-3}),$$

where, as before, the diagonal entries are set to zero whenever $|p|_i = 0$.

6.2.2 Numerical performance. For the numerical experiments we consider the material data $\nu = 1$ and $\alpha = 10$ and the penalty parameter $\varepsilon = \sqrt{\epsilon ps}$. The solution of the corresponding problem with $\gamma = 10^4$ and $h = 1/121$ is shown in Figure 3, which depicts the velocity vector field (left) and the resulting solid-liquid zones (right).

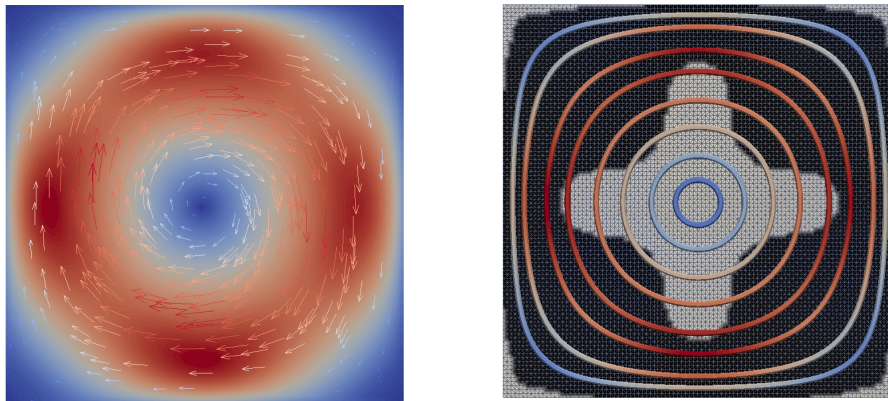


FIG. 3. Example 2: Velocity vector field (left) and solid-liquid zones (right) of a two-dimensional Bingham fluid with data $\nu = 1$, $\alpha = 10$ and $f = 300(x_2 - 0.5, 0.5 - x_1)^T$, and regularization parameter $\gamma = 10^4$.

The numerical behavior of the method for these parameter choices is displayed in Table 4. The superlinear convergence rate in the final iterations can be inferred from our results, as well as the monotone decrease of the size of the active set in the last iterations.

6.2.3 Line search. Although the global convergence result of Section 4 relies on the use of a line search, we found in our tests that the semismooth Newton method converges globally even without such a line search scheme. However, using the Armijo alternative step (iii') instead of (iii) in the semismooth Newton algorithm, the number of iterations until successful termination is significantly increased; see Table 5. This behavior can be attributed to the fact that the Armijo

TABLE 4

Example 2: $\alpha = 10$, $f = 300(x_2 - 0.5, 0.5 - x_1)^\top$, $\gamma = 10^4$, $\nu = 1$, and $h = 1/121$.

Iteration	$ \mathcal{A}_k $	ϱ_k	ϑ_k
1	57600	4.979277	-
2	30484	11.138236	2.236918
3	53720	7.002363	0.628677
4	49656	2.316722	0.330848
5	51536	0.645666	0.278698
6	46964	1.315373	2.037231
7	47912	0.181672	0.138114
8	45440	0.524227	2.885562
9	45540	0.058119	0.110867
10	44272	0.381752	6.568363
11	44320	0.049032	0.128440
12	43496	0.202268	4.125209
13	43520	0.013150	0.065016
14	42964	0.059831	4.549660
15	42968	0.006041	0.100968
16	42724	0.011167	1.848516
17	42724	0.002543	0.227766
18	42700	0.001978	0.777842
19	42696	0.000793	0.401097
20	42680	0.000289	0.364833
21	42680	1.5149e-5	0.052329
22	42680	1.2557e-7	0.008288

TABLE 5

Example 2: $\alpha = 5$, $f = 300(x_2 - 0.5, 0.5 - x_1)^\top$, $\nu = 1$ and $h = 1/121$.

