Semi-Smooth Newton Methods for the Signorini Problem

Kazufumi Ito[1] Karl Kunisch[2]

July 23, 2008

[1]Department of Mathematics, North Carolina State University, Raleigh, North Carolina, 27695-8205, USA; research partially supported by the Army Research Office under DAAD19-02-1-0394

[2]Institut für Mathematik und wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, A-8010 Graz, Austria, supported in part by the Fonds zur Förderung der wissenschaftlichen Forschung under SFB 32, "Mathematical Optimization and Applications in the Biomedical Sciences".

Abstract Semi-smooth Newton methods are analyzed for the Signorini problem. A proper regularization is introduced which guarantees that the semi-smooth Newton method is superlinearly convergent for each regularized problem. Utilizing a shift motivated by an augmented Lagrangian framework, to the regularization term, the solution to each regularized problem is feasible. Convergence of the regularized problems is shown and a report on numerical experiments is given.

Keywords Signorini problem, variational inequality, semi-smooth Newton method, primal-dual active set strategy.

1 Introduction

The objective of this paper is to analyze a Newton type method for the following Signorini problem:

$$(Sig) \qquad \begin{cases} \min \frac{1}{2}\int_\Omega |\nabla u|^2 - \int_{\Gamma_N} qu - \int_\Omega fu \\ \text{subject to } \; u \in H^1(\Omega), \quad u = 0 \text{ on } \Gamma_D, \quad u \le \psi \text{ on } \Gamma, \end{cases}$$

where Ω is a bounded domain with boundary consisting of the disjoint subsets Γ_N, Γ_D and Γ. The inequality constraint $u \le \psi$ appears at first sight to impede the Newton method. But following the recent developments of semi-smooth Newton methods in functions spaces, see e.g. [HIK, HK, IK, U], we shall show that superlinear methods for solving (Sig) can be developed. We shall introduce a Lagrangian framework for a family of regularized problems and prove their convergence as the regularization parameter tends to its limit. Each of the regularized problems can be solved by a semi-smooth Newton method with local superlinear convergence rate. The regularization differs from penalty type methods by involving a shift $\bar{u}$ which is the solution to the following auxiliary problem.

$$(Aux) \qquad \begin{cases} -\Delta \bar{u} = f \text{ in } \Omega \\ \bar{u} = 0 \text{ on } \Gamma_D, \quad \dfrac{\partial \bar{u}}{\partial n} = q \text{ on } \Gamma_N, \quad \bar{u} = \psi \text{ on } \Gamma. \end{cases}$$

.

Introducing the shift is suggested by augmented Lagrangian concepts. For the problem under consideration it will guarantee that the approximating solutions are all feasible. Section 2 contains the exact problem formulation and the convergence of the regularized problems. The semi-smooth Newton method is developed in Section 3. A short description of numerical experiments is given in the final Section 4

2 Problem formulation and monotone, feasible approximation

Let $\Omega \subset \mathbb{R}^2$ be a rectangular domain with lateral boundaries Γ_D, top boundary Γ_N and bottom boundary Γ and consider the Signorini problem

$$(2.1)\qquad \begin{cases} \min \frac{1}{2}\int_\Omega |\nabla u|^2 - \int_{\Gamma_N} qu - \int_\Omega fu \\ \text{subject to } \; u \in H^1(\Omega), \quad u = 0 \text{ on } \Gamma_D, \quad u \leq \psi \text{ on } \Gamma. \end{cases}$$

Here $f \in L^2(\Omega)$, $q \in L^2(\Gamma_N)$ and

$$\psi = \hat{\psi}|\Gamma \text{ with } \hat{\psi} \in H^1(\Omega) \text{ and } \hat{\psi}|(\Gamma_N \cup \Gamma_D) = 0.$$

In particular this implies that $\psi \in H^{\frac{1}{2}}_{0,0}(\Gamma)$, i.e. $\psi \in H^{\frac{1}{2}}(\Gamma)$ and $\psi = 0$ in an integral sense on the boundaries of Γ, [G], pg. 44. Associated to (2.1) we define the Lagrangian $\mathcal{L} : H^1_{0,D}(\Omega) \times H^{-\frac{1}{2}}_{0,0}(\Gamma) \to \mathbb{R}$ by

$$\mathcal{L}(u,\lambda) = \frac{1}{2}\int_\Omega |\nabla u|^2 - \int_{\Gamma_N} qu - \int_\Omega fu + \langle \lambda, u - \psi \rangle_\Gamma,$$

where $H^1_{0,D}(\Omega) = \{\varphi \in H^1(\Omega) : \varphi|\Gamma_D = 0\}$, and $\langle \cdot,\cdot \rangle_\Gamma$ denotes the duality pairing between $H^{\frac{1}{2}}_{0,0}(\Gamma)$ and $H^{-\frac{1}{2}}_{0,0}(\Gamma)$.

