Proceedings of the International Congress of Mathematicians
Hyderabad, India, 2010

Novel Concepts for Nonsmooth Optimization and their Impact on Science and Technology

Kazufumi Ito, Karl Kunisch *

Abstract. A multitude of important problems can be cast as nonsmooth variational problems in function spaces, and hence in an infinite-dimensional, setting. Traditionally numerical approaches to such problems are based on first order methods. Only more recently Newton-type methods are systematically investigated and their numerical efficiency is explored. The notion of Newton differentiability combined with path following is of central importance. It will be demonstrated how these techniques are applicable to problems in mathematical imaging, and variational inequalities. Special attention is paid to optimal control with partial differential equations as constraints.

Mathematics Subject Classification (2000). 35Q93, 46N10, 49K20, 65K10.

Keywords. Non-smooth optimization, semi-smooth Newton methods, optimal control, complementarity problems, ill-posed problems.

1. Introduction

Let X, H be real Hilbert spaces and K a closed convex subspace of X. Identify H with H^* and let $\langle \cdot, \cdot \rangle$ denote the duality product on $X^* \times X$. We consider the minimization problem

$$\min \quad f(x) + \varphi(\Lambda x) \quad \text{over } x \in K, \tag{P}$$

where $f : X \to \mathbb{R}$ is a lower semi-continuous, continuously differentiable, convex function, $\Lambda \in \mathcal{L}(X, H)$ and $\varphi : H \to (-\infty, \infty]$ is a proper, lower semi-continuous, convex function. Typically X and H will be real-valued function spaces over a bounded domain $\Omega \subset \mathbb{R}^n$ with smooth boundary $\partial\Omega$. This is a problem that is well-studied within convex analysis framework. This aspect, as well as first order numerical iterative solution methods are reviewed in part from a non-classical perspective in Section 2. Since φ is not assumed to be regular, classical Newton methods are not directly applicable. In Section 3 the concept of Newton-differentiability and semi-smooth Newton methods are introduced. In the subsequent sections the

*The work of this author was supported by the Austrian Science Fund (FWF) under grant SFB F32 (SFB "Mathematical Optimization and Applications in Biomedical Sciences").

applicability of these tools is demonstrated for a wide variety of topics, including optimal boundary control in Section 4, optimal control with sparsity constraints in Section 5, time optimal control in Section 6, and data fitting problems in Section 7. The final Section 8 is devoted to a general class of non-linear, non-differentiable complementarity problems. Most of these applications involve differential equations.

2. First order Augmented Lagrangian Method

In this section we summarize convex analysis techniques for solving (P). For basic convex analysis concepts see [ET, ETu]. Throughout we assume that

$$f,\ \varphi \text{ are bounded below by zero on } K \tag{$A1$}$$

$$\langle f'(x_1)-f'(x_2),\, x_1-x_2\rangle \geq \sigma\, |x_1-x_2|_X^2 \text{ for all } x_1,\ x_2\in K \text{ and } \sigma>0, \tag{$A2$}$$

$$\varphi(\Lambda x_0)<\infty \quad \text{for some } x_0\in K. \tag{$A3$}$$

Note that

$$\begin{aligned} &f(x)-f(x_0)-\langle f'(x_0),\, x-x_0\rangle \\ &\quad = \int_0^1 \langle f'(x_0+t(x-x_0))-f'(x_0),\, x-x_0\rangle\, dt \geq \tfrac{\sigma}{2}\, |x-x_0|^2. \end{aligned}$$

Since φ is proper there exists an element $y_0\in D(\partial\varphi)$ and

$$\varphi(\Lambda x)-\varphi(y_0)\geq (y_0^*, \Lambda x - y_0)_H \quad \text{for } y_0^*\in\partial\varphi(y_0),$$

where $\partial\varphi$ denotes the subdifferential of φ. Hence, $\lim f(x)+\varphi(\Lambda x)\to\infty$ as $|x|_X\to\infty$ and it follows that there exists a unique minimizer $x^*\in K$ for (P).