γ	Iter. with line search	Iter. without line search
10^2	15	13
10^3	30	16
10^4	41	17

line search reduces the step length in many iterations so that the full Newton step is only accepted later in the iteration history. Utilizing a non-monotone line search technique (see, e.g., [14]) may overcome this adverse effect.

7. Extensions

Several extensions of our framework are possible. Here we confine ourselves to the following two:

First, the ℓ_2 -norm in the definition of $\int_S |By|_2 ds$ may be generalized to $\int_S |By|_t ds$ for $t \geq 1$. Then the dual contains the associated dual norm $|\cdot|_{t'}$ with $t^{-1} + t'^{-1} = 1$. As this change only affects finite dimensions, all of our function space results remain true.

Secondly, occasionally the ‘forward’ dualization, i.e., starting with (P) and arriving at (DP) upon employing the Fenchel–Legendre calculus, is not possible due to function space regularity. In

such cases it might be worthwhile to study the Fenchel pre-dual of (P) , which is a function space problem whose dual is given by (P) . In some situations the pre-dual admits a structure similar to the one of (DP) with $p \in \mathbb{H}_p(S) \subset \mathbb{L}^2(S)$ where $\mathbb{H}_p(S)$ denotes some Hilbert space which is continuously embedded in $\mathbb{L}^2(S)$. As a result, our framework may be applied to the pre-dual problem yielding an efficient solution algorithm in function space.

Such a situation occurs, for instance, in total variation based image restoration. Here the pre-dual is given by

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \|\operatorname{div} p + K^* y_d\|_{L^2(\Omega)}^2 && \text{over } p \in \mathbb{H}_p(\Omega) = H_0(\operatorname{div}, \Omega) \\ & \text{subject to} && |p|_2 \leq \alpha \text{ a.e. in } \Omega, \end{aligned}$$

where $H_0(\operatorname{div}, \Omega) = \{p \in L^2(\Omega, \mathbb{R}^2) : \operatorname{div} p \in L^2(\Omega), p \cdot \eta = 0 \text{ on } \partial\Omega\}$ with η the outward unit normal on $\partial\Omega$, the boundary of Ω . The Fenchel dual of the above problem is the total variation regularized problem

$$\text{minimize} \quad \frac{1}{2} \|Ky - y_d\|_{L^2(\Omega)}^2 + \alpha \int_{\Omega} |Dy| \quad \text{over } y \in \operatorname{BV}(\Omega),$$

where $\operatorname{BV}(\Omega)$ denotes the space of functions of bounded variation and $\int_{\Omega} |Dy|$ the usual BV-seminorm; see [11].

Acknowledgements

Research supported by the Austrian Federal Ministry of Science and Research and the FWF under START-program Y305 “Interfaces and Free Boundaries”, the Alexander von Humboldt Foundation, the DFG research center Matheon under subproject C28 and the DFG priority program SPP1253 on “Optimization with PDE constraints”.

REFERENCES

1. AOSPORIDIS, A., HABER, E., OLSHANSKII, M. A., & VENEZIANI, A. A mixed formulation of the Bingham fluid flow problem: Analysis and numerical solution. *Comput. Methods Appl. Mech. Engrg.* **200** (2011), 2434–2446. [MR 2803637](#)
2. CHENG, X.-I., & HAN, W. Inexact Uzawa algorithms for variational inequalities of the second kind. *Comput. Methods Appl. Mech. Engrg.* **192** (2003), 1451–1462. [Zbl 1033.65045](#) [MR 1963060](#)
3. BREZIS, H. Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations. In: *Contributions to Non-linear Functional Analysis*, E. Zarantonello (ed.), Academic Press (1971), 101–156. [Zbl 0278.47033](#) [MR 0394323](#)
4. DAUTRAY, R., & LIONS, J.-L. *Mathematical Analysis and Numerical Methods in Science and Technology*. Springer (2000).
5. DEAN, E. J., GLOWINSKI, R., & GUIDOBONI, G. On the numerical simulation of Bingham visco-plastic flow: old and new results. *J. Non-Newtonian Fluid Mech.* **142** (2007), 36–62. [Zbl 1107.76061](#)
6. DE LOS REYES, J. C., & GONZÁLEZ, S. Path following methods for steady laminar Bingham flow in cylindrical pipes. *Math. Model. Numer. Anal.* **43** (2009), 81–117. [Zbl 1159.76033](#) [MR 2494795](#)
7. DE LOS REYES, J. C., & GONZÁLEZ, S. Numerical simulation of two-dimensional Bingham fluid flow by semismooth Newton methods. *J. Comput. Appl. Math.* **235** (2010), 11–32. [Zbl pre05804863](#)

