Mathematical Modelling and Numerical Analysis
Modélisation Mathématique et Analyse Numérique

ESAIM: M2AN
M2AN, Vol. 37, N° 1, 2003, pp. 41–62
DOI: 10.1051/m2an:2003021

SEMI–SMOOTH NEWTON METHODS FOR VARIATIONAL INEQUALITIES OF THE FIRST KIND

Kazufumi Ito[1,*] and Karl Kunisch[2,†]

Abstract. Semi–smooth Newton methods are analyzed for a class of variational inequalities in infinite dimensions. It is shown that they are equivalent to certain active set strategies. Global and local super-linear convergence are proved. To overcome the phenomenon of finite speed of propagation of discretized problems a penalty version is used as the basis for a continuation procedure to speed up convergence. The choice of the penalty parameter can be made on the basis of an L^∞ estimate for the penalized solutions. Unilateral as well as bilateral problems are considered.

Mathematics Subject Classification. 49J40, 65K10.

Received: April 15, 2002. Revised: July 30, 2002.

1. Introduction

This paper is devoted to the convergence analysis of iterative algorithms for solving variational inequalities of the form

$$\begin{cases} \min \ \dfrac{1}{2}\, a(y,y) - (f,y) \\ y \in H_0^1(\Omega) \\ y \le \psi \ \text{ a.e. in } \Omega, \end{cases} \tag{1.1}$$

where $a(\cdot,\cdot)$ is a coercive bilinear form on $H_0^1(\Omega) \times H_0^1(\Omega)$, and $(\cdot,\cdot)$ denotes the inner product in $L^2(\Omega)$. The precise assumptions on the quantities appearing in (1.1) are given in Section 2. While iterative methods for solving finite dimensional discretization of (1.1) are extensively studied see *e.g.* [4, 8, 9] and the references therein, little attention has been paid to the infinite–dimensional counter–parts. Our contribution will focus on the convergence of the infinite dimensional algorithms. More precisely we shall analyze primal–dual active set algorithms or – as we shall argue – equivalently semi–smooth Newton algorithms. To briefly describe this class

Keywords and phrases. Semi-smooth Newton methods, contact problems, variational inequalities, bilateral constraints, super-linear convergence.

[1] Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University, USA.

* Supported in part by AFSOR under contract F-49620-95-1-0447.

[2] Institut für Mathematik, Universität Graz, Graz, Austria. e-mail: karl.kunisch@uni-graz.at

† Supported in part by the "Fonds zur Förderung der wissenschaftlichen Forschung under SFB 03, Optimierung und Kontrolle".

© EDP Sciences, SMAI 2003

of algorithms let y^* denote the solution to (1.1) and let λ^* be the associated Lagrange multiplier. As we shall recall in Section 2, the optimality system associated to (1.1) can be expressed as

$$\begin{cases} a(y^*, v) + (\lambda^*, v) = (f, v), \quad \text{for all} \quad v \in H_0^1(\Omega), \\ \lambda^* = \max(0, \lambda^* + c(y^* - \psi)), \end{cases} \tag{1.2}$$

for each $c > 0$, where max denotes the pointwise a.e. maximum operation. The second order augmented Lagrangian method in [1, 10] employs the primal–dual active set strategy based on the second equality in (1.2) and is given as the following iterative method: given a current pair (y_k, λ_k) of primal and dual variables, predict the active set $\mathcal{A}_{k+1}$ as

$$\mathcal{A}_{k+1} = \{x \colon \lambda_k(x) + c(y_k(x) - \psi(x)) > 0\} \cdot \tag{1.3}$$

We arrive at the following formal algorithm:

Algorithm.

(i) *Choose* $c > 0$, (y_o, λ_0), *set* $k = 0$.
(ii) *Determine* $\mathcal{A}_{k+1}$ *according to (1.2).*
(iii) *Solve for* $y_{k+1} = \arg\min\{\frac{1}{2}a(y, y) - (f, y) \colon y = \psi$ *on* $\mathcal{A}_{k+1}\}$.
(iv) *Let* λ_{k+1} *be the Lagrange multiplier associated to the constraint in* (iii) *with* $\lambda_{k+1} = 0$ *on* $\Omega \setminus \mathcal{A}_{k+1}$.
(v) *Set* $k = k + 1$ *and goto* (ii).

Let us observe that the optimality system for the variational problem in (iii) is given by

$$\begin{cases} a(y_{k+1}, v) + \langle \lambda_{k+1}, v \rangle_{H^{-1}, H_0^1} = (f, v) \quad \text{for all} \quad v \in H_0^1(\Omega), \\ y_{k+1} = \psi \quad \text{on} \quad \mathcal{A}_{k+1}, \ \lambda_{k+1} = 0 \quad \text{on} \quad \mathcal{I}_{k+1} = \Omega \setminus \mathcal{A}_{k+1}. \end{cases} \tag{1.4}$$

In particular, the Lagrange multiplier associated to the constraint $y = \psi$ on $\mathcal{A}_{k+1}$, is in general only a distribution in $H^{-1}(\Omega)$. This results from the fact that $\frac{\partial y_{k+1}}{\partial n}$ is not continuous across the boundaries between active and inactive sets. Rather λ_{k+1} contains jumps of magnitude $\frac{\partial y_{k+1}}{\partial n_{\mathrm{J}}^+} - \frac{\partial y_{k+1}}{\partial n_{\mathrm{J}}^-}$, where $n_{\mathrm{J}}^{\pm}$ stands for the normal directions to either side of the boundary between active and inactive set. These jumps are not present in the solution of the limit–problem (1.1), since under mild assumptions [12, 13] we have $y^* \in H^2(\Omega)$ and $\lambda^* \in L^2(\Omega)$. The fact that the Lagrange multipliers λ_{k+1} of the auxiliary problems in (iii) of the algorithm are not contained in the pivot space $L^2(\Omega)$ between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$ presents a serious difficulty, both from the point of view of numerical implementation and convergence analysis. In order to remedy this difficulty we consider a one-parameter family of regularized problems based on smoothing of the complementarity condition by

$$\lambda = \alpha \max(0, \lambda + c(y - \psi)), \quad 0 < \alpha < 1$$

which replaces the second equation in (1.2). The motivation for this regularization is that it is a relaxation of the second equation in (1.2). We analyze (i) the convergence of the active set strategy to the regularized problem, (ii) the monotone convergence property and L^∞ rate of convergence of solutions to the regularized problem to the original variational inequality and then (iii) develop and test a continuation method for the second order augmented Lagrangian method based on (i) and (ii).

The outline of the paper is as follows. In Section 2 we first introduce an equivalent but much more convenient form of the regularized problems and subsequently an iteration method based on the primal–dual active set strategy. We show that the method based on the active set strategy is equivalent to a semi-smooth Newton method [7]. Global as well as local super–linear convergence of the iteration method for the regularized problems is proven. The equivalence to the semi-smooth Newton is used to prove local super–linear convergence. Section 3 is devoted to the asymptotic analysis with respect to the regularization parameter. Monotone convergence

properties of the solutions of the regularized problems towards the solution of the original problem are proven and an L^∞-error estimate for this convergence is obtained. It is important to note that the L^∞-error estimate can be used as a guideline for the choice of the penalty in terms of the mesh-size. In Section 5 we present our numerical examples to demonstrate the structural results obtained in this paper. Moreover we demonstrate that the algorithm allows to determine the boundary of the active set within grid–size accuracy. We also show that regularization can be used to overcome an essential drawback of active set strategies applied to (1.1), *i.e.*, when the bilinear form a is discretized by finite differences (the five point stencil in dimension two) then changes from one iteration to the next occur along layers between active and inactive sets which have only the width of one mesh-size. For fine mesh-sizes this results in large iteration numbers. This difficulty can be overcome by multigrid methods, for example. Here we show that regularization techniques provide an alternative to deal with this shortcoming of active set strategies for (1.1). A regularized version of the above algorithm converges within a very few iteration due to its capability to change large sets of active indices to inactive ones and vice versa. We shall demonstrate that this property can advantageously be used in a continuation procedure with respect to the regularization parameter. The focus of our numerical test is not to compete with the most efficient implementations for this frequently tested class of obstacle problems, but rather to validate the structural results of the paper and to show the potential of a systematic use of regularization.

Our theoretical results provide a framework for an efficient second order iterative process for solving a regularized form of (1.2). It should also be noted that solving the regularized problem is equivalent to solving a single step of the first order augmented Lagrangian method, *e.g.*, see [11] and thus semi-smooth Newton methods should also improve the original implementation of the first order augmented Lagrangian method reported in [11]. This can be the focus of future investigations.

Beyond the motivation of overcoming the difficulty due to lack of regularity of the Lagrange multiplier our interest in analyzing primal–dual active set strategies for (1.1) also stems from our desire to investigate these algorithms separately for classes of problems which differ with respect to the regularity properties of the Lagrange multipliers. The abstract results are contained in [10]. In [3] we considered optimal control problems with control constraints. In this case the Lagrange multipliers of the original problem as well as those arising in the auxiliary problems of the primal–dual active set algorithm are in $L^2(\Omega)$ or $L^2(\partial\Omega)$, depending on whether distributed or boundary control problems are considered. For such problems large sets of active and inactive indices are moved from one iteration to the next. In [7] we established the strong relationship of these methods with superlinearly convergent semi–smooth Newton methods. For variational inequalities of the form (1.1) the Lagrange multipliers of the limit problem are L^2 but those of the auxiliary problems are not. Finally, for state constrained optimal control problems as well as for control of variational inequalities the Lagrange multipliers of the limit–problems themselves are not L^2 smooth but are in general only measures. Numerical results for these classes of problems are contained in [2, 11]. Convergence results for the latter are only available in the case of discretized state constrained optimal control problems.

We briefly summarize those facts on semi–smooth Newton methods which are relevant for our analysis in Section 2. 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 1.1. The mapping $F\colon D \subset X \to Z$ is called generalized–differentiable on the open subset $U \subset D$ if there exists a family of generalized derivatives $G\colon U \to \mathcal{L}(X, Z)$ such that

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

for every $x \in U$.

We shall refer to mappings F which allow a generalized derivative on U in the sense of Definition 1.1 as Newton–differentiable.

Theorem 1.1. *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.*

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

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

where $y \in X$.

Proposition 1.1. *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* $G_{m,\delta}$ *is a generalized derivative.*

For the proofs of Theorem 1.1 and Proposition 1.1 we refer to [7]. Related results can be found in [14]. The following chain rule will be utilized in Section 2. We utilize a third Banach space Y.

Proposition 1.2. *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 derivative* 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 derivative given by* $G(By + b)B \in \mathcal{L}(Y, Z)$, *for* $y \in F_2^{-1}(U)$.