Problem (2.1) admits a unique solution denoted by $u^* \in H^1_{0,D}(\Omega)$. Let $g : H^1_{0,D}(\Omega) \to H^{\frac{1}{2}}_{0,0}(\Gamma)$ denote the mapping describing the inequality constraint in (2.1), i.e. $g(u) = u|\Gamma - \psi$. Its linearization at u^* is surjective and hence there exists a Lagrange multiplier $\lambda^* \in H^{-\frac{1}{2}}_{0,0}(\Gamma)$ which renders $\mathcal{L}$ stationary at (u^*, λ^*), i.e.

(2.2)
$$\begin{cases} \displaystyle\int_\Omega \nabla u^* \nabla v - \int_{\Gamma_N} qv - \int_\Omega fv + \langle \lambda^*, v\rangle_\Gamma = 0 \quad \text{for all } v \in H^1_{0,D}(\Omega) \\ \langle \lambda^*, u^* - \psi\rangle_\Gamma = 0, \quad u^* \le \psi, \quad \langle \lambda^*, v\rangle_\Gamma \ge 0 \quad \text{for all } v \in H^{\frac{1}{2}}_{00}(\Gamma),\ v \ge 0, \end{cases}$$

which can formally be expressed as

$$\begin{cases} -\Delta u^* = f \text{ in } \Omega \\ u^* = 0 \text{ on } \Gamma_D,\ \dfrac{\partial u^*}{\partial n} = q \text{ on } \Gamma_N, \quad \dfrac{\partial u^*}{\partial n} = -\lambda^* \text{ on } \Gamma \\ u^* \le \psi, \quad \lambda^* \ge 0, \quad \lambda^*(u^* - \psi) = 0. \end{cases}$$

The solution $\bar{u} \in H^1_{0,D}(\Omega)$ of the following problem will play a significant role

$$\begin{cases} -\Delta \bar{u} = f \text{ in } \Omega \\ \bar{u} = 0 \text{ on } \Gamma_D, \quad \dfrac{\partial \bar{u}}{\partial n} = q \text{ on } \Gamma_N, \quad \bar{u} = \psi \text{ on } \Gamma. \end{cases} \tag{2.3}$$

We recall from e.g. [G] pg. 27 that $\frac{\partial \bar{u}}{\partial n} \in H^{-\frac{1}{2}}_{0,0}(\Gamma)$. Moreover, if

$$q \in H^{\frac{1}{2}}_{0,0}(\Gamma), \tag{2.4}$$

then

$$\bar{u} \in H^2(\Omega), \tag{2.5}$$

and in particular

$$\frac{\partial \bar{u}}{\partial n}|\Gamma \in H^{\frac{1}{2}}(\Gamma). \tag{2.6}$$

Similarly

$$\frac{\partial u^*}{\partial n}|\Gamma \in H^{\frac{1}{2}}(\Gamma) \subset L^q(\Gamma) \quad \text{for every } \ q \ge 1, \tag{2.7}$$

if (2.4) holds. In what follows we shall utilize a function $\bar{\lambda} \in L^2(\Gamma)$ satisfying

$$\bar{\lambda} \ge 0 \text{ and } \big\langle \bar{\lambda} + \frac{\partial \bar{u}}{\partial n},\ v\big\rangle \ge 0 \text{ for all } v \in H^{\frac{1}{2}}_{0,0}(\Gamma), v \ge 0. \tag{2.8}$$

In case of (2.6) we can choose

$$\bar{\lambda} = \max(0, \frac{-\partial \bar{u}}{\partial n}), \tag{2.9}$$

where max denotes the pointwise a.e. maximum along Γ.

For every $c > 0$ we consider the regularized problem

$$\min_{u \in H^1_{0,D}(\Omega)} \frac{1}{2} \int_\Omega |\nabla u|^2 - \int_{\Gamma_N} qu - \int_\Omega uf + \frac{1}{2c} \int_\Gamma |\lambda_c|^2, \tag{2.10}$$

where $\lambda_c = \max(0, \bar{\lambda} + c(u - \psi))$. Clearly (2.10) admits a unique solution $u_c \in H^1_{0,D}(\Omega)$. It satisfies the variational form of the following equation

$$\begin{cases} -\Delta u_c = f \text{ in } \Omega \\ \frac{\partial u_c}{\partial n} = q \text{ on } \Gamma_N, \quad u_c = 0 \text{ on } \Gamma_D \\ \frac{\partial u_c}{\partial n} = -\lambda_c \text{ on } \Gamma. \end{cases} \tag{2.11}$$

If

$$\bar{\lambda} \in H^{\frac{1}{2}}_{0,0}(\Gamma), \tag{2.12}$$

then $\lambda_c \in H^{\frac{1}{2}}_{0,0}(\Gamma)$, and if (2.4) and (2.12) hold then $u_c \in H^2(\Omega)$.