Theorem 2.1. *The necessary and sufficient condition for $x^*\in K$ to be the minimizer of (P) is given by*

$$\langle f'(x^*),\, x-x^*\rangle + \varphi(\Lambda x)-\varphi(\Lambda x^*)\geq 0 \quad \textit{for all } x\in K. \tag{1}$$

Proofs to the results of this section can be found in [IK1]. Next a Lagrangian associated to the nonsmooth summand φ in (P) will be introduced, while the condition $x\in K$ is kept as explicit constraint. For this purpose we consider

$$f(x)+\varphi_c(\Lambda x,\lambda) \quad \text{over } x\in K, \tag{P_c}$$

where the regularization φ_c of φ is defined as the shifted inf-convolution

$$\varphi_c(y,\lambda)=\inf\,\{\varphi(y-u)+(\lambda,\, u)+\frac{c}{2}\,|u|^2\} \quad \text{over } u\in H, \tag{2}$$

for $y,\ \lambda\in H$ and $c>0$.

Before we return to the necessary optimality condition, properties of the smooth approximation $\varphi_c(x,\lambda)$ to φ are addressed. For $\lambda > 0$ let $J_\lambda = (I + \lambda\,\partial\varphi)^{-1}$ denote the resolvent of $\partial\varphi$ and let

$$A_\lambda x = \lambda^{-1}\,(x - J_\lambda x).$$

stand for the Yosida approxiamtion of $\partial\varphi$.

Theorem 2.2. *For $x,\ \lambda \in H$ the infimum in* (2) *is attained at a unique point $u_c(x,\lambda)$ where $u_c(x,\lambda) = x - J_{1/c}\,(x + c^{-1}\lambda)$. Further $\varphi_c(x,\lambda)$ is convex, (Lipschitz-) continuously Fréchet differentiable in x and $\varphi_c'(x,\lambda) = \lambda + c\,u_c(x,\lambda) = A_{1/c}\,(x + c^{-1}\lambda)$. Moreover,* $\lim_{c\to\infty}\ \varphi_c(x,\lambda) = \varphi(x)$ *and*

$$\varphi(J_{1/c}\,(x + c^{-1}\lambda)) - \frac{1}{2c}\,|\lambda|^2 \le \varphi_c(x,\lambda) \le \varphi(x)$$

for every $x,\ \lambda \in H$.

In the above statement the prime denotes differentiation with respect to the primal variable x.

Theorem 2.3. *For $x,\ \lambda \in H$ we have*

$$\varphi_c(x,\lambda) = \sup_{y^*\in H}\ \{(x,\,y^*) - \varphi^*(y^*) - \frac{1}{2c}\,|y^* - \lambda|^2\}, \tag{3}$$

where the supremum is attained at the unique point $\lambda_c(x,\lambda) = \varphi_c'(x,\lambda)$.

Above φ^* denotes the conjugate of φ defined by

$$\varphi^*(y^*) = \sup_{y\in H}\ \{(y,\,y^*) - \varphi(y)\} \quad \text{for } y^* \in H.$$

Remark 2.1. If $\varphi = I_{\{y=0\}}$, where and I_S is the indicator function of a set S:

$$I_S(x) = \begin{cases} 0 & \text{if } x \in S \\ \\ \infty & \text{if } x \notin S \end{cases},$$

then $\varphi_c(y,\lambda) = (\lambda, y) + \frac{c}{2}|y|^2$ which is the classical augmented Lagrangian functional associated to equality constraints, [Be, IK1].

Remark 2.2. In many applications the conjugate function φ^* is given by

$$\varphi^*(v) = I_{C^*}(v),$$

where C^* is a closed convex set in H. In this case it follows from Theorem 2.3 that for $v,\ \lambda \in H$

$$\varphi_c(v,\lambda) = \sup_{y^*\in C^*}\ \{-\frac{1}{2c}|y^* - (\lambda + c\,v)|_H^2\} + \frac{1}{2c}\,(|\lambda + c\,v|_H^2 - |\lambda|_H^2). \tag{4}$$

Hence the supremum is attained at $\lambda_c(v,\lambda) = \mathrm{Proj}_{C^*}(\lambda + c\,v)$ where $\mathrm{Proj}_{C^*}(\phi)$ denotes the projection of $\phi \in H$ onto C^*.