Proof. By assumption $F_2^{-1}(U)$ is nonempty and due to continuity of F_2 the set $F_2^{-1}(U)$ is open. Note that $G(By + b)B \in \mathcal{L}(Y, Z)$ for each $y \in F_2^{-1}(U)$ since $G(x) \in \mathcal{L}(X, Z)$ for each $x \in U$. For $y \in F_2^{-1}(U)$ we find

$$\lim_{\substack{h \to 0 \\ h \in Y}} \frac{1}{\|h\|} \|(F_1 \circ F_2)(y + h) - (F_1 \circ F_2)(y) - G(F_2(y + h))Bh\| =$$

$$\lim_{\substack{h \to 0 \\ h \in Y}} \frac{1}{\|Bh\|} \|F_1(By + b + Bh) - F_1(By + b) - G(By + b + Bh)Bh\| \frac{\|Bh\|}{\|h\|} = 0,$$

and hence the claim follows.

2. Global and local convergence of the iterative method for the regularized problems

We consider

$$\begin{cases} \min \ \dfrac{1}{2}\, a(y, y) - (f, y) \\ y \in H_0^1(\Omega) \\ y \le \psi \ \text{ a.e. in } \Omega, \end{cases} \tag{2.1}$$

where $a\,(\cdot, \cdot)$ is a bilinear form on $H_0^1(\Omega) \times H_0^1(\Omega)$ satisfying

$$a(v, v) \ge \nu |v|_{H_0^1}^2, \quad a(w, z) \le \mu |w|_{H^1} |z|_{H^1}, \tag{2.2}$$

for some $\nu > 0$ and $\mu > 0$ independently of $v \in H_0^1(\Omega)$ and $w, z \in H^1(\Omega)$. Further it is assumed that $f \in L^2(\Omega)$, $\psi \in H^1(\Omega)$ with $\psi|_{\partial\Omega} \ge 0$. Throughout Ω is a bounded domain in R^n with Lipschitzian boundary $\partial\Omega$. Since $\psi \in H^1(\Omega)$ the trace $\psi|_{\partial\Omega}$ is well-defined. The assumption $\psi|_{\partial\Omega} \ge 0$ implies that the set of admissible functions y for (2.1) is nonempty. We shall also require that a satisfies the weak maximum principle, *i.e.* for all $v \in H_0^1(\Omega)$

$$a(v, v^+) \le 0 \ \text{ implies } \ v^+ = 0, \tag{2.3}$$

where $v^+ = \max(0, v)$. We set $K = \{v \in H_0^1(\Omega) \colon v \geq 0 \text{ a.e.}\}$. It will further be convenient to introduce the representation operator

$$A \colon H_0^1(\Omega) \to H^{-1}(\Omega)$$

associated to $a(\cdot,\cdot)$. Utilizing (2.2) it is well-known [12] that (2.1) admits a unique solution $y^* \in H_0^1(\Omega)$ and an associated Lagrange multiplier $\lambda^* \in H^{-1}(\Omega)$. Under well-known additional regularity assumptions which we recall in Remark 2.3 below $\lambda^* \in L^2(\Omega)$ and the following optimality system characterizes y^*:

$$\begin{cases} a(y^*, v) + (\lambda^*, v) = (f, v), \quad \text{for all } \ v \in H_0^1(\Omega) \\ (\lambda^*, y^* - \psi) = 0, \ y^* \leq \psi, \ (\lambda^*, v) \geq 0 \ \text{ for all } \ v \in K. \end{cases} \tag{2.4}$$

By inspection (2.4) can equivalently be expressed as

$$\begin{cases} a(y^*, v) + (\lambda^*, v) = (f, v) \ \text{ for all } \ v \in H_0^1(\Omega) \\ \lambda^* = \max(0, \lambda^* + c(y^* - \psi)), \end{cases} \tag{2.5}$$

for arbitrary $c > 0$. (More precisely, (2.4) implies (2.5) for every c, and (2.5) for some $c > 0$ implies (2.4)). Next we turn to the regularization of the max–function in (2.5). We have motivated the necessity for regularization for the primal–dual active set method by the abstract algorithm in Section 1. Concerning the semi–smooth Newton approach we have from Proposition 1.1 that the max operation is Newton differentiable from $L^p(\Omega)$ to $L^2(\Omega)$ if $p > 2$. If we were to consider both y and λ as independent variables in a semi–smooth Newton approach to (2.5), then we can expect to obtain the necessary smoothing for the y component due to the first equation in (2.5) but we lack the smoothing property with respect to λ.

In our first attempt to regularize the max–function in (2.5) we are tempted to use the well–known smoothing

$$\max_\sigma(x) = \begin{cases} 0 & \text{for} \quad x < -\dfrac{\sigma}{2} \\ \dfrac{1}{2\sigma}\left(x + \dfrac{\sigma}{2}\right)^2 & \text{for} \quad -\dfrac{\sigma}{2} \leq x \leq \dfrac{\sigma}{2} \\ x & \text{for} \quad x > \dfrac{\sigma}{2}, \end{cases}$$

with $\sigma > 0$, see *e.g.* [1]. After a short computation we obtain an explicit expression for $\lambda = \lambda_\sigma(z)$ satisfying $\lambda = \max_\sigma(0, \lambda + c\,z)$ as

$$\lambda_\sigma(z) \begin{cases} = 0 & \text{for} \quad \lambda + c\,z < -\dfrac{\sigma}{2} \\ = \dfrac{\sigma}{2} - c\,z - \sqrt{-2c\sigma z} & \text{for} \quad -\dfrac{\sigma}{2} \leq \lambda + c\,z \leq \dfrac{\sigma}{2} \\ \in \left[\dfrac{\sigma}{2}, \infty\right) & \text{for} \quad \lambda + c\,z > \dfrac{\sigma}{2}. \end{cases}$$

Thus we obtain an equation $Ay + \lambda_\sigma(y - \psi) = f$ for $y \in H_0^1(\Omega)$, where λ_σ is a multi–valued function defined above. This smoothing has some nice properties but it is much less convenient than penalty–type smoothing that we turn to next.

As stated in introduction we shall use

$$\lambda = \alpha \ \max(0, \lambda + c(y^* - \psi)), \quad 0 < \alpha < 1 \tag{2.6}$$

to regularize the second equation in (2.5). This is equivalent to

$$\lambda = \max\left(0, \bar{\lambda} + \gamma(y - \psi)\right), \ \gamma \in (0, \infty), \tag{2.7}$$

where $\bar{\lambda} \in L^2(\Omega)$, if we set $\bar{\lambda} = 0$ and $\gamma = c\alpha/(1-\alpha)$. Note that $\gamma \to \infty^+$ as $\alpha \to 1^-$. This type of regularization will allow us to prove global monotone convergence of the primal–dual active set method. The introduction of $\bar{\lambda}$ in (2.7), which does not appear in the original regularization, was motivated by augmented Lagrangians, [10,11]. We shall see in Section 3 that depending on its choice the feasibility of the approximations can be controlled. Note that if $\bar{\lambda} = 0$ on $\{x\colon y(x) \geq \psi(x)\}$, then (2.7) can be regarded as a penalty–type formulation of the complementarity condition

$$y - \psi \leq 0, \ \lambda \geq 0, \ (y - \psi, \lambda) = 0,$$

as $\gamma \to \infty$. In the remainder of this section $\gamma > 0$ is a fixed constant and we consider an active set strategy or alternatively a semi–smooth Newton method to solve

$$\begin{cases} a(y, v) + (\lambda, v) = (f, v) \ \text{ for all } \ v \in H_0^1(\Omega) \\ \lambda = \ \max\left(0, \bar{\lambda} + \gamma(y - \psi)\right). \end{cases} \tag{2.8}$$

Monotone operator theory provides the existence of a unique solution $(y_\gamma, \lambda_\gamma) \in H_0^1(\Omega) \times L^2(\Omega)$. An independent existence proof will follow from the results of this section.

We turn to the description of the algorithm.

Primal–dual active set (PDAS) algorithm

(i) Choose y_0 , set $k = 0$.
(ii) Set $\mathcal{A}_{k+1} = \{x\colon (\bar{\lambda} + \gamma(y_k - \psi))(x) > 0\}$, $\mathcal{I}_{k+1} = \Omega \setminus \mathcal{A}_{k+1}$.
(iii) Solve for $y_{k+1} \in H_0^1(\Omega)$:

$$a(y, v) + (\bar{\lambda} + \gamma(y - \psi), \chi_{\mathcal{A}_{k+1}} v) = (f, v) \ \text{ for all } \ v \in H_0^1(\Omega). \tag{2.9}$$

(iv) Set

$$\lambda_{k+1} = \begin{cases} 0 & \text{on} \quad \mathcal{I}_{k+1} \\ \bar{\lambda} + \gamma(y_{k+1} - \psi) & \text{on} \quad \mathcal{A}_{k+1}. \end{cases}$$

(v) Stop or $k = k + 1$, goto (ii).

Remark 2.1. Here we establish the relationship between the above algorithm and a semi–smooth Newton method applied to (2.8). Recall the definition $A\colon H_0^1(\Omega) \to H^{-1}(\Omega)$ and introduce the nonlinear mapping $F\colon H_0^1(\Omega) \times L^2(\Omega) \to H^{-1}(\Omega) \times L^2(\Omega)$, by

$$F(y, \lambda) = \begin{pmatrix} A\,y + \lambda - f \\ \lambda - \max(0, \bar{\lambda} + \gamma(y - \psi)) \end{pmatrix}.$$

A generalized derivative G of F in the sense of Definition 1.1 and Proposition 1.1 with $\delta = 0$ is given by

$$G(y_k, \lambda_k)h = \begin{pmatrix} A\,h_1 + h_2 \\ h_2 - \gamma\chi_{\mathcal{A}_{k+1}} h_1 \end{pmatrix}.$$

where $h = (h_1, h_2) \in H_0^1(\Omega) \times L^2(\Omega)$.

The resulting semi–smooth Newton–update is thus given by

$$\begin{cases} A\,\delta y + \delta\lambda = -Ay_k - \lambda_k + f \\ \delta\lambda = -\lambda_k \ \text{ on } \ \mathcal{I}_{k+1}, \\ \delta\lambda - \gamma\delta y = -\lambda_k + \bar{\lambda} + \gamma(y_k - \psi) \ \text{ on } \ \mathcal{A}_{k+1} \end{cases} \tag{2.10}$$

where $\delta y = y_{k+1} - y_k$ and $\delta\lambda = \lambda_{k+1} - \lambda_k$, and coincides with step (iii)–(iv) of the primal–dual active set algorithm.