Proposition 2.1. *Let* (2.8) *hold. Then for each* $c > 0$ *we have*

$$u_c \leq \psi \text{ on } \Gamma_c.$$

Proof. Note that $(u_c - \bar{u})^+ \in H^1_{0,D}(\Omega)$ and $(u_c - \bar{u})^+ |\Gamma_N \in H^{\frac{1}{2}}_{0,0}(\Gamma_N)$, $(u_c - \bar{u})^+ \,|\Gamma \in |H^{\frac{1}{2}}_{0,0}(\Gamma)$. Consequently

$$|\nabla (u_c - \bar{u})^+|^2_\Omega - \langle \frac{\partial}{\partial n}(u_c - \bar{u}), (u_c - \bar{u})^+ \rangle_{\frac{1}{2}} = 0,$$

and

$$|\nabla (u_c - \bar{u})^+|^2_\Omega - \langle \bar{\lambda} + \frac{\partial \bar{u}}{\partial n},\, (u_c - \bar{u})^+ \rangle_{\frac{1}{2}} = 0.$$

By (2.8) and since $\Gamma_D \neq \emptyset$ this implies that $u_c \leq \bar{u}$ in $H^1_{0,D}(\Omega)$. Consequently $u_c \leq \bar{u} = \psi$ on Γ. $\square$

Corollary 2.1. *If* (2.8) *holds, then*

$$0 \le \lambda_c = \max\big(0, \bar\lambda + c(u_c - \psi)\big) \le \bar\lambda \textit{ for each } c > 0.$$

Proposition 2.2. *Let* (2.8) *hold. Then for any $c \le \bar c$ we have*

$$u_c \le u_{\bar c} \quad in \quad H^1_{0,D}(\Omega).$$

Proof. Again $(u_c - u_{\bar c})^+ \in H^1_{0,D}(\Omega)$ and $(u_c - u_{\bar c})^+ \,|\Gamma_N \in H^{\frac12}_{0,0}(\Gamma_N)$, $(u_c - u_{\bar u})^+|\,\Gamma \in H^{\frac12}_{0,0}(\Gamma)$. Consequently

$$|\nabla(u_c - u_{\bar c})^+|^2_\Omega - \langle \frac{\partial}{\partial n}(u_c - u_{\bar c}),\, (u_c - u_c)^+\rangle_{\frac12} = 0,$$

and

$$|\nabla(u_c - u_{\bar c})^+|^2_\Omega + (\lambda_c - \lambda_{\bar c},\, (u_c - u_c)^+)_\Gamma = 0. \tag{2.13}$$

Note that

$$\begin{aligned}
&(\lambda_c - \lambda_{\bar c}, (u_c - u_{\bar c})^+)_\Gamma \\
&\qquad = \big(\max(0, \bar\lambda + c(u_c - \psi)\big) - \max\big(0, \bar\lambda + \bar c(u_{\bar c} - \psi)\big),\, (u_c - u_{\bar c})^+\big)_\Gamma \\
&\qquad = (\max(0, \bar\lambda + c(u_c - \psi)) - \max(0, \bar\lambda + c(u_{\bar c - \psi})), (u_c - u_{\bar c)})^+)_\Gamma \\
&\qquad + (\max(0, \bar\lambda + c(u_c - \psi)) - \max(0, \bar\lambda + \bar c(u_{\bar c} - \psi)), (u_c - u_{\bar c})^+)_\Gamma \\
&\qquad \ge (\max(0, \bar\lambda + c(u_c - \psi)) - \max(0, \bar\lambda + c(u_{\bar c} - \psi)), (u_c - u_{\bar c})^+_\Gamma, \ge 0,
\end{aligned}$$

where in the next to last step we used Proposition 2.1. The claim now follows from (2.13). □

Theorem 2.1. *If* (2.8) *holds then* (u_c, λ_c) *converges to* (u^*, λ^*) *in the sense that* $u_c \to u^*$ *in* $H^1_D(\Omega)$ *and* $\lambda_c \rightharpoonup \lambda^*$ *weakly in* $L^2(\Gamma)$*, as* $c \to \infty$.

Proof. From (2.11) we have

$$|\nabla u_c|^2_\Omega = (f, u_c)_\Omega + (q, u_c)_{\Gamma_N} - (\lambda_c, u_c)_\Gamma. \tag{2.14}$$

Since

$$
\begin{aligned}
(\lambda_c, u_c)_\Gamma = &\big(\max(0, \bar\lambda + c(u_c - \psi)),\, u_c - \psi\big)_\Gamma \\
&+\big(\max(0, \bar\lambda + c(u_c - \psi)), \psi\big)_\Gamma \le (\bar\lambda, \psi)_\Gamma,
\end{aligned}
$$

it follows from (2.14) that $\{u_c\}_{c\ge 1}$ is bounded in $H^1_{0,D}(\Omega)$. Together with Corollary (2.1) this implies that $\{(u_c, \lambda_c)\}_{c\ge 1}$ is bounded in $H^1_{0,D}(\Omega) \times L^2(\Omega)$ and hence there exist $(\hat u, \hat\lambda) \in H^1_{0,D}(\Omega) \times L^2(\Omega)$ and a subsequence such that $(u_c, \lambda_c) \rightharpoonup (\hat u, \hat\lambda)$ weakly in $H^1_{0,D}(\Omega) \times L^2(\Omega)$. Clearly $\hat u \le \psi$ and $\hat\lambda \ge 0$ on Γ. By (2.2), (2.11) and $u_c \le \psi$ on Γ