The following theorem provides an equivalent characterization of $\lambda \in \partial\varphi(x)$.

Theorem 2.4. *If $\lambda \in \partial\varphi(y)$ for $y,\ \lambda \in H$, then $\lambda = \varphi_c'(y,\lambda)$ for all $c > 0$. Conversely, if $\lambda = \varphi_c'(y,\lambda)$ for some $c > 0$, then $\lambda \in \partial\varphi(y)$.*

We return to (P_c). Since $x \to \varphi_c(\Lambda x, \lambda)$ is bounded from below by $-\frac{1}{2c}\,|\lambda|_H^2$, the regularized problems (P_c) admit a unique solution $x_c \in K$. The necessary and sufficient optimality condition is given by

$$\langle f'(x_c),\, x - x_c\rangle + (\varphi_c'(\Lambda x_c, \lambda),\, \Lambda(x - x_c))_H \geq 0 \quad \text{for all } x \in K. \tag{5}$$

Theorem 2.5. *Assume that there exist $\lambda_c^* \in \partial\varphi(\Lambda x_c)$ for $c \geq 1$ such that $\{|\lambda_c^*|_H\}_{c\geq 1}$ is bounded. Then, x_c converges strongly to x^* in X as $c \to \infty$ and for each weak cluster point λ^* of $\{\lambda_c\}_{c\geq 1}$ in H*

$$\lambda^* \in \partial\varphi(\Lambda x^*) \quad \textit{and} \quad \langle f'(x^*),\, x - x^*\rangle + (\lambda^*,\, \Lambda(x - x^*))_H \geq 0 \textit{ for all } x \in K. \tag{6}$$

Conversely, if $x^* \in K$ satisfies (6) then x^* solves (P).

The following lemma addresses the assumption of Theorem 2.5.

Lemma 2.1. *(1) If $dom(\varphi) = H$ then $\partial\varphi(\Lambda x_c)$ is non-empty and $|\partial\varphi(\Lambda x_c)|_H$ is uniformly bounded for $c \geq 1$.*
(2) If $\varphi = \chi_C$ with C a closed convex set in H and $\Lambda x_c \in C$ for all $c > 0$, then λ_c^* can be chosen to be 0 for all $c > 0$.

Theorem 2.6. *Assume that there exists a pair $(x^*, \lambda^*) \in K \times H$ that satisfies (6). Then the complementarity condition $\lambda^* \in \partial\varphi(\Lambda x^*)$ can equivalently be expressed as*

$$\lambda^* = \varphi_c'(\Lambda x^*, \lambda^*) \tag{7}$$

and x^* is the unique solution of

$$\min\ f(x) + \varphi_c(\Lambda x, \lambda^*) \quad \textit{over } x \in K \tag{8}$$

for every $c > 0$.

Note that (7) follows directly from Theorem 2.4. The importance of Theorem 2.6 is given by the fact that the complementarity condition in the form of a differential inclusion is replaced by a nonlinear equation, which is preferable for computations. In the case of Remark (2.2), $\varphi_c'(\Lambda x, \lambda)$ is a projection.