Remark 2.2. The semi–smooth Newton can be applied to (2.6) without reformulation as (2.7). Based on (2.6) it coincides with the one we specified above, with $\bar{\lambda} = 0$, except for the initialization phase, where now y_0 and λ_0 must be prescribed. In case of (2.6) the active set in step (ii) of the algorithm would be set $\tilde{\mathcal{A}}_{k+1} = \{x\colon (\lambda_k + \gamma(y_k - \psi))(x) > 0\}$ and the update on the basis of (2.6) for λ_{k+1} coincides with the one of step (iv) in the algorithm. Note that sgn $(\lambda_k + \gamma(y_k - \psi))(x) =$ sgn $(y_k - \psi)(x)$ for all $x \in \Omega$, and $k \geq 1$, and hence $\tilde{\mathcal{A}}_{k+1} = \mathcal{A}_{k+1}$ for all $k \geq 2$. A similar remark applies in case $\bar{\lambda} \neq 0$.

Properties of the semi–smooth Newton algorithm or equivalently the PDAS are analyzed next.

Proposition 2.1. *If $\mathcal{A}_{k+1} = \mathcal{A}_k$ for $k \geq 1$, then (y_k, λ_k) is the solution to (2.9).*

Proof. Since for given $\mathcal{A}_{k+1}$ the solution to (2.9) is unique it follows from $\mathcal{A}_k = \mathcal{A}_{k+1}$ that $y_k = y_{k+1}$ and consequently $\lambda_{k+1} = \lambda_k$.

Proposition 2.2. *The sequence $\{y_k\}_{k=1}^{\infty}$ is monotonically decreasing, i.e. $y_{k+1} \leq y_k$, a.e. on Ω for all $k \geq 1$.*

Proof. Let $\delta y = y_{k+1} - y_k$ for $k \geq 1$ and observe that

$$a(\delta y, \delta y^+) + (\lambda_{k+1} - \lambda_k, \delta y^+) = 0. \tag{2.11}$$

We have

$$\lambda_{k+1}(x) - \lambda_k(x) \begin{cases} = 0 & \text{for} \quad x \in \mathcal{I}_{k+1} \cap \mathcal{I}_k, \\ = \gamma(y_{k+1}(x) - y_k(x)) & \text{for} \quad x \in \mathcal{A}_{k+1} \cap \mathcal{A}_k, \\ = -\bar{\lambda} - \gamma(y_k - \psi)(x) \geq 0 & \text{for} \quad x \in \mathcal{I}_{k+1} \cap \mathcal{A}_k, \\ > \gamma(y_{k+1} - y_k)(x) & \text{for} \quad x \in \mathcal{A}_{k+1} \cap \mathcal{I}_k. \end{cases}$$

It follows that $(\lambda_{k+1} - \lambda_k, \delta y^+) \geq 0$ and by (2.11)

$$a(\delta y, \delta y^+) \leq 0.$$

Consequently $\delta y^+ = 0$ by (2.3) and $y_{k+1} \leq y_k$ follows. □

Proposition 2.3. *For every $k \geq 1$ we have $\mathcal{I}_k \subset \mathcal{I}_{k+1}$.*

Proof. If not, then there exists a set $S \subset \Omega$ of positive measure and $x \in \mathcal{I}_k \cap \mathcal{A}_{k+1}$ for every $x \in S$. It follows that $\left(\bar{\lambda} + (y_{k-1} - \psi)\right)(x) \leq 0$ and by Proposition 2.2 $\left(\bar{\lambda} + (y_k - \psi)\right)(x) \leq 0$. On the other hand $x \in \mathcal{A}_{k+1}$, and hence $\left(\bar{\lambda} + (y_k - \psi)(x)\right) > 0$. This gives the desired contradiction. □

Proposition 2.4. *For every $k \geq 1$ we have $y_\gamma \leq y_k$.*

Proof. We consider the sign of $\lambda_k - \lambda_\gamma$. Let $\mathcal{A}_\gamma = \{x\colon \left(\bar\lambda + \gamma(y_\gamma - \psi)\right)(x) > 0\}$, and $\mathcal{I}_\gamma = \Omega \setminus A_\gamma$. For $x \in \mathcal{I}_\gamma \cap \mathcal{I}_k$ we have $(\lambda_k - \lambda_\gamma)(x) = 0$, and for $x \in \mathcal{A}_\gamma \cap \mathcal{A}_k$ we find $(\lambda_k - \lambda_\gamma)(x) = \gamma(y_k - y_\gamma)(x)$. If $x \in \mathcal{I}_\gamma \cap \mathcal{A}_k$ then $(\lambda_k - \lambda_\gamma)(x) = \left(\bar\lambda + \gamma(y_k - \psi)\right)(x) \le \gamma(y_k - y_\gamma)$. Finally, if $x \in \mathcal{A}_\gamma \cap \mathcal{I}_k$ then $(\lambda_k - \lambda_\gamma)(x) = -\left(\bar\lambda + \gamma(y_\gamma - \psi)\right)(x) \le 0$. We find

$$a(y_\gamma - y_k, (y_\gamma - y_k)^+) = -(\lambda_\gamma - \lambda_k, (y_\gamma - y_k)^+) \le 0.$$

By (2.2) it follows that $(y_\gamma - y_k)^+ = 0$ and hence $y_\gamma \le y_k$. □

Proposition 2.5. *For every $k \ge 1$ we have $0 \le \lambda_{k+1} \le \lambda_k$.*

Proof. The claim follows from Propositions 2.2 and 2.3. □

Note that Propositions 2.2–2.5 hold for $k \ge 1$ and are in general not valid for the initialization step with $k = 0$.

Theorem 2.1. *For every $\gamma > 0$ we have $\lim_{k\to\infty}(y_k, \lambda_k) = (y_\gamma, \lambda_\gamma)$ in $H_0^1(\Omega) \times L^2(\Omega)$.*

Proof. The sequences $\{y_k\}_{k=1}^\infty$ and $\{\lambda_k\}_{k=1}^\infty$ are decreasing pointwise almost everywhere and are uniformly bounded by $L^2(\Omega)$ functions. By (2.2) and (2.9) moreover, $\{y_k\}_{k=1}^\infty$ is bounded in $H_0^1(\Omega)$. Hence there exist $\hat y \in H_0^1(\Omega)$ and $\hat\lambda \in L^2(\Omega)$ such that a subsequence of y_k converges weakly in $H_0^1(\Omega)$ to $\hat y$, and $\lim_{k\to\infty} y_k = \hat y$ a.e. and $\lim_{k\to\infty} \lambda_k = \hat\lambda$ a.e.. Since $\mathcal{I}_k \subset \mathcal{I}_{k+1}$ and $\lambda_k = 0$ on $\mathcal{I}_k$ it follows that $\hat\lambda = 0$ on $\mathcal{I} = \bigcup_{k=1}^\infty \mathcal{I}_k$ and $\hat\lambda = \bar\lambda + \gamma(\hat y - \psi)$ on $\mathcal{A} = \bigcap_{k=1}^\infty \mathcal{A}_k$. Moreover, if $x \in \mathcal{I}$ then $(\bar\lambda + (\hat y - \psi))(x) \le 0$, and for $x \in \mathcal{A}$ we have $(\bar\lambda + \gamma(y_k - \psi))(x) > 0$ for all k and hence $(\bar\lambda + (\hat y - \psi))(x) \ge 0$. Consequently $\hat\lambda = \max(0, \bar\lambda + \gamma(\hat y - \psi))$. By Lebesgue's bounded convergence theorem $\lim_{k\to\infty} \lambda_k = \hat\lambda$ in $L^2(\Omega)$. Taking the limit in

$$a(y_k, v) + (\lambda_k, v) = (f, v),$$

we arrive at

$$\begin{aligned} &a(\bar y, v) + (\hat\lambda, v) = (f, v) \quad \text{for all} \quad v \in H_0^1(\Omega) \\ &\hat\lambda = \max(0, \bar\lambda + \gamma(\hat y - \psi)). \end{aligned}$$

Since the solution to this system is unique we have $\left(\hat y, \hat\lambda\right) = (y_\gamma, \lambda_\gamma)$. Finally, setting $v = y_k$ in (2.6) and using (2.2) we find $|y_k|_{H_0^1} \to |y_\gamma|_{H_0^1}$. Together with weak convergence of y_k to y_γ in H_0^1 this implies $\lim_{k\to\infty} y_k = y_\gamma$ in $H_0^1(\Omega)$. □

Remark 2.3. Under additional regularity assumptions the above result can be strengthened. We shall repeatedly refer to these assumptions which we now summarize. The bilinear form has the form

$$a(v, w) = \int_\Omega [a_{ij}\partial_{x_i} v \partial_{x_j} w + dw]\mathrm{d}x,$$

for $v, w \in H^1(\Omega)$, where we use the summation convention, the leading differential operator is uniformly elliptic and $a_{ij} \in C^1(\bar\Omega)$, $d \in L^\infty(\Omega)$, $d \ge 0$. Moreover $\psi \in H^2(\Omega)$, $\partial\Omega$ is $C^{1,1}$ or Ω is a polyhedron.

Under these requirements the representation operator A is a homeomorphism from $H^2(\Omega) \cap H_0^1(\Omega)$ to $L^2(\Omega)$. The solution to (2.1) satisfies $(y^*, \lambda^*) \in (H^2(\Omega) \cap H_0^1(\Omega)) \times L^2(\Omega)$, see *e.g.* [11–13], or as corollary to the results of Section 3. Moreover $\lim_{k\to\infty} y_k = y_\gamma$ in $H_0^1(\Omega) \cap H^2(\Omega)$ in the statement of Theorem 2.1. □

Theorem 2.1 guarantees global convergence of the semi–smooth Newton method, *i.e.* the algorithm converges from any initial condition. Next we establish that once the iterates are sufficiently close to the solution, then the convergence is superlinear.

For this purpose we introduce the mapping $F\colon L^2(\Omega) \to L^2(\Omega)$ by

$$F(\lambda) = \lambda - \max\left(0, \bar{\lambda} + \gamma\left(A^{-1}(f-\lambda) - \psi\right)\right). \tag{2.12}$$

Note that $F(\lambda) = 0$ is equivalent to system (2.8). We consider the following reduced algorithm in the variable λ. It arises from applying the quasi–Newton method to $F(\lambda) = 0$. It turns out that the reduced algorithm is equivalent to the primal-dual active set algorithm.

Reduced algorithm

(i) Choose $y_0 \in H_0^1(\Omega)$, set $\lambda_0 = f - Ay_0$ and $k = 0$.
(ii) Set $\mathcal{A}_{k+1} = \left\{x\colon \left[\bar{\lambda} + \gamma A^{-1}(f-\lambda_k) - \gamma\,\psi\right](x) > 0\right\}$, $\mathcal{I}_{k+1} = \Omega \setminus \mathcal{A}_{k+1}$.
(iii) Set $\delta\lambda = \lambda_k$ on $\mathcal{I}_{k+1}$, and solve for $\delta\lambda \in L^2(\Omega)$

$$\left(\delta\lambda + \gamma A^{-1}(\delta\lambda)\right)(x) = \left[-\lambda_k + \bar{\lambda} - \gamma\,\psi + \gamma A^{-1}(f-\lambda_k)\right](x), \quad x \in \mathcal{A}_{k+1}.$$

(iv) Set $\lambda_{k+1} = \lambda_k + \delta\lambda$ and goto (ii).