$$
\begin{aligned}
|\nabla(u_c - u^*)|^2 &= \langle \lambda^* - \lambda_c,\, u_c - u^* \rangle_{\frac12} = \langle \lambda^* - \lambda_c,\, u_c - \psi + \psi - u^* \rangle_{\frac12} \\
&\le -(\lambda_c,\, u_c - \psi + \psi - u^*) \le (\lambda_c,\, u_c - \psi)_\Gamma,
\end{aligned}
$$

and hence

$$
0 \le \overline{\lim}_{c\to\infty} |\nabla(u_c - u^*)|^2 \le (\hat\lambda,\, \hat u - \psi)_\Gamma \le 0,
$$

where we use that $(u_c - \hat u)\,|\Gamma \to 0$ strongly in $L^2(\Gamma)$. As a consequence $\lim_{c\to\infty} u_c \to u^*$ strongly in $H^1_{0,D}(\Omega)$ and $(\hat u,\, \hat\lambda)$ satisfies the complementarity system

$$
\hat u \le \psi, \quad \hat\lambda \ge 0, \quad \hat\lambda(\hat u - \psi) \quad \text{on}\, \Gamma. \tag{2.15}
$$

Taking the limit in

$$
(\nabla u_c,\, \nabla v)_\Omega = (f, v)_\Omega + (q, v)_\Gamma - (\lambda_\gamma, v), \quad \text{for all } v \in H^1_{0,D}(\Omega).
$$

implies that

$$
(\nabla \hat u,\, \nabla v)_\Omega = (f, v)_\Omega + (q, v)_\Gamma - (\hat\lambda, v), \quad \text{for all } v \in H^1_{0,D}(\Omega).
$$

Together with (2.15) this implies that $(\hat u, \hat\lambda)$ satisfies (2.2). Since the solution to (2.2) is unique we have $(\hat u, \hat\lambda) = (u^*, \lambda^*)$. □

Remark 2.1. *Let $\mathcal{A}_c = \{x \in \Gamma : u_c(x) = \psi(x)\}$ and $\mathcal{A} = \{x : u^*(x) = \psi(x)\}$. Proposition 2.2 and Theorem 2.1 imply that $\mathcal{A}_c$ is monotonically increasing and that*

$$
\mathcal{A}_c \to \mathcal{A}^* \quad as \quad c \to \infty.
$$

3 Semi-smooth Newton method for regularized problem.

This section is devoted to the discussion of an iterative algorithm for solving (2.10). Note that the direct application of a Newton algorithm is impeded by the fact that the max-operation is not differentiable. Alternatively we shall apply a semi-smooth Newton method to the mapping $F : L^2(\Gamma) \to L^2(\Gamma)$ defined by

$$F(\lambda) = \lambda - \max\big(0, \bar{\lambda} + c(u(\lambda)|\Gamma - \psi)\big), \tag{3.1}$$

where $u(\lambda)$ is the solution to (2.11), which we repeat for convenience, dropping the index c,

$$\begin{cases} -\Delta u = f \ \text{ in } \ \Omega, \\ \frac{\partial u}{\partial u} = q \ \text{ on } \ \Gamma_N, \quad u = 0 \ \text{ on } \ \Gamma_D \\ \frac{\partial u}{\partial u} = -\lambda \ \text{ on } \ \Gamma. \end{cases} \tag{3.2}$$

The solution to $F(\lambda) = 0$ provides the unique solution to (2.10). We now briefly recall those facts from semi-smooth Newton methods which are relevant for present the context.

Let X and Z be Banach spaces and let $F\colon D \subset X \to Z$ be a nonlinear mapping with open domain D.

Definition 3.1. *The mapping $F\colon D \subset X \to Z$ is called Newton differentiable on the open subset $U \subset D$ if there exists a family of generalized gradients $G\colon U \to \mathcal{L}(X, Z)$ such that*

$$\lim_{h\to 0} \frac{1}{\|h\|} \|F(x+h) - F(x) - G(x+h)h\| = 0, \tag{A}$$

for every $x \in U$.

Theorem 3.1. *Suppose that $x^* \in D$ is a solution to $F(x) = 0$ and that F is Newton–differentiable in an open neighborhood U containing x^* and that $\{\|G(x)^{-1}\|\colon x \in U\}$ is bounded. Then the Newton–iteration $x_{k+1} = x_k - G(x_k)^{-1}F(x_k)$ converges superlinearly to x^* provided that $\|x_0 - x^*\|$ is sufficiently small.*

We refer to mappings F which allow a Newton derivative on U in the sense of Definition 3.1 as Newton–differentiable.