We turn to the discussion of the first order augmented Lagrangian method. Problem (P) is equivalent to

$$\begin{cases} \min f(x) + \varphi(\Lambda x - u) \\ \text{subject to } x \in K \quad \text{and} \quad u = 0 \text{ in } H. \end{cases} \tag{9}$$

To treat the constraint $u = 0$ in (9) by the augmented Lagrangian method we consider the sequential minimization over $x \in K$ and $u \in H$ of the form

$$\min\ f(x) + \varphi(\Lambda x - u) + (\lambda,\, u)_H + \frac{c}{2}\,|u|_H^2, \tag{10}$$

where $\lambda \in H$ is a multiplier and c is a positive scalar penalty parameter [Be, IK1]. Equivalently (10) can be expressed as

$$\min\ L_c(x,\lambda) = f(x) + \varphi_c(\Lambda x, \lambda) \quad \text{over } x \in K, \tag{11}$$

where $\varphi_c(v,\lambda)$ is defined in (2). The (first-order) augmented Lagrangian method is given next:

Augmented Lagrangian Method

(i) Choose a starting value $\lambda_1 \in H$, a positive number c and set $k = 1$.

(ii) Given $\lambda_k \in H$ determine $x_k \in K$ from

$$L_c(x_k, \lambda_k) = \min\ L_c(x, \lambda_k) \quad \text{over } x \in K.$$

(iii) Update λ_k by $\lambda_{k+1} = \varphi_c'(\Lambda x_k, \lambda_k)$.

(iv) If the convergence criterion is not satisfied then set $k = k+1$ and go to (ii).

The following theorem asserts unconditional convergence with respect to c of the augmented Lagrangian method.

Theorem 2.7. *Assume that there exists $\lambda^* \in \partial\varphi(\Lambda x^*)$ such that (6) is satisfied. Then the sequence (x_k, λ_k) is well-defined and satisfies*

$$\frac{\sigma}{2}\,|x_k - x^*|_X^2 + \frac{1}{2c}\,|\lambda_{k+1} - \lambda^*|_H^2 \le \frac{1}{2c}\,|\lambda_k - \lambda^*|_H^2, \tag{12}$$

and

$$\sum_{k=1}^{\infty} \frac{\sigma}{2}\,|x_k - x^*|_X^2 \le \frac{1}{2c}\,|\lambda_1 - \lambda^*|_H^2, \tag{13}$$

which implies that $|x_k - x^*|_X \to 0$ as $k \to \infty$.

Example 2.1 (Obstacle problem). We consider the problem

$$\begin{cases} \min \quad \int_\Omega (\frac{1}{2}\,|\nabla u|^2 - \tilde f\, u)\,dx \quad \text{over } u \in H_0^1(\Omega) \\ \text{subject to} \quad \phi \le u \le \psi \ \text{ a.e. in } \Omega, \end{cases} \tag{14}$$

with $\tilde f \in L^2(\Omega)$ and ϕ, ψ given obstacles. In the context of the general framework we choose $X = H_0^1(\Omega)$, $H = L^2(\Omega)$ and $\Lambda =$ the natural injection, and define $f : X \to \mathbb{R}$ and $\varphi : H \to \mathbb{R}$ by

$$f(u) = \int_\Omega (|\nabla u|^2 - \tilde f\, u)\,dx \quad \text{and} \quad \varphi(v) = I_C,$$

where $C \subset H$ is the closed convex set defined by $C = \{v \in H : \phi \le v \le \psi \text{ a.e. in } \Omega\}$. For one the sided constraint $u \le \psi$ (i.e., $\phi = -\infty$) it is shown from the literature, see e.g. [GLT, IK3], that there exists a unique $\lambda^* \in \partial\varphi(u^*)$ such that (6) is satisfied provided that $\psi \in H^1(\Omega)$, $\psi|_\Gamma \ge 0$ and $\sup(0, \tilde f + \Delta\psi) \in L^2(\Omega)$. Let us set $C_\psi = \{v \in H : v \le \psi \text{ a.e. in } \Omega\}$. Then we have $I^*_{C_\psi}(v) = (\psi, v)$ if $v \ge 0$

a.e. and $I^*_{C_\psi}(v) = \infty$ otherwise. By Theorems 2.2, 2.3, for example, we can argue that $\lambda_c(u,\lambda) = \max(0, \lambda + c(u-\psi))$, where max is the pointwise a.e. operation in Ω. Therefore the optimal pair $(u^*,\lambda^*) \in (H^2 \cap H^1_0) \times L^2$ satisfies

$$\begin{cases} -\Delta u^* + \lambda^* = \tilde{f} \\ \quad \lambda^* = \max(0, \lambda^* + c\,(u^* - \psi)). \end{cases} \tag{15}$$