In fact (iii)–(iv) of the reduced algorithm is equivalent to

$$\left(\lambda_{k+1} - \bar{\lambda} + \gamma\,\psi\right)(x) = \gamma\left(A^{-1}(f-\lambda_{k+1})\right)(x) \ \text{ for } \ x \in \mathcal{A}_{k+1},\ \lambda_{k+1} = 0 \ \text{ in } \ \mathcal{I}_{k+1}$$

and thus it is equivalent to (iii)–(iv) of the primal–dual active set algorithm with $y_{k+1} = A^{-1}(f - \lambda_{k+1})$. Since the initializations for both algorithms are the same the two algorithms give identical iterates. Note that while λ_0 may only be in $H^{-1}(\Omega)$, the iterates satisfy $\lambda_k \in L^2(\Omega)$ for $k \geq 1$.

Theorem 2.2. *If $\lambda_0 \in L^2(\Omega)$ and $|\lambda_0 - \lambda_\gamma|_{L^2(\Omega)}$ is sufficiently small then $(y_k, \lambda_k) \to (y_\gamma, \lambda_\gamma)$ superlinearly in $H_0^1(\Omega) \times L^2(\Omega)$.*

Proof. First we show superlinear convergence of λ_k to λ_γ by applying Theorem 1.1 to F defined in (2.12). Let $q = \frac{1}{2} - \frac{1}{n}$, if $n \geq 3$ and $q \in (2, \infty)$ if $n = 2$. Then $H^1(\Omega)$ is continuously injected into $L^q(\Omega)$, and $q > 2$ for each n. From Propositions 1.1 and 1.2 it follows that F is Newton–differentiable on $L^2(\Omega)$. For this purpose we set $B = \gamma A^{-1}$ and $b = \bar{\lambda} + \gamma\left(A^{-1}f - \psi\right)$ in Proposition 1.2. To apply Theorem 1.1 it remains to verify that the generalized derivatives $G(\lambda) \in \mathcal{L}(L^2(\Omega))$ of F have uniformly bounded inverses. We define

$$\mathcal{A} = \left\{x\colon \left[\bar{\lambda} + \gamma\left(A^{-1}(f-\lambda) - \psi\right)\right](x) > 0\right\}, \quad \mathcal{I} = \Omega \setminus \mathcal{A}.$$

Further let $E_{\mathcal{A}}\colon L^2(\mathcal{A}) \to L^2(\Omega)$ and $E_{\mathcal{I}}\colon L^2(\mathcal{I}) \to L^2(\Omega)$ denote the extension – by – zero operators from $\mathcal{A}$ and $\mathcal{I}$ to Ω, respectively. Their adjoints $E_{\mathcal{A}}^*\colon L^2(\Omega) \to L^2(\mathcal{A})$ and $E_{\mathcal{I}}^*\colon L^2(\Omega) \to L^2(\mathcal{I})$ are restriction operators. The mapping $(E_{\mathcal{A}}^*, E_{\mathcal{I}}^*)\colon L^2(\Omega) \to L^2(\mathcal{A}) \times L^2(\mathcal{I})$ determines an isometric isomorphism and every $\lambda \in L^2(\Omega)$ can uniquely be expressed as $(E_{\mathcal{A}}^*\lambda, E_{\mathcal{I}}^*\lambda)$. A generalized derivative of F in the sense of Definition 1.1 is obtained by setting $\delta = 0$ in the definition $G_{m,\delta}$ for generalized derivatives of the max–operation. We obtain

$$G(\lambda) = I + \gamma\chi_{\mathcal{A}}A^{-1}.$$

This operator can equivalently be expressed as

$$G(\lambda) = \begin{pmatrix} I_{\mathcal{A}} & 0 \\ 0 & I_{\mathcal{I}} \end{pmatrix} + \gamma \begin{pmatrix} E_{\mathcal{A}}^* A^{-1} E_{\mathcal{A}} & E_{\mathcal{A}}^* A^{-1} E_{\mathcal{I}} \\ 0 & 0 \end{pmatrix},$$

where $I_{\mathcal{A}}$ and $I_{\mathcal{I}}$ denote the identity operators on $L^2(\mathcal{A})$ and $L^2(\mathcal{I})$. Let $(g_{\mathcal{A}}, g_{\mathcal{I}}) \in L^2(\mathcal{A}) \times L^2(\mathcal{I})$ be arbitrary and consider the equation

$$G(\lambda)((\delta\lambda)_{\mathcal{A}}, (\delta\lambda)_{\mathcal{I}}) = (g_{\mathcal{A}}, g_{\mathcal{I}}). \tag{2.13}$$

Then necessarily $(\delta\lambda)_{\mathcal{I}} = g_{\mathcal{I}}$ and (2.13) is equivalent to

$$(\delta\lambda)_{\mathcal{A}} + \gamma E_{\mathcal{A}}^* A^{-1} E_{\mathcal{A}} (\delta\lambda)_{\mathcal{A}} = g_{\mathcal{A}} - \gamma E_{\mathcal{A}}^* A^{-1} E_{\mathcal{I}} g_{\mathcal{I}}. \tag{2.14}$$

The Lax–Milgram theorem and positivity of A^{-1} imply the existence of a unique solution $(\delta\lambda)_{\mathcal{A}}$ to (2.14) and consequently (2.13) has a unique solution for every $(g_{\mathcal{A}}, g_{\mathcal{I}})$ and every λ. Moreover these solutions are uniformly bounded w.r.t. $\lambda \in L^2$. This follows from $(\delta\lambda)_{\mathcal{I}} = g_{\mathcal{I}}$ and

$$|\delta\lambda_{\mathcal{A}}|_{L^2(\mathcal{A})} \le |g_{\mathcal{A}}|_{L^2(\Omega)} + \gamma \|A^{-1}\|_{\mathcal{L}(L^2(\Omega))}\, |g_{\mathcal{I}}|_{L^2(\mathcal{I})}.$$

This proves superlinear convergence $\lambda_k \to \lambda_\gamma$ in $L^2(\Omega)$. Superlinear convergence of y_k to y_γ in $H_0^1(\Omega)$ follows from $Ay_k + \lambda_k = f$ and the fact that $A : H_0^1(\Omega) \to H^{-1}(\Omega)$ is a homeomorphism. □

If the problem data are sufficiently regular as specified in Remark 2.3 such that $A\colon H_0^1(\Omega) \cap H^2(\Omega) \to L^2(\Omega)$ is a homeomorphism, then $y_k \to y_\gamma$ in $H^2(\Omega)$ under the assumptions of Theorem 2.2.

3. Convergence of regularized problems

First we establish a general convergence result with respect to the penalty parameter γ. For related results we refer to [6], for example.

Theorem 3.1. *The solutions $(y_\gamma, \lambda_\gamma)$ to the regularized problem (2.8) converge to (y^*, λ^*) in the sense that $y_\gamma \to y^*$ strongly in $H_0^1(\Omega)$ and $\lambda_\gamma \rightharpoonup \lambda^*$ weakly in $H^{-1}(\Omega)$ as $\gamma \to \infty$.*

Proof. From (2.4) and (2.8) we have for every $\gamma > 0$

$$a(y_\gamma, y_\gamma - y^*) + (\lambda_\gamma, y_\gamma - y^*) = (f, y_\gamma - y^*)$$

where $\lambda_\gamma = \max\big(0, \bar\lambda + \gamma(y_\gamma - \psi)\big)$. Since $\lambda_\gamma \ge 0$ and $\psi - y^* \ge 0$ we have

$$(\lambda_\gamma, y_\gamma - y^*) = \left(\lambda_\gamma, \frac{\bar\lambda}{\gamma} + y_\gamma - \psi + \psi - y^* - \frac{\bar\lambda}{\gamma}\right) \ge \frac{1}{\gamma}\left(\lambda_\gamma, \bar\lambda + \gamma(y_\gamma - \psi)\right) - \frac{1}{\gamma}\left(\lambda_\gamma, \bar\lambda\right),$$

and hence

$$(\lambda_\gamma, y_\gamma - y^*) \ge \frac{1}{\gamma}|\lambda_\gamma|_{L^2}^2 - \frac{1}{\gamma}(\lambda_\gamma, \bar\lambda). \tag{3.1}$$

Using this inequality and the equation derived from (2.4) and (2.8) we have

$$a(y_\gamma, y_\gamma) + \frac{1}{\gamma}|\lambda_\gamma|_{L^2}^2 \le a(y_\gamma, y^*) + (f, y_\gamma - y^*) + \frac{1}{\gamma}\left(\bar\lambda, \lambda_\gamma\right).$$

It thus follows from (2.2) that

$$\nu\, |y_\gamma|_{H_0^1}^2 + \frac{1}{\gamma}|\lambda_\gamma|_{L^2}^2$$

is uniformly bounded with respect to $\gamma \ge 1$ and hence by (2.8) the family $\{\lambda_\gamma\}_{\gamma\ge1}$ is bounded in $H^{-1}(\Omega)$. Consequently there exist $\left(\hat y, \hat\lambda\right) \in H_0^1(\Omega) \times H^{-1}(\Omega)$ and a sequence $\{(y_{\gamma_n}, \lambda_{\gamma_n}\}$ with $\lim \gamma_n = \infty$ such that

$$w - \lim(y_{\gamma_n}, \lambda_{\gamma_n}) = \left(\hat y, \hat\lambda\right) \quad \text{in} \quad H_0^1(\Omega) \times H^{-1}(\Omega).$$

Henceforth we drop the subscript n with γ_n. Note that

$$\frac{1}{\gamma}|\lambda_\gamma|_{L^2}^2 = \gamma \left|\max\left(0, \frac{\bar{\lambda}}{\gamma} + y_\gamma - \psi\right)\right|_{L^2}^2.$$

Since $H_0^1(\Omega)$ is embedded compactly into $L^2(\Omega)$, we can assume without loss of the generality that y_γ converges to $\hat{y}$ a.e. in Ω. From the above equality and Fatou's lemma we conclude that $|(\hat{y} - \psi)^+| = 0$ and therefore $\hat{y} \le \psi$. From (2.4) and (2.8) we also have

$$a(y_\gamma - y^*, y_\gamma - y^*) + \langle \lambda_\gamma - \lambda^*, y_\gamma - y^* \rangle_{H^{-1}, H_0^1} = 0,$$

and by (3.1)

$$(\lambda_\gamma, y_\gamma - y^*) \ge -\frac{1}{4\gamma}|\bar{\lambda}|_{L^2}^2.$$

Hence

$$0 \le \overline{\lim_{\gamma\to\infty}}\, \nu\, |y_\gamma - y^*|_{H_0^1}^2 \le \lim_{\gamma\to\infty} \langle \lambda^*, y_\gamma - y^* \rangle_{H^{-1}, H_o^1} = \langle \lambda^*, \hat{y} - \psi \rangle_{H^{-1}, H_0^1} \le 0,$$

where we used the complementarity condition $\langle \lambda^*, y^* - \psi \rangle_{H^{-1}, H_0^1} = 0$ and $\hat{y} \le \psi$. It follows that $\lim_{\gamma\to\infty} y_\gamma = y^*$ in $H_0^1(\Omega)$ and hence $\hat{y} = y^*$. Taking the limit in

$$a(y_\gamma, v) + (\lambda_\gamma, v) = (f, v) \quad \text{for all} \quad v \in H_0^1,$$

we find

$$a(y^*, v) + \left\langle \hat{\lambda}, v \right\rangle_{H^{-1}, H_0^1} = (f, v) \quad \text{for all} \quad v \in H_0^1.$$

This equation is also satisfied with $\hat{\lambda}$ replaced by λ^* and consequently $\lambda^* = \hat{\lambda}$. Since (y^*, λ^*) is the unique solution to (2.5) the whole family $\{(y_\gamma, \lambda_\gamma)\}$ converges in the sense given in the statement of the theorem. □

In the next two sections we establish monotonicity for the family $\{y_\gamma\}_{\gamma>0}$ and the rate of convergence to y^* in $L^\infty(\Omega)$ for two specific selections of $\bar{\lambda}$. We believe that such results are new and they play an important role in developing a fast algorithm in Section 5.