8. DONG, Y., & HINTERMÜLLER, M. Multi-scale vectorial total variation with automated regularization parameter selection for color image restoration. In: *Scale Space and Variational Methods in Computer Vision* (Voss, 2009), Lecture Notes in Comput. Sci. 5567, Springer (2003), 271–281.
9. DUVAUT, G., & LIONS, J.-L. *Inequalities in Mechanics and Physics*. Springer, Berlin (1976). [Zbl 0331.35002](#) [MR 0521262](#)
10. EKELAND, I., & TEMAM, R. *Convex Analysis and Variational Problems*. Corrected republication of the 1976 English original, Classics Appl. Math. 28, SIAM, Philadelphia (1999). [Zbl 0939.49002](#) [MR 1727362](#)
11. GIUSTI, E. *Minimal Surfaces and Functions of Bounded Variation*. Birkhäuser, Boston (1984). [Zbl 0545.49018](#) [MR 0775682](#)
12. GLOWINSKI, R. *Numerical Methods for Nonlinear Variational Problems*. Springer (1982).
13. GLOWINSKI, R., LIONS, J.-L., & TREMOLIERS, R. *Analyse Numérique des Inéquations Variationnelles, Vol. 2: Applications aux Phénomènes Stationnaires et d'Évolution*. Dunod-Bordas, Paris (1976). [Zbl 0358.65091](#) [MR 0655455](#)
14. GRIPPO, L., LAMPARIELLO, F., & LUCIDI, S. A nonmonotone line search technique for Newton's method. *SIAM J. Numer. Anal.* **23** (1986), 707–716. [Zbl 0616.65067](#) [MR 0849278](#)
15. GRISVARD, P. *Elliptic Problems in Nonsmooth Domains*. Monographs Stud. Math. 24, Pitman, London (1985). [Zbl 0695.35060](#) [MR 0775683](#)
16. GUNZBURGER, M. Navier–Stokes equations for incompressible flows: finite-element methods. In: *Handbook of Computational Fluid Mechanics*, Academic Press, San Diego, (2000), 99–158. [Zbl 1107.76351](#)
17. HACKBUSCH, W. *Multigrid Methods and Applications*. Springer, Berlin (1985). [MR 0814495](#)
18. HINTERMÜLLER, M., ITO, K., & KUNISCH, K. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.* **13** (2002), 865–888. [Zbl 1080.90074](#) [MR 1972219](#)
19. HINTERMÜLLER, M., & KOPACKA, I. Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm. *SIAM J. Optim.* **20** (2009), 868–902. [Zbl 1189.49032](#) [MR 2515801](#)
20. HINTERMÜLLER, M., & KUNISCH, K. Total bounded variation regularization as a bilaterally constrained optimization problem. *SIAM J. Appl. Math.* **64** (2004), 1311–1333. [Zbl 1055.94504](#) [MR 2068672](#)
21. HINTERMÜLLER, M., & KUNISCH, K. Path-following methods for a class of constrained minimization problems in function space. *SIAM J. Optim.* **17** (2006), 159–187. [Zbl 1137.49028](#) [MR 2219149](#)
22. HINTERMÜLLER, M., & KUNISCH, K. PDE-constrained optimization subject to pointwise constraints on the control, the state, and its derivative. *SIAM J. Optim.* **20** (2009), 1133–1156. [Zbl 1195.49037](#) [MR 2546337](#)
23. HINTERMÜLLER, M., & STADLER, G. An infeasible primal-dual algorithm for TV-based inf-convolution-type image restoration. *SIAM J. Sci. Comput.* **28** (2006), 1–23. [Zbl 1136.94302](#) [MR 2219285](#)
24. HINTERMÜLLER, M., & ULBRICH, M. A mesh-independence result for semismooth Newton methods, *Math. Program.* **101** (2004), 151–184. [Zbl 1079.65065](#) [MR 2085262](#)
25. HIRIART-URRUTY, J.-P., & LEMARÉCHAL, C. *Fundamentals of Convex Analysis*. Springer Grundlehren Text Editions, Springer, Berlin (2001). [Zbl 0998.49001](#) [MR 1865628](#)
26. HOPPE, R. H. W. Multigrid algorithms for variational inequalities. *SIAM J. Numer. Anal.* **24** (1987), 1046–1065. [Zbl 0628.65046](#) [MR 0909064](#)
27. HOPPE, R. H. W., & KORNHUBER, R. Adaptive multilevel methods for obstacle problems. *SIAM J. Numer. Anal.* **31** (1994), 301–323. [Zbl 0806.65064](#) [MR 1276702](#)
28. KIKUCHI, N., & ODEN, J. T. *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*. SIAM Stud. Appl. Math. 8, SIAM (1988). [Zbl 0685.73002](#) [MR 0961258](#)