In this case Steps 2−3 in the augmented Lagrangian method is given by

$$-\Delta u_k + \lambda_{k+1} = \tilde{f}$$

$$\lambda_{k+1} = \max(0, \lambda_k + c\,(u_k - \psi)).$$

For bilateral constraints the existence of a multiplier is much more delicate. We refer to [IK1, IK2, IK3] and assume that $\phi,\ \psi \in H^1(\Omega)$ satisfy

$$\begin{aligned} &\phi \le 0 \le \psi \quad \text{on } \Gamma, \quad \text{and} \quad \max(0, \Delta\psi + \tilde{f}),\ \min(0, \Delta\phi + \tilde{f}) \in L^2(\Omega), \\ &S_1 = \{x \in \Omega : \Delta\psi + \tilde{f} > 0\} \cap S_2 = \{x \in \Omega : \Delta\phi + \tilde{f} < 0\} \text{ is empty}, \\ &-\Delta(\psi - \phi) + c_0\,(\psi - \phi) \ge 0 \quad \text{a.e. in } \Omega \text{ for some } c_0 > 0. \end{aligned}$$

Once existence of a multiplier in $L^2(\Omega)$ guaranteed, see [IK1] p.123, the optimality system we can use Theorem 2.3 and Theorem 2.6 to express the optimality condition can be expressed as

$$-\Delta u^* + \lambda^* = \tilde{f}, \quad \text{with } \lambda^* \in \partial I_C(\Lambda u^*). \tag{16}$$

The latter expression is equivalent to $u^* \in \partial^*_C(\lambda^*)$. By Remark 2.2 and Theorem 2.4, this is equivalent to $u^* = \mathrm{Proj}_C(\lambda^* + cu^*)$, which after some manipulation can be expressed as

$$\lambda^* = \max(0, \lambda^* + c\,(u^* - \psi)) + \min(0, \lambda^* + c(u^* - \phi)).$$

The augmented Lagrangian method for the two-sided constraint can be expressed as

$$-\Delta u_k + \lambda_{k+1} = \tilde{f}, \quad \lambda_{k+1} = \max(0, \lambda_k + c\,(u_k - \psi)) + \min(0, \lambda_k + c\,(u_k - \psi)).$$

Example 2.2 (Bingham fluid and imaging denoising). The simplified Bingham fluid problem is given by

$$\min \int_\Omega (\frac{\lambda}{2}|\nabla u|^2 - \tilde{f}\,u)\,dx + g\int_\Omega |\nabla u|\,dx \quad \text{over } u \in H^1_0(\Omega) \tag{17}$$

where Ω is a bounded open set in R^2 with Lipschitz boundary and $\tilde{f} \in L^2(\Omega)$. In the context of the general theory we choose

$$X = H^1_0(\Omega),\ H = L^2(\Omega) \times L^2(\Omega),\ K = X, \quad \text{and} \quad \Lambda = g\,\mathrm{grad},$$

and define $f : X \to R$ and $\varphi : H \to R$ by

$$f(u) = \frac{1}{g}\int_\Omega (\frac{\lambda}{2}\,|\nabla u|^2 - \tilde{f}\,u)\,dx, \text{ and } \varphi(v_1, v_2) = \int_\Omega \sqrt{v_1^2 + v_2^2}\,dx.$$