3.1. **Infeasible case**

Here we choose $\bar{\lambda} = 0$. For $\gamma > 0$ we set

$$\mathcal{A}_\gamma = \{x \in \Omega \colon (y_\gamma - \psi)(x) > 0\} \quad \text{and} \quad \mathcal{I}_\gamma = \Omega \setminus \mathcal{A}_\gamma.$$

Proposition 3.1. *If* $0 < \alpha < \beta$ *then*

$$y^* \le y_\beta \le y_\alpha, \quad \textit{a.e. in} \ \ \Omega.$$

Proof. By (2.8) we have

$$\lambda_\alpha - \lambda_\beta = \max(0, \alpha(y_\alpha - \psi)) - \max(0, \beta(y_\beta - \psi)).$$

It follows that

$$\begin{aligned}&(\lambda_\alpha-\lambda_\beta)(x)=0 \ \text{ for } \ x\in\mathcal{I}_\alpha\cap\mathcal{I}_\beta\\&(\lambda_\alpha-\lambda_\beta)(x)\le\beta(y_\alpha-y_\beta)(x) \ \text{ for } \ x\in\mathcal{A}_\alpha\cap\mathcal{A}_\beta.\end{aligned}\tag{3.2}$$

For $x\in I_\beta\cap\mathcal{A}_\alpha$ we find $(\lambda_\alpha-\lambda_\beta)(x)=\alpha(y_\alpha-\psi)(x)\le\beta(y_\alpha-\psi)(x)-\beta(y_\beta-\psi)(x)=\beta(y_\alpha-y_\beta)(x)$, and thus

$$(\lambda_\alpha-\lambda_\beta)(x)\le\beta(y_\alpha-y_\beta)(x) \ \text{ for } \ x\in\mathcal{I}_\beta\cap\mathcal{A}_\alpha.\tag{3.3}$$

Moreover

$$(\lambda_\alpha-\lambda_\beta)(x)\le 0 \ \text{ for } \ x\in\mathcal{A}_\beta\cap\mathcal{I}_\alpha.\tag{3.4}$$

For (3.2)–(3.4) and (2.8) we find

$$a(y_\beta-y_\alpha,(y_\beta-y_\alpha)^+)=(\lambda_\alpha-\lambda_\beta,(y_\beta-y_\alpha)^+)\le 0$$

and hence $y_\beta\le y_\alpha$. The verification that $y^*\le y_\alpha$ is quite similar. □

Proposition 3.2. *For $0<\alpha<\beta$ we have*

$$\mathcal{I}^*\supset\mathcal{I}_\beta\supset\mathcal{I}_\alpha.$$

Proof. If $x\in\mathcal{A}_\beta\cap\mathcal{I}_\alpha$ then $(y_\alpha-\psi)(x)\le 0$ and $(y_\beta-\psi)(x)>0$. Hence $y_\alpha(x)<y_\beta(x)$ which contradicts Proposition 3.1 and therefore $\mathcal{I}_\beta\supset\mathcal{I}_\alpha$. □

Our next objective is to prove convergence of y_γ to y^* in $L^\infty(\Omega)$ with rate γ^{-1}, provided certain regularity conditions are satisfied. We require a technical lemma which we describe first. For this purpose let ω denote a subdomain of Ω with Lipschitzian boundary $\partial\omega$. The restriction of a to $H^1(\omega)\times H^1(\omega)$ will be denoted by the same symbol.

Lemma 3.1. *Assume that $g\in L^\infty(\omega)$ and that $a(1,v)\ge 0$, for all $v\in H^1(\omega)$ with $v\ge 0$. For $c>0, c\in\mathbb{R}$, let $y_c\in H_0^1(\omega)$ denote the solution to*

$$a(y,v)+c(y,v)_{L^2(\omega)}=(g,v)_{L^2(\omega)} \ \text{ for all } \ v\in H_0^1(\omega).\tag{3.5}$$

Then $y_c\in L^\infty(\omega)$ and $|y_c|_{L^\infty(\omega)}\le\frac{1}{c}|g|_{L^\infty(\omega)}$.

Proof. For the sake of completeness we include the proof which can be obtained with known techniques. Let $\bar g=\max(0,\sup\{g(x)\colon x\in\omega\})$.

For all $v\in H_0^1(\omega)$ we have

$$a\left(y_c-\frac{1}{c}\bar g,v\right)+(\bar g-g,v)_{L^2(\omega)}=(\bar g-cy_c,v)_{L^2(\omega)}-a\left(\frac{\bar g}{c},v\right).\tag{3.6}$$

Set $v=\left(y_c-\frac{1}{c}\bar g\right)^+$ and observe that $v\in H_0^1(\omega)$ since $\bar g\ge 0$. Since $a(1,v)\ge 0$ for all $v\in H^1(\omega)$ and $v\ge 0$, it follows from (3.6) that

$$a\left(y_c-\frac{1}{c}\,\bar g,\left(y_c-\frac{1}{c}\,\bar g\right)^+\right)\le 0$$

and consequently $y(x) \leq \frac{1}{c}|g|_{L^\infty(\omega)}$ for a.e. $x \in \omega$. Analogously it can be verified that $-\frac{1}{c}|g|_{L^\infty(\omega)} \leq y(x)$ for a.e. $x \in \omega$. □

Let us introduce the active and inactive sets associated to the solution y^* of (1.1):

$$\mathcal{A}^* = \{x \in \Omega \colon y^*(x) = \psi(x)\}, \quad \mathcal{I}^* = \{x \in \Omega \colon y^*(x) < \psi(x)\},$$

with boundaries $\partial\mathcal{A}^*$ and $\partial\mathcal{I}^*$ respectively.

Theorem 3.2. *Let the regularity requirements of Remark 2.3 be satisfied and assume that $f \in L^\infty(\Omega)$ and $A\psi \in L^\infty(\Omega)$. If, moreover, the boundary $\partial\mathcal{A}^*$ of the active set is $C^{1,1}$ manifold in R^{n-1} and for every $\gamma > 0$ the boundary $\partial\mathcal{A}_\gamma$ of $\mathcal{A}_\gamma$ is Lipschitzian manifold in R^{n-1} , then*

$$|y_\gamma - y^*|_{L^\infty(\Omega)} \leq \frac{1}{\gamma}|f - A\psi|_{L^\infty(\Omega)}.$$

Proof. The regularity assumption imply that $y^* \in W^{2,p}(\Omega)$ and $y_\gamma \in W^{2,p}(\Omega)$ with $p > n$. Recall that $W^{2,p}(\Omega) \subset C(\bar{\Omega})$ if $p > n$. From Proposition 3.2 it follows that $\mathcal{A}^* \subset \mathcal{A}_\gamma$ for every $\gamma > 0$. From the definition of $\mathcal{A}_\gamma$ we have

$$\begin{cases} Ay_\gamma + \gamma(y_\gamma - \psi) = f & \text{in } \mathcal{A}_\gamma \\ y_\gamma - \psi = 0 & \text{on } \partial\mathcal{A}_\gamma. \end{cases}$$

Consequently

$$\begin{cases} A(y_\gamma - \psi) + \gamma(y_\gamma - \psi) = f - A\psi & \text{in } \mathcal{A}_\gamma \\ y_\gamma - \psi = 0 & \text{on } \partial\mathcal{A}_\gamma. \end{cases}$$

From Lemma 3.1 with $\omega = \mathcal{A}_\gamma$ and $g = f$ we find

$$|y_\gamma - \psi|_{L^\infty(\mathcal{A}_\gamma)} \leq \frac{1}{\gamma}|f - A\psi|_{L^\infty(\Omega)}$$

and in particular

$$|y_\gamma - \psi|_{L^\infty(\mathcal{A}^*)} \leq \frac{1}{\gamma}|f - A\psi|_{L^\infty(\Omega)}. \tag{3.7}$$

Note further that on $\mathcal{I}^*$ we have

$$\begin{cases} A(y_\gamma - y^*) = \lambda^* - \lambda_\gamma \leq 0 & \text{in } \mathcal{I}^* \\ y_\gamma - y^* = y_\gamma - \psi \geq 0 & \text{on } \partial\mathcal{I}^*. \end{cases} \tag{3.8}$$

From the maximum principle applied to (3.8), and (3.7) it follows that

$$|y_\gamma - y^*|_{L^\infty(\mathcal{I}^*)} \leq |y_\gamma - \psi|_{L^\infty(\partial\mathcal{I}^*)} \leq \frac{1}{\gamma}|f - A\psi|_{L^\infty(\Omega)}, \tag{3.9}$$

see *e.g.* [13] (p. 191). Combining (3.7) and (3.9) gives the desired conclusion. □

To justify the terminology to refer to $\bar{\lambda} = 0$ as the infeasible case note that if $y_\gamma < \psi$ for some $\gamma > 0$ then $\mathcal{I}_\gamma = \Omega$, $\lambda_\gamma = 0$ and $(y_\gamma, \lambda_\gamma)$ satisfy the optimality system (2.4). Consequently $(y^*, \lambda^*) = (y_\gamma, \lambda_\gamma)$ and y^* is also a solution of the unconstrained problem. Thus unless y^* is also a solution to the unconstrained problem, $y_\gamma \leq \psi$ for some finite γ is impossible. In the following section it will be shown that proper choice of $\bar{\lambda}$ guarantees feasibility of the solutions y_γ to (2.8).