Let us consider Newton–differentiability of the max–operator. For this purpose X denotes a function space of real–valued functions on a bounded domain $\omega \subset \mathbb{R}^n$ and $\max(0, y)$ is the pointwise max–operation. For $\delta > 0$ we introduce candidates for the generalized gradients of the form

$$G_m(y)(x) = \begin{cases} 1 & \text{if} \quad y(x) > 0 \\ 0 & \text{if} \quad y(x) < 0 \\ \delta & \text{if} \quad y(x) = 0, \end{cases} \tag{3.3}$$

where $y \in X$.

Proposition 3.1. *The mapping* $\max(0, \cdot) \colon L^q(\omega) \to L^p(\omega)$ *with* $1 \le p < q < \infty$ *is Newton differentiable on* $L^q(\omega)$ *and* G_m *is a generalized gradient.*

For the proofs of Theorem 3.1 and Proposition 3.1 we refer to [HIK]. Related results can be found in [U]. The following chain rule will be needed in the proof of Theorem 3.2 below . We utilize a third Banach space Y.

Proposition 3.2. *Let* $F_2 \colon Y \to X$ *be an affine mapping with* $F_2\, y = By + b$, $B \in \mathcal{L}(Y, X)$, $b \in X$, *and assume that* $F_1 \colon D \subset X \to Z$ *is Newton–differentiable on the open subset* $U \subset D$ *with generalized gradient* G. *If* $F_2^{-1}(U)$ *is nonempty, then* $F = F_1 \circ F_2$ *is Newton–differentiable on* $F_2^{-1}(U)$ *with generalized gradient given by* $G(By + b)B \in \mathcal{L}(Y, Z)$, *for* $y \in F_2^{-1}(U)$.

We are now prepared to address super-linear convergence of a semi-smooth Newton method applied to (3.1).

Theorem 3.2. *Let* (2.8) *hold and let* $\bar{\lambda} \in L^p(\Gamma)$ *for some* $p > 2$. *Then semi-smooth Newton-iteration applied to* F *given in* (3.1) *with generalized gradient for the max-operator given in* (3.3) *and* $\omega = \Gamma$, *converges locally superlinearly.*

Proof. We apply Proposition 3.2 with $Y = Z = L^2(\Gamma)$ and $X = L^P(\Gamma)$, where $p > 2$. Moreover B is given by $\lambda \to cu(\lambda)|\Gamma$ and $b = \bar{\lambda} - c\psi$. Since $H^{\frac{1}{2}}(\Gamma)$ embeds continuously into every $L^p(\Gamma), p \in [1, \infty)$, we have $b \in L^p(\Gamma)$.

Moreover $\lambda \to u(\lambda)|\Gamma$ is a continuous mapping from $L^2(\Gamma)$ to $H^{\frac{1}{2}}(\Gamma)$ and hence $B \in \mathcal{L}(L^2(\Gamma), L^p(\Gamma))$.

We still have to verify that the generalized gradients G_F of F are uniformly bounded in $\mathcal{L}(L^2(\Gamma))$ for λ in a neighborhood of λ^*. Here

$$G_F(\lambda)\,\delta\lambda = \delta\lambda - c\chi_{\mathcal{A}_\lambda}\delta u(\delta\lambda), \tag{3.4}$$

where $\chi_{\mathcal{A}_\lambda}$ is the characteristic function of the set

$$\mathcal{A}_\lambda = \{x \in \Gamma : \bar{\lambda} + c(u(\lambda) - \psi) > 0\},$$

with $u(\lambda)$ the solution to (3.2) and $\delta u = \delta u(\delta\lambda) = u'(\lambda)\delta\lambda$ the solution to

$$\begin{cases} -\Delta\,\delta u = 0 \ \text{ in } \ \Omega \\ \frac{\partial}{\partial n}\,\delta u = 0 \text{ on } \Gamma_N, \quad \delta u = 0 \text{ on } \Gamma_D \\ \frac{\partial}{\partial n}\,\delta u = -\delta\lambda \text{ on } \Gamma. \end{cases} \tag{3.5}$$

Let λ and g in $L^2(\Gamma)$ be arbitrary, and note that

$$G_F(\lambda)\,\delta\lambda = g$$

can equivalently be expressed as

$$\delta\lambda - c\chi_{\mathcal{A}_\lambda}\,\delta u(\delta\lambda) = g$$

or

$$\begin{cases} \delta\lambda\,\chi_{\mathcal{I}} = g\,\chi_{\mathcal{I}} \\ \delta\lambda\,\chi_{\mathcal{A}} - c\chi_{\mathcal{A}}\,\delta u(\delta\lambda\,\chi_{\mathcal{A}}) = g\,\chi_{\mathcal{A}} + c\,\chi_{\mathcal{A}}\,\delta u(\delta\lambda\,\chi_{\mathcal{I}}), \end{cases} \tag{3.6}$$