Since $\mathrm{dom}(\varphi) = H$ it follows from Theorem 2.5 and Lemma 2.1 that there exists λ^* such that (6) holds. Moreover $\varphi^*(v) = \chi_{C^*}(v)$, where $C^* = \{v \in H : |v(x)|_{\mathbb{R}^2} \leq 1 \text{ a.e. in } \Omega\}$. Hence it follows that the optimality system for (17) is given by

$$\begin{cases} \int_\Omega (\lambda \nabla u^*\,\nabla v - \tilde{f} v)\,dx + \int_\Omega (\lambda^* \nabla u^*)\,dx = 0 \text{ for all } v \in X \\ \lambda^* = \mathrm{Proj}_{C^*}(\lambda^* + c\nabla u^*) = \frac{\lambda^* + c\nabla u^*}{|\lambda^* + c\nabla u^*|_{\mathbb{R}^2}} \; a.e. \in \Omega. \end{cases} \tag{18}$$

Moreover steps (ii)-(iii) in the augmented Lagrangian method are given by

$$-\lambda\,\Delta u_k - g\,\mathrm{div}\,\lambda_{k+1} = \tilde{f}, \tag{19}$$

where

$$\lambda_{k+1} = \begin{cases} \lambda_k + c\,\nabla u_k & \text{on } A_k = \{x : |\lambda_k(x) + c\,\nabla u_k(x)|_{R^2} \leq 1\} \\[2ex] \dfrac{\lambda_k + c\,\nabla u_k}{|\lambda_k + c\,\nabla u_k|} & \text{on } \Omega \setminus A_k. \end{cases} \tag{20}$$

Equation (19) is a nonlinear equation for $u_k \in H_0^1(\Omega)$. The augmented Lagrangian method is thus closely related to the explicit duality (Uzawa-) method, where λ_{k+1} in (19) is replaced by λ_k. The Uzawa method is conditionally convergent in the sense that there exist $0 < \underline{\rho} < \bar{\rho}$ such that it converges for $\rho \in [\underline{\rho}, \bar{\rho}]$, [ET], [GLT]. On the other hand the augmented Lagrangian method converges unconditionally by Theorem 2.7.

The image denoising problem based on BV-regularisation and an additional H^1 semi-norm regularisation term (λ much smaller than g) is given by

$$\min \quad \int_\Omega (\frac{\lambda}{2}\,|\nabla u|^2 + g\,|\nabla u|)\,dx + \frac{1}{2}\int_\Omega |u - z|^2\,dx \quad \text{over } u \in H^1(\Omega), \tag{21}$$

where z denotes the noise corrupted data. It can be treated analogously as the Bingham fuid problem. For a duality based treatment expressing BV-regularized problems as bilateral obstacle problems we refer to [HK1].

In the simplified friction problem the functional φ is given by $\varphi = \int_{\partial\Omega} |u|\,ds$. It also can be treated with the concepts of this section, see [IK4].

3. Semi-Smooth Newton Method in Function Spaces

In the previous section we discussed how equations such as (16) and (18) can be solved by the augmented Lagrangian method. Due to lack of Fréchet differentiability of the involved operations they are not directly amenable for treatment by the Newton algorithm. Therefore in this section we focus on the notion of Newton differentiability.

Let X and Z be Banach spaces and consider the nonlinear equation

$$F(x) = 0\,, \tag{22}$$

where $F\colon D \subset X \to Z$, and D is an open subset of X.

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

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

for every $x \in U$.

We refer to [CNQ, K, HIK] for work related (A). In [CNQ] the term slant differentiability and in [K], for a slightly different notion, the term Newton map were used. Note that it is not required that the mapping G serving as generalized (or Newton) derivative is not required to be unique. The following convergence result is well known [CNQ, HIK].

Theorem 3.1. *Suppose that x^* is a solution to (22) and that F is Newton differentiable in an open neighborhood U containing x^* with Newton derivative $G(x)$. If $G(x)$ is nonsingular for all $x \in U$ and $\{\|G(x)^{-1}\| : x \in U\}$ 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.

Proof. Note that the Newton iterates satisfy

$$|x^{k+1} - x^*| \le |G(x^k)^{-1}|\; |F(x^k) - F(x^*) - G(x^k)(x^k - x^*)|, \tag{23}$$

provided that $x^k \in U$. Let $B(x^*, r)$ denote a ball of radius r centered at x^* contained in