29. KORNHUBER, R. Monotone multigrid methods for elliptic variational inequalities I. *Numer. Math.* **69** (1994), 167–184. [Zbl 0817.65051](#) [MR 1310316](#)
30. KORNHUBER, R. Monotone multigrid methods for elliptic variational inequalities II. *Numer. Math.* **72** (1996), 481–499. [Zbl 0861.65056](#) [MR 1376109](#)
31. MIFFLIN, R. Semismooth and semiconvex functions in constrained optimization. *SIAM J. Control Optim.* **15** (1977), 959–972. [Zbl 0376.90081](#) [MR 0461556](#)
32. MOSOLOV, P. P., & MIASNIKOV, V. P. Variational methods in the theory of the fluidity of a viscous-plastic medium. *J. Appl. Math. Mech.* **29** (1965), 545–577. [Zbl 0168.45505](#)
33. MURAVLEVA, E. A., & OLSHANSKII, M. A. Two finite-difference schemes for calculation of Bingham fluid flows in a cavity. *Russian J. Numer. Anal. Math. Modelling* **23** (2008), 615–634. [Zbl 1151.76539](#) [MR 2482805](#)
34. NOCEDAL, J., & WRIGHT, S. J. *Numerical Optimization*. Springer Ser. Oper. Res., Springer, New York (1999). [Zbl 0930.65067](#) [MR 1713114](#)
35. QI, L., & SUN, J. A nonsmooth version of Newton’s method. *Math. Program. Ser. A* **58** (1993), 353–367. [Zbl 0780.90090](#) [MR 1216791](#)
36. QUARTERONI, A., SACCO, R., & SALERI, F. *Numerical Mathematics*. Springer, Berlin (2007). [Zbl 1136.65001](#) [MR 2265914](#)
37. REDDY, B. D. Mixed variational inequalities arising in elastoplasticity. *Nonlinear Anal.* **19** (1992), 1071–1089. [Zbl 0768.49008](#) [MR 1194146](#)
38. RUDIN, L. I., OSHER, S. J., & FATEMI, E. Nonlinear total variation based noise removal algorithms. *Phys. D* **60** (1992), 259–268. [Zbl 0780.49028](#)
39. STADLER, G. Semismooth Newton and augmented Lagrangian methods for a simplified friction problem. *SIAM J. Optim.* **18** (2004), 39–62. [Zbl 1106.90078](#) [MR 2112975](#)
40. WLOKA, J. *Partial Differential Equations*. Cambridge Univ. Press, Cambridge (1987). [Zbl 0623.35006](#) [MR 0895589](#)