where the notation of the dependence of $\mathcal{A}$ and $\mathcal{I}$ on λ is dropped. The first equation in (3.6) determined $\delta\lambda$ uniquely on $\mathcal{I}$. The Lax Milgram lemma can be used to solve the second equation in $L^2(\mathcal{A})$. In fact, note that $|\nabla\,\delta u(\delta\lambda\,\chi_{\mathcal{A}})|^2_{L^2(\Omega)} = -\int_\Gamma \delta\lambda\,\chi_{\mathcal{A}}\,\delta u(\delta\lambda\,\chi_{\mathcal{A}})$. Therefore, taking the inner product of the second equation in (3.6) with $\delta\lambda\,\chi_{\mathcal{A}}$ we obtain for a constant K independent of $\delta\lambda$

$$\begin{aligned} |\delta\lambda\,\chi_{\mathcal{A}}|_{L^2(\Gamma)} \le |g|_{L^2(\Gamma)} + c|\delta u(\delta\lambda\,\chi_{\mathcal{I}})|_{L^2(\Gamma)} \le |g|_{L^2(\Gamma)} + cK|\delta\lambda\,\chi_{\mathcal{I}}|_{L^2(\Gamma)} \\ \le |g|_{L^2(\Gamma)} + cK|g\,\chi_{\mathcal{I}}|_{L^2(\Gamma)} = (1 + cK)|g|_{L^2(\Gamma)}. \end{aligned}$$

This implies that (3.6) admits a solution $\delta\lambda$ for any $g \in L^2(\Gamma)$ and $|\delta\lambda|_{L^2(\Gamma)} \leq (1 + cK)|g|_{L^2(\Gamma)}$. The claim now follows from Theorem 3.1.

$\square$

To express the Newton step

$$G_F(\lambda_k)\delta\lambda = -F(\lambda_k) \tag{3.7}$$

in an alternative way let

$$\mathcal{A}_k = \{x \in \Gamma : \bar{\lambda}(x) + c(u_k(x) - \psi(x)) > 0\}, \quad \mathcal{I}_k = \Gamma \backslash \mathcal{A}_k,$$

where $u_k = u(\lambda_k)$. Then (3.7) can be equivalently expressed as

$$\delta\lambda - cu'(\lambda_k)\,\delta\lambda\,\chi_{\mathcal{A}_k} = -\lambda_k + \max(0, \bar{\lambda} + c(u_k - \psi))$$

or

$$\lambda_{k+1} = (\bar{\lambda} + c(u_{k+1} - \psi))\chi_{\mathcal{A}_k}. \tag{3.8}$$

Thus $\lambda_{k+1},\ u_{k+1} = u(\lambda_{k+1})$ are the solution to

$$\begin{cases} -\Delta u_{k+1} = f \text{ in } \Omega \\ \frac{\partial}{\partial n} u_{k+1} = q, \quad \text{on } \Gamma_N, \quad u_{k+1} = 0 \text{ on } \Gamma_D \\ \lambda_{k+1} = -\frac{\partial}{\partial n} u_{k+1} = 0 \ \text{ on } \ \Gamma_{\mathcal{I}_k}, \\ \lambda_{k+1} = -\frac{\partial}{\partial n} u_{k+1} = \bar{\lambda} + c(u_{k+1} - \psi) \text{ on } \Gamma_{\mathcal{A}_k}. \end{cases} \tag{3.9}$$

The semi-smooth Newton algorithm can now be expressed as the following active set strategy with respect to the inequality $u \leq \psi$:

Primal Dual Active set algorithm

(i) Determine $\bar{u}$, $\bar{\lambda}$ according to (2.3) and (2.9), set $c > 0,\ k = 0$.

(ii) Set $u_0 = \bar{u}$.

(iii) Determine $\mathcal{A}_k,\ \mathcal{I}_k$.

(iv) Solve (3.9) for u_{k+1}. Set $\lambda_{k+1} = -\frac{\partial}{\partial n} u_{k+1}$ on Γ.

(v) Stop or set $k = k+1$ and go to (iii).

Clearly alternative initializations are possible. By Theorem 3.2 this algorithm converges superlinearly if the initialization is sufficiently close to the solution u_c of (2.10). The algorithm also converges globally.

Theorem 3.3. *Let* (2.8) *hold and* $c > 0$. *Then* $\lim_{k\to\infty}(y_k, \lambda_k) = (y_c, \lambda_c)$ *in* $H^1(\Omega) \times L^2(\Gamma)$ *as* $k \to \infty$.