3.2. **Feasible case**

Here we choose $\bar{\lambda} \in L^2(\Omega)$ such that

$$\begin{cases} \bar{\lambda} \geq 0 \ \text{ and } \ \bar{\lambda} - (f - A\psi) \geq 0, \ \text{ a.e. in } \ \Omega \\ \langle \bar{\lambda} - (f - A\psi), v \rangle_{H^{-1}, H_0^1} \geq 0 \ \text{ for all } \ v \in K. \end{cases} \tag{3.10}$$

Note that if $\psi \in H^2(\Omega)$ then for the choice $\bar{\lambda} = \max(0, f - A\psi)$ (3.10) is satisfied.

Proposition 3.3. *If (3.10) holds and* $0 < \alpha < \beta$ *then*

$$y_\alpha \leq y_\beta \leq \psi \quad \textit{a.e. in} \ \ \Omega.$$

In particular y_α *is feasible for every* $\alpha > 0$.

Proof. From (2.8) we have by (3.10)

$$\begin{aligned} a(y_\alpha - \psi, (y_\alpha - \psi)^+) &= (f - \lambda_\alpha, (y_\alpha - \psi)^+) - a(\psi, (y_\alpha - \psi)^+) \\ &= \langle f - A\psi - \max(0, \bar{\lambda} + \alpha(y_\alpha - \psi)), (y_\alpha - \psi)^+ \rangle \\ &= \langle f - A\psi - \bar{\lambda}, (y_\alpha - \psi)^+ \rangle - \alpha(y_\alpha - \psi, (y_\alpha - \psi)^+) \\ &\leq -\alpha |\, (y_\alpha - \psi)^+|^2 \leq 0 \end{aligned}$$

and hence by (2.3)

$$y_\alpha \leq \psi.$$

It follows that y_α is feasible for every $\alpha > 0$.

Next let $0 < \alpha < \beta$. By (2.8)

$$a(y_\alpha - y_\beta, (y_\alpha - y_\beta)^+) = (\lambda_\beta - \lambda_\alpha, (y_\alpha - y_\beta)^+). \tag{3.11}$$

We introduce the set

$$S = \{x \colon y_\alpha(x) - y_\beta(x) > 0\}$$

and decompose this set as $S = S_1 \cap S_2 \cup S_3$, where

$$\begin{aligned} S_1 &= \left\{x \in S \colon \ \left(\bar{\lambda} + \beta(y_\beta - \psi)\right)(x) \leq 0\right\} \\ S_2 &= \left\{x \in S \colon \ \left(\bar{\lambda} + \beta(y_\beta - \psi)\right)(x) > 0, \left(\bar{\lambda} + \alpha(y_\alpha - \psi)\right)(x) \leq 0\right\} \\ S_3 &= \left\{x \in S \colon \ \left(\bar{\lambda} + \beta(y_\beta - \psi)\right)(x) > 0, \left(\bar{\lambda} + \alpha(y_\alpha - \psi)\right)(x) > 0\right\}. \end{aligned}$$

To estimate the right hand side of (3.11) recall that

$$\lambda_\beta - \lambda_\alpha = \max\left(0, \bar{\lambda} + \beta(y_\beta - \psi)\right) - \max\left(0, \bar{\lambda} + \alpha(y_\alpha - \psi)\right).$$

We find

$$\begin{aligned}(\lambda_\beta - \lambda_\alpha, (y_\alpha - y_\beta)^+) &= (\lambda_\beta - \lambda_\alpha, y_\alpha - y_\beta)_{L^2(S_1)} + (\lambda_\beta - \lambda_\alpha, y_\alpha - y_\beta)_{L^2(S_2)} + (\lambda_\beta - \lambda_\alpha, y_\alpha - y_\beta)_{L^2(S_3)} \\ &\le (\beta(y_\beta - \psi) - \alpha(y_\alpha - \psi), (y_\alpha - y_\beta)_{L^2(S_2)} + (\beta(y_\beta - y_\alpha), y_\alpha - y_\beta)_{L^2(S_3)} \\ &\quad +(\beta(y_\alpha - \psi) - \alpha(y_\alpha - \psi), y_\alpha - y_\beta)_{L^2(S_3)}.\end{aligned}$$

Utilizing the fact that $y_\alpha \le \psi$ and $y_\beta \le \psi$ we find

$$(\lambda_\beta - \lambda_\alpha, (y_\alpha - y_\beta)^+) \le \beta(y_\beta - y_\alpha, y_\alpha - y_\beta)_{L^2(S_2)} + (\beta - \alpha)(y_\alpha - \psi, y_\alpha - y_\beta)_{L^2(S_3)} \le 0.$$

Inserting this estimate into (3.11) and using the weak maximum principle implies that $y_\alpha \le y_\beta$. □

Corollary 3.1. *If (3.10) holds and* $0 < \alpha < \beta$ *then*

$$0 \le \lambda_\alpha \le \lambda_\beta \le \max\left(0, \bar\lambda\right)$$

and $\mathcal{I}_\alpha \supset \mathcal{I}_\beta$.

Proof. From the representation $\lambda_\gamma = \max\left(0, \bar\lambda + \gamma(y_\gamma - \psi)\right)$ and the fact that $\gamma \to \gamma(y_\gamma - \psi)(x)$ is increasing with respect to γ for a.e. $x \in \Omega$, it follows that λ_γ is increasing and $\mathcal{I}_\gamma$ is decreasing with respect to γ. The estimate $\lambda_\gamma \le \max\left(0, \bar\lambda\right)$ is a consequence of the feasibility of y_γ for every γ. □

As in the infeasible case we can consider the question of rate of convergence with respect to γ if additional regularity requirements are satisfied.

Remark 3.1. From Theorem 3.1, Corollary 3.1 and Lebesgue's monotone convergence theorem it follow that $\lambda_\gamma \to \lambda^*$ strongly in $L^2(\Omega)$.

Theorem 3.3. *Assume that* $f \in L^\infty(\Omega)$, $A\psi \in L^\infty(\Omega)$, $\bar\lambda \in L^\infty(\Omega)$ *and that the assumptions of Remark 2.3 hold. If in addition* $\mathcal{A}_\gamma$ *is a domain with a* $C^{1,1}$ *boundary, then*

$$|y_\gamma - y^*|_{L^\infty(\Omega)} \le \frac{1}{\gamma}\,|\bar\lambda|_{L^\infty(\Omega)}.$$

Proof. By the assumptions of the theorem y^* and $y_\gamma \in W^{2,p}(\Omega)$, $p > n$. On $\overline{\mathcal{A}_\gamma}$ we have $\bar\lambda + \gamma(y_\gamma - \psi) \ge 0$ and $y_\gamma \le \psi$, and hence

$$|y_\gamma - \psi|_{L^\infty(\mathcal{A}_\gamma)} \le \frac{1}{\gamma}\,|\bar\lambda|_{L^\infty(\Omega)}.$$

Since $\mathcal{A}_\gamma \subset \mathcal{A}^*$ by Corollary 3.1 this implies that

$$|y_\gamma - y^*|_{L^\infty(\mathcal{A}_\gamma)} \le \frac{1}{\gamma}\,|\bar\lambda|_{L^\infty(\Omega)}.$$

Moreover we have

$$\begin{cases} A(y^* - y_\gamma) = \lambda_\gamma - \lambda^* \le 0 & \text{in } \mathcal{I}_\gamma \\ y^* - y_\gamma \le 0 & \text{on } \partial\mathcal{I}_\gamma. \end{cases}$$

From the maximum principle and the regularity assumption on $\partial\mathcal{I}_\gamma$ it follows that

$$|y^* - y_\gamma|_{L^\infty(\mathcal{I}_\gamma)} \le \frac{1}{\gamma}\,|\bar\lambda|_{L^\infty(\Omega)}.$$ □

4. Bilateral constraints

The treatment of bilateral constraints gives rise to some additional difficulties. Here we consider

$$\begin{cases} \min \ \dfrac{1}{2}\, a(y,y) - (f,y) \\ \text{over } y \in H_0^1(\Omega) \\ \varphi \le y \le \psi \ \text{ in } \Omega. \end{cases} \tag{4.1}$$

Throughout this section we assume that

$$a(\varphi,\psi) = (\nabla\varphi, \nabla\psi) \text{ for all } \varphi,\psi \in H_0^1(\Omega),$$

that φ and ψ are in $H^1(\Omega)$, that $\partial\Omega$ is $C^{1,1}$ and

$$\begin{aligned} \varphi \le 0 \le \psi \text{ on } \partial\Omega, \quad & \max(0, \Delta\psi + f) \in L^2(\Omega), \\ & \min(0, \Delta\varphi + f) \in L^2(\Omega), \end{aligned} \tag{4.2}$$

$$S_1 := \{x \in \Omega \colon \Delta\psi + f > 0\} \cap S_2 := \{x \in \Omega \colon \Delta\varphi + f < 0\} = \emptyset, \tag{4.3}$$

and that there exists $c_0 > 0$ such that

$$-\Delta(\psi - \varphi) + c_0(\psi - \varphi) \ge 0 \ \text{ a.e. in } \Omega. \tag{4.4}$$

Under these assumptions it was shown in [10] that there exists a solution $y^* \in H_0^1(\Omega) \cap H^2(\Omega)$ to (4.1) with associated Lagrange multiplier $\lambda^* \in L^2(\Omega)$. This was verified by passing to the limit $\gamma \to \infty$ in

$$-\Delta y_\gamma + \lambda_\gamma = f, \quad \lambda_\gamma = \begin{cases} \bar\lambda + \gamma\,(y_\gamma - \psi) & \text{if} \quad \bar\lambda + \gamma\,(y_\gamma - \psi) > 0, \\ \bar\lambda + \gamma\,(y_\gamma - \varphi) & \text{if} \quad \bar\lambda + \gamma\,(y_\gamma - \varphi) < 0, \\ 0 & \text{otherwise}, \end{cases} \tag{4.5}$$

where

$$\bar\lambda = \begin{cases} \Delta\psi + f & \text{on } S_1 \\ \Delta\varphi + f & \text{on } S_2 \\ 0 & \text{otherwise.} \end{cases} \tag{4.6}$$

The weak limit of $(y_\gamma, \lambda_\gamma)$ in $H^2(\Omega) \times L^2(\Omega)$ satisfies

$$-\Delta y^* + \lambda^* = f, \quad \lambda^* = \begin{cases} \lambda^* + c\,(y^* - \psi) & \text{if} \quad \lambda^* + c\,(y^* - \psi) > 0, \\ \lambda^* + c\,(y^* - \varphi) & \text{if} \quad \lambda^* + c\,(y^* - \varphi) < 0, \\ 0 & \text{otherwise}, \end{cases}$$

for every $c > 0$. The latter equation can be equivalently expressed as

$$\begin{cases} -\Delta y^* + \lambda^* = f, \\ \lambda^* = \max(0, \lambda^* + c(y^* - \psi)) + \min(0, \lambda^* + c(y^* - \varphi)), \end{cases} \tag{4.7}$$

for every arbitrary fixed $c > 0$.