Proof. On $\mathcal{I}_k$

$$-\frac{\partial}{\partial n}(u_{k+1} - u_k) = \begin{cases} 0 - 0 & \text{on } \mathcal{I}_k \cap \mathcal{I}_{k-1} \\ 0 - (\bar\lambda + c(u_k - \psi)) & \text{on } \mathcal{I}_k \cap \mathcal{A}_{k-1} \end{cases} \geq 0.$$

Similarly on $\mathcal{A}_k$

$$-\frac{\partial}{\partial n}(u_{k+1} - u_k) = \begin{cases} c(u_{k+1} - u_k) & \text{on } \mathcal{A}_k \cap \mathcal{A}_{k-1} \\ \bar\lambda + c(u_{k+1} - \psi) - 0 & \text{on } \mathcal{A}_k \cap \mathcal{I}_{k-1} \end{cases} \geq c(u_{k+1} - u_k).$$

Thus, it follows that

$$\begin{aligned} 0 &= -\int_\Omega \Delta(u_{k+1} - u_k)(u_{k+1} - u_k)^+ \, dx \\ &= \int_\Omega |\nabla(u_{k+1} - u_k)^+|^2 - \int_\Gamma \tfrac{\partial}{\partial n}(u_{k+1} - u_k)(u_{k+1} - u_k)^+ \, ds \\ &\geq \int_\Omega |\nabla(u_{k+1} - u_k)^+|^2 + \int_{\mathcal{A}_k} c|(u_{k+1} - u_k)^+|^2 \, ds. \end{aligned}$$

Consequently $u_{k+1} - u_k \leq 0$ a.e. in Ω. We can now proceed as in [IK] to verify the desired convergence.

In fact, as in Propositions 2.4 and 2.5 and of [IK] we show that $u_c \leq u_k$ and $0 \leq \lambda_{k+1} \leq \lambda_k$ for all k. Since u_k is the solution to

$$\min_{u \in H^1_{0,D}(\Omega)} \frac{1}{2}\int_\Omega |\nabla u|^2 - \int_{\Gamma_N} qu - \int_\Omega uf + \frac{1}{2c}\int_{\mathcal{A}_{k-1}} |\lambda_k|^2,$$

it follows that $\{u_k\}_{k=1}^\infty$ is bounded in $H^1(\Omega)$. Extracting subsequences and using Lebegue's bounded convergence theorem, the proof can now be completed as that of Theorem 2.1. in [IK]. □

4 Numerical tests

The feasibility of the proposed active set method was tested numerically by means a finite difference approximation on a uniform grid. The second order operator was discretized by a five point stencil and the Neumann boundary conditions were realized by a second order discretization.

The iteration wan be terminated by means of the criterion that two consecutive active sets coincide. In this case the exact solution of the discretized problem is found.

For several examples with smooth problem data, we made the following common observations.

- The number of iterations increases with c and with the number of grid points. However, the increase is very moderate.
- The active sets increase as c is increased.
- For the examples that we ran, the active set did not change any more for $c \geq 10^4$.
- Choosing $\bar{\lambda}$ different from (2.9) may lead to chattering of the iterates, higher iteration numbers and in any case, to unfeasible solutions. Chattering can possibly be eliminated by taking into consideration that the determination of the active sets involves manipulation with numerical zeroes. In [BHHK] a method was proposed in a related situation, which allows to cope with this difficulty. Here we took the point of view that using $\bar{\lambda}$ the situation did not arise.
- The angle between the obstacle and the solution at the points of contact can be very small. Consequently, the determination of the active set on the basis of logic statements involving ≥ 0 can be sensitive with respect to discretization errors.

Let us turn to a specific example next. We chose $q(s) = -7s(1-s)$, $f(x_1, x_2) = \cos(\frac{\pi}{2} + \pi x_1) + 1$, and $\psi(s) = 5s(1-s)(.5-x)\max(s, 1-s)$. The solution $\bar{u}$ to the initialization phase is depicted in Figure 1, the final solution, for $c = 10^4$ and mesh size $h = \frac{1}{n} = \frac{1}{128}$ in Figure 2.

c	*iter*	active set $\mathcal{A}_c$	$\max(u_c - \psi)$
1	2	$\{\}$	≤ 0
10	3	$(0, .008) \cup (.760, .820)$	$4.4 * 10^{-5}$
100	5	$(0, .027) \cup (.656, .863)$	$8.8 * 10^{-5}$
1000	6	$(0, .031) \cup (.664, .867)$	$1.9 * 10^{-5}$
10000	6	$(0, .031) \cup (.664, .867)$	$2.1 * 10^{-6}$
100000	6	$(0, .031) \cup (.664, .867)$	$2.1 * 10^{-7}$

Table 1: n=256, increasing c

n	*iter*	active set $\mathcal{A}_c$	$\max(u_c - \psi)$
16	2	$(0, .0625) \cup (.563, .875)$	$2.98 * 10^{-5}$
32	3	$(0, .0625) \cup (.531, .875)$	$1.6 * 10^{-5}$
64	4	$(0, .0625) \cup (.547, .891)$	$7.5 * 10^{-6}$
128	6	$(0, .0625) \cup (.586, .883)$	$3.8 * 10^{-6}$
256	6	$(0, .0625) \cup (.664, .867)$	$2.1 * 10^{-6}$