Primal–dual active set algorithm

(i) Choose y_0, set $k = 0$.
(ii) Set

$$\begin{aligned}
\mathcal{A}_{k+1}^{\psi} &= \left\{x\colon \left(\bar\lambda + \gamma(y_k - \psi)\right)(x) > 0\right\},\\
\mathcal{A}_{k+1}^{\varphi} &= \left\{x\colon \left(\bar\lambda + \gamma(y_k - \varphi)\right)(x) < 0\right\}\\
\mathcal{I}_{k+1} &= \Omega \setminus \left(\mathcal{A}_{k+1}^{\psi} \cup \mathcal{A}_{k+1}^{\varphi}\right).
\end{aligned}$$

(iii) Solve for $y_{k+1} \in H_0^1(\Omega)$.

$$a(y,v) + \left(\left(\bar\lambda + (y-\psi)\right)\chi_{\mathcal{A}_{k+1}^{\psi}}, v\right) + \left(\left(\bar\lambda + (y-\varphi)\right)\chi_{\mathcal{A}_{k+1}^{\varphi}}, v\right) = (f,v), \tag{4.8}$$

for all $v \in H_0^1(\Omega)$. Set

$$\lambda_{k+1} = \begin{cases} 0 & \text{on } \mathcal{I}_{k+1}\\ \bar\lambda + \gamma(y_{k+1} - \psi) & \text{on } \mathcal{A}_{k+1}^{\psi}\\ \bar\lambda + \gamma(y_{k+1} - \varphi) & \text{on } \mathcal{A}_{k+1}^{\varphi}. \end{cases}$$

(iv) Stop, or $k = k+1$ and goto (ii).

For the following local convergence result the choice of $\bar\lambda$ as in (4.6) is not essential.

Proposition 4.1. *If $|\lambda_0 - \lambda_\gamma|_{L^2(\Omega)}$ is sufficiently small then $(y_k, \lambda_k) \to (y_\gamma, \lambda_\gamma)$ superlinearly in $H_0^1(\Omega) \times L^2(\Omega)$.*

Proof. The proof is quite similar to that of Theorem 2.2 and we therefore only give a brief outline. Again the algorithm is expressed in the variable λ only. The resulting iteration map $F\colon L^2(\Omega) \to L^2(\Omega)$ is given by

$$F(\lambda) = \lambda - \max\left(0, \bar\lambda + \gamma\left(A^{-1}(f-\lambda) - \psi\right)\right) - \min\left(0, \bar\lambda + \gamma\left(A^{-1}(f-\lambda) - \varphi\right)\right),$$

and (4.7) is equivalent to $F(\lambda) = 0$. Steps (ii) and (iii) of the reduced algorithm are replaced by:

(ii') Set

$$\begin{aligned}
\mathcal{A}_{k+1}^{\psi} &= \left\{x\colon \left(\bar\lambda + \gamma A^{-1}(f - \lambda_k) - \gamma\,\psi\right)(x) > 0\right\},\\
\mathcal{A}_{k+1}^{\varphi} &= \left\{x\colon \left(\bar\lambda + \gamma A^{-1}(f - \lambda_k) - \gamma\,\varphi\right)(x) < 0\right\},\\
\mathcal{I}_{k+1} &= \Omega \setminus \left(\mathcal{A}_{k+1}^{\psi} \cup \mathcal{A}_{k+1}^{\varphi}\right).
\end{aligned}$$

(iii') Set

$$\begin{aligned}
&\delta\,\lambda = -\lambda_k \ \text{ on } \mathcal{I}_{k+1} \ \text{ and solve for } \ \delta\,\lambda \in H^{-1}\\
&\delta\,\lambda + \gamma A^{-1}(\delta\,\lambda) = -\lambda_k + \bar\lambda - \gamma\,\psi + \gamma A^{-1}(f - \lambda_k) \ \text{ in } \ \mathcal{A}_{k+1}^{\psi}\\
&\delta\,\lambda + \gamma A^{-1}(\delta\,\lambda) = -\lambda_k + \bar\lambda - \gamma\,\varphi + \gamma A^{-1}(f - \lambda_k) \ \text{ in } \ \mathcal{A}_{k+1}^{\varphi}.
\end{aligned}$$

As in the proof of Theorem 2.2 one argues that F is Newton–differentiable. To characterize the generalized derivative we set

$$c_\psi = \bar\lambda + \gamma\left(A^{-1}(f-\lambda) - \psi\right),\ c_\varphi = \bar\lambda + \gamma\left(A^{-1}(f-\lambda) - \varphi\right),$$

and

$$\mathcal{A}_\psi = \{x\colon c_\psi(x) > 0\},\ \mathcal{A}_\varphi = \{x\colon c_\varphi(x) < 0\},\ \mathcal{I} = \Omega \setminus (\mathcal{A}_\psi \cup \mathcal{A}_\varphi).$$

A generalized derivative is given by

$$G(\lambda) = I + \gamma\,\chi_{\mathcal{A}_\psi} A^{-1} + \gamma\,\chi_{\mathcal{A}_\varphi} A^{-1} = I + \gamma\,\chi_{\mathcal{A}} A^{-1},$$

where $\mathcal{A} = \mathcal{A}_\varphi \cup A_\psi$. Existence and uniform boundedness of the inverses of $G(\lambda)$ is verified as in the proof of Theorem 2.2. □

5. Numerical experiments

In this section we describe some numerical experiments to illustrate and confirm our results. The problem under consideration is

$$\begin{cases} -\Delta y + \lambda = f & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega \\ y \le \psi,\ \lambda \ge 0, & (\lambda, y-\psi)_{L^2(\Omega)} = 0 \end{cases} \tag{5.1}$$

which is discretized by means of node–based finite differences. In the one-dimensional case $\Omega = (0,1)$ with grid $\{x_i\} = \{\frac{i}{m}\}_{i=0}^m$ and in the two–dimensional case $\Omega = (0,1)\times(0,1)$ with grid $\{x_{i,j}\} = \{(\frac{i}{m}, \frac{j}{m}\}_{i,j=0}^m$. The functions y, λ, f and ψ are discretized by grid functions denoted by the same symbol and $-\Delta$ is discretized by the three-, respectively five-point finite difference stencil. It is well-known that the resulting discretized operator $-\Delta_h$ satisfies the discrete maximum principle. Unless specified otherwise the algorithms are initialized with the unconstrained solution to (5.1), *i.e.* $\psi = \infty$.

Example 5.1. Here $\Omega = (0,1)$, $f = \frac{1}{4} \times \sin(5x)$, $\psi = \frac{1}{4}$, and $m = 100$. For all runs that we report upon the primal–dual active set algorithm converges in finitely many steps, *i.e.* the situation discussed in Proposition 2.1 occurs. We denote the number of iterations that are required until the algorithm reaches the solution of the discretized problem by iter. For $\gamma > 0$ the iterates of the algorithm are denoted by y_k, the solution by y_γ. Similarly $\mathcal{A}_k$ stands for the active sets of the iterates, $\mathcal{A}_\delta$ for the active set corresponding to solution y_δ. In this example as well as in Examples 5.2 and 5.3 below the algorithm was terminated when in two successive iterations the active sets coincide. The current variables then give the solution of the discretized problem.

Let us start with some general observations for the numerical solution:

- $y_{\gamma_2} \le y_{\gamma_1}$ for $\gamma_1 \le \gamma_2$ and $\bar\lambda = 0$;
- $y_{\gamma_1} \le y_{\gamma_2}$ for $\gamma_1 \le \gamma_2$ and $\bar\lambda = \max(0, f + \Delta_k\psi)$;
- $\mathcal{A}_{k+1} \subset \mathcal{A}_k$ for $\gamma > 0$ and $\bar\lambda = 0$ or $\bar\lambda = \max(0, f + \Delta_h\psi)$;
- iter $(\gamma_1) \ge$ iter (γ_2) if $\gamma_1 \le \gamma_2$;
- for large γ changes after the initialization phase from active to inactive occur only along the boundary of $\mathcal{A}_k$. This is not the case for small γ.

In Table 1 we report the required number of iterations and the cardinality of the active set $\mathcal{A}$ as a function of γ, for $\bar\lambda = 0$.

Table 1.

γ	2.5	5	10	20	100	1000	10000
iter	4	4	6	6	15	20	20
$\#(\mathcal{A}_\gamma)$	29	26	23	22	19	18	18

The results of Table 1 suggest to combine the primal–dual active set strategy with a continuation procedure with respect to γ: Thus we start with small γ and use the solution as initialization for the algorithm with larger γ. Table 2 shows that this continuation method is effective.

TABLE 2.

γ	5	20	10^4
iter	5	3	4

Concerning superlinear convergence of the algorithm for fixed $\gamma \in (1, \infty)$ it is not obvious whether the continuous result can be used as indicator for the discrete one, due to finite speed of propagation of the discrete Laplacian. In Table 3 we report the results for the quotients

$$q_k = (y_{k+1} - y_{20})^T \Delta_h (y_{k+1} - y_{20}) \big/ (y_k - y_{20}) \Delta_h (y_k - y_{20}),$$

for selected values of k, where $\bar{\lambda} = 0$, $\gamma = 10^4$.

TABLE 3.

k	2	6	10	13	14	15	16	17	18
q_k	0.84	0.80	0.72	0.62	0.58	0.51	0.43	0.31	0.13

It is quite typical for the runs that we tested that the quotients decrease approximately by one power of 10, between initialization and final result.

Example 5.2. For this example $\Omega = (0,1) \times (0,1)$, $f = 500\, x \sin(5x) \cos(2y)$, $\psi = 1$ on the annulus, $R = \{(x,y) \colon .2 < \sqrt{x^2 + y^2} < .4\}$, $\psi = 10$ on $\Omega \quad R$, and $m = 200$. Again the algorithm with $\bar{\lambda} = 0$ and $\bar{\lambda} = \max(0, f + \Delta_h \psi)$ terminates after finitely many iterations. The same observations can be made as for the one-dimensional example above. Typical results for $\bar{\lambda} = 0$ and various values of γ are given in Table 4.

TABLE 4.

γ	10^3	10^4	10^5	10^6	10^7	10^8	10^9
iter	5	8	12	27	35	36	36
$\#(\mathcal{A}_\gamma)$	3117	2530	2348	2306	2302	2301	2301

For $\bar{\lambda} = \max(0, f + \Delta_h \psi)$ the number of required iterations is comparable and the final active sets for $\gamma \geq 10^8$ is the same. For $\gamma = 10^3$ changes from active to inactive sets take place along the boundaries of these sets in layers up to the depth of 16 pixels. Continuation procedures with respect to γ as explained in Example 5.1 again reduce the total number of iterations significantly, see Table 5.

TABLE 5.

γ	10^4	10^6	10^8
iter	8	5	1

We carried out computations with the same specifications as in Table 5 with a series of mesh-sizes characterized by $m = (100, 200, 300, 400)$. The resulting number of total iterations are $(11, 14, 16, 20)$. Again superlinear convergence of the iterates can be observed. In Table 6 we give selected results for the quotients q_k with $m = 200$, $\gamma = 10^8$ and $\bar{\lambda} = 0$. Since in this case the algorithm terminates in 36 iterations we set $q_f = q_{36}$.

TABLE 6.

k	1	8	15	22	29	31	33	35
q_k	0.86	0.82	0.79	0.75	0.55	0.42	0.21	0.17

Tests with the smooth obstacle $\psi = 8\left(\left(x-\frac{1}{2}\right)^2+\left(y-\frac{1}{2}\right)^2\right)-1$ give quite similar results. The iteration procedure with the same values for γ as in Table 5 requires 16 iterations to obtain the solution, without continuation procedure 44, for $\gamma = 10^8$.

Example 5.3. This is an example with lack of strict complementarity. The choice for Ω, and f is as in Example 5.2. We set m=40. Let y_h^* denote the solution to the unconstrained problem $-\Delta_h y_h = f$, and define $\psi = 10$ except on $S = \left(\frac{3}{8}, \frac{5}{8}\right)$, where it is set equal to y_h^*. The algorithm with both $\bar{\lambda} = 0$ and $\bar{\lambda} = \max(0, f + \Delta_h\psi)$ terminates in 1 iteration for a large range of γ–values. In a similar experiment with $m = 30$ and S replaced by $\left(\frac{1}{3}, \frac{2}{3}\right) \times \left(\frac{1}{3}, \frac{2}{3}\right)$ the algorithm starts to chatter if $\bar{\lambda} = 0$, while it converges in finitely many iterations comparable to those in Table 4 for $\bar{\lambda} = \max(0, f + \Delta_h\psi)$. Due to finite precision arithmetic and the fact that the active/inactive set structure and the stopping rule are determined by commands involving machine zero, chattering in the case of lack of strict complementarity comes as no surprise. There are various remedies to avoid chattering based on stopping rules involving machine epsilon. The alternative choice of using $\bar{\lambda} = \max(0, f + \Delta_h\psi)$ rather than $\bar{\lambda} = 0$ has consistently eliminated chattering in this and other examples. For instance, again with $m = 30$, we chose $\psi = 10$ on $\Omega\backslash S$ and $\psi = y_h^* - 1$. In the interior of the active set we have lack of strict complementarity and for $\bar{\lambda} = 0$ and $\gamma > 10^6$ the iterates chatter. With $\bar{\lambda} = \max(0, f + \Delta_h\psi)$ no chattering occurs.

In Examples 5.1 and 5.2 we investigated the case when the penalty parameter tends to ∞. For a specific application it may be desirable to compute with a fixed penalty parameter. For this purpose the penalty parameter should be chosen such that the error due to penalization is of the same order as that due to discretization. Theorems 3.2 and 3.3 then suggest to choose γ proportional to h^{-2}. The success of this procedure is illustrated in the following example.

Example 5.4. The following example from [5] (pp. 44–45). It represents an elasto–plastic torsion problem formulated as obstacle problem on the unit disc $\widetilde{\Omega}$ with center at $(0.5, 0.5)$. Let $r = \sqrt{(x_1-0.5)^2+(x_2-0.5)^2}$, $d > 2$ be a constant and set

$$\psi(x) = 1 - r, \quad \text{and} \quad f(x) = d.$$

Then the solution to the obstacle problem on $\widetilde{\Omega}$ is given by

$$y(x) = \begin{cases} 1 - r & \text{if } \dfrac{2}{d} \le r \le 1 \\ \dfrac{d}{4}\left[\left(1-r^2\right) - \left(r - \dfrac{2}{d}\right)^2\right] & \text{if } 0 \le r \le \dfrac{2}{d}. \end{cases}$$

In our calculation we chose $d = 5.123$.

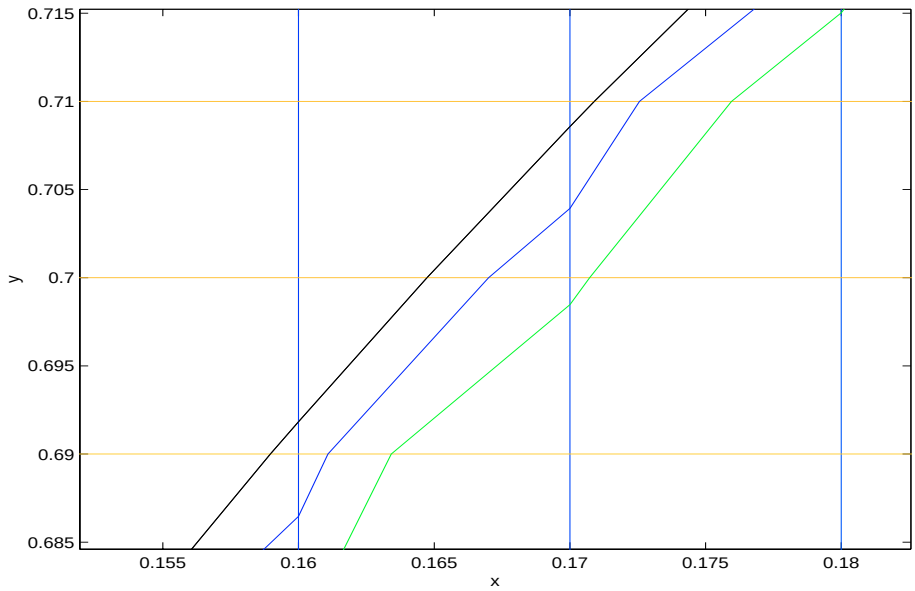

FIGURE 1. Infeasible method.

To compute on the unit square Ω we used exact non–homogeneous boundary conditions assigned at the boundary. The regularization parameter was $\gamma = \left|\bar{\lambda}\right|_\infty / h^2$ with $h = 0.01$. The exact interface Γ between the active and inactive sets is given by $r = 1 - \frac{2}{d}$. We use this example to demonstrate how the interface can be approximated by our proposed algorithms. In the following figures we show the estimated interface for both the infeasible (Γ_1) and feasible (Γ_2) method determined by means of

$$\begin{aligned} \Gamma_1 &= \{x \in R^2 \colon y - \psi = 0\} \\ \Gamma_2 &= \{x \in R^2 \colon \bar{\lambda} + \gamma(y - \psi) = 0\} \end{aligned} \tag{5.2}$$

or alternatively by

$$\begin{aligned} \Gamma_1 &= \left\{x \in R^2 \colon y - \psi = \frac{1}{\gamma}\right\} \\ \Gamma_2 &= \left\{x \in R^2 \colon \bar{\lambda} + \gamma(y - \psi) = \frac{1}{\gamma}\right\}. \end{aligned} \tag{5.3}$$

Figure 1 shows a blow–up section of Γ and the Γ_1's for the infeasible method and Figure 2 shows a blow–up section of Γ and the Γ_2's for the feasible method. The most outer curve is for the exact interface Γ both in Figures 1 and 2. In this example the second estimates by (5.3) provide better and smoother estimates of the interface Γ both for the infeasible and feasible methods than (5.2).

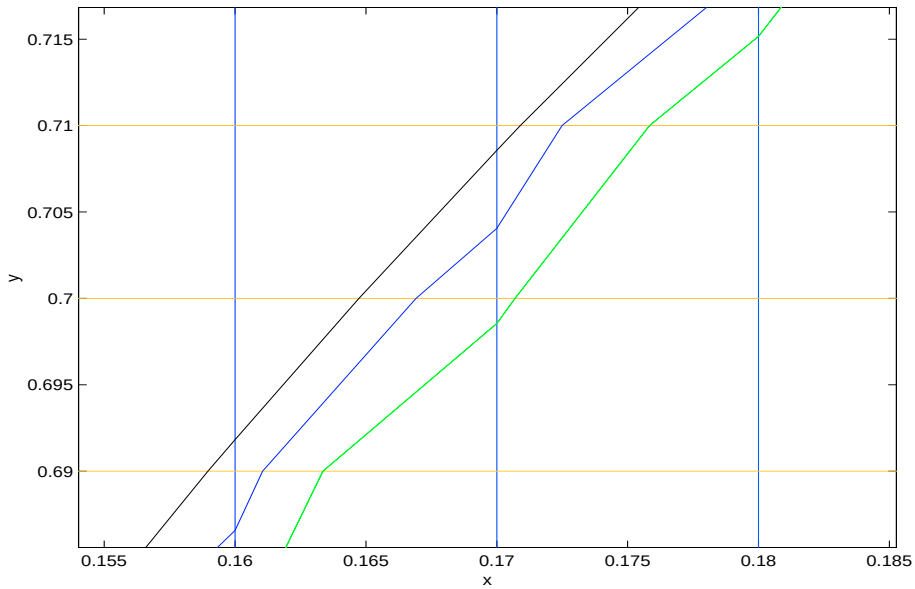

FIGURE 2. Feasible method.

REFERENCES

[1] D.P. Bertsekas, *Constrained Optimization and Lagrange Mulitpliers*. Academic Press, New York (1982).

[2] 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. Optim.* **11** (2000) 495–521.

[3] M. Bergounioux, K. Ito and K. Kunisch, Primal–dual strategy for constrained optimal control problems. *SIAM J. Control Optim.* **37** (1999) 1176–1194.

[4] Z. Dostal, Box constrained quadratic programming with proportioning and projections. *SIAM J. Optim.* **7** (1997) 871–887.

[5] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*. Springer Verlag, New York (1984).

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

[7] M. Hintermüller, K. Ito and K. Kunisch, The primal–dual active set strategy as semi–smooth Newton method. *SIAM J. Optim.* (to appear).

[8] R. Hoppe, Multigrid algorithms for variational inequalities. *SIAM J. Numer. Anal.* **24** (1987) 1046–1065.

[9] R. Hoppe and R. Kornhuber, Adaptive multigrid methods for obstacle problems. *SIAM J. Numer. Anal.* **31** (1994) 301–323.

[10] K. Ito and K. Kunisch, Augmented Lagrangian methods for nonsmooth convex optimization in Hilbert spaces. *Nonlinear Anal.* **41** (2000) 573–589.

[11] K. Ito and K. Kunisch, Optimal control of elliptic variational inequalities. *Appl. Math. Optim.* **41** (2000) 343–364.

[12] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*. Academic Press, New York (1980).

[13] D.M. Troianello, *Elliptic Differential Equations and Obstacle Problems*. Plenum Press, New York (1987).

[14] M. Ulbrich, Semi-smooth Newton methods for operator equations in function space. *SIAM J. Optim.* (to appear).

To access this journal online:
www.edpsciences.org