Table 2: c=10000, increasing n

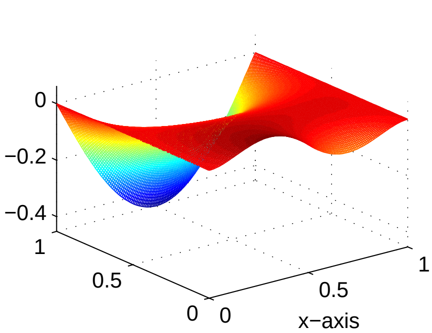

Figure 1

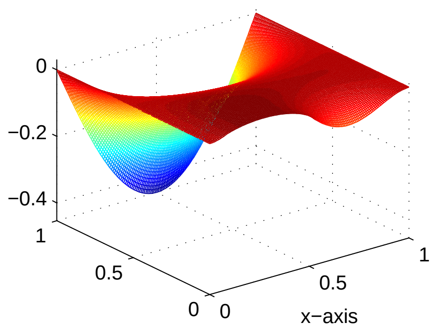

Figure 2

In Table 1 we present results for increasing values of c and fixed mesh size $h = \frac{1}{256}$. Here *iter* refers to the number of iterations that are required before two consecutive active sets $\mathcal{A}_k$ coincide. Further $\max(u_c - \psi)$ refers to the value of this expression along Γ. We note that consistent with Remark 2.1 the active sets are increasing as c increases.

In Table 2 the results for decreasing mesh size. As claimed earlier, the dependence of the iteration number on n is small. For this reason we do not propose to use specific techniques, such as path following methods for this class of problems to determine c. The second component of the active set determined on the basis $u \geq \psi$ is sensitive with respect to the meshsize.

In Figure 3 we present $u - \psi$ along the boundary Γ for four consecutive mesh sizes, exhibiting the two components of the active set.

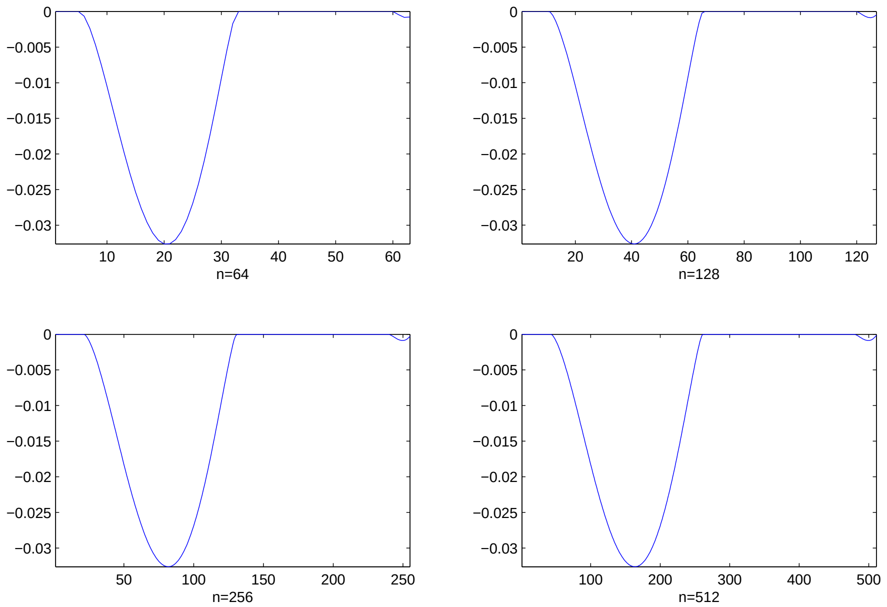

Figure 3

References

[BHHK] M. Bergounioux, M. Haddou, M. Hintermüller and K. Kunisch: A Comparison of a Moreau-Yosida Based Active Strategy and Interior Point Methods for Constrained Optimal Control Problems, SIAM J. on Optimization, 11(2000), 495–521.

[G] R. Glowinski: *Numerical Methods for Nonlinear Variational Problems*, Springer Verlag, New York, 1984.

[GLT] R. Glowinski, J.L. Lions and T. Tremolieres: *Analyse Numerique des Inequations Variationnelles, Vol. 1*, Dunod, Paris, 1976.

[G1] P. Grisvard: *Elliptic Problems in Nonsmooth Domains*, Pitman: Boston, 1985.

[G2] P. Grisvard: *Singularities in Boundary Value Problems*, Recherches en mathmatiques appliques, 22, Masson, Paris, 1992.

[HIK] M. Hintermüller, K. Ito and K. Kunisch: The primal–dual active set strategy as semi–smooth Newton method, SIAM Journal on Optimization, 13(2002), 865-888.

[HK] M. Hintermüller and K. Kunisch: Feasible and Non-Interior Path-Following in Constrained Minimization with Low Multiplier Regularity, SIAM J. Control and Optim., 45(2006), 1198-1221.

[IK] K. Ito and K. Kunisch: Semi–smooth Newton methods for variational inequalities of the first kind, Mathematical Modelling and Numerical Analysis, R.A.R.I.O, Mathematical Modelling and Numerical Analysis, 37(2002), 41-62.

[U] M. Ulbrich: Semismooth Newton methods for operator equations in function spaces, SIAM J. Optimization, 13(2002), 805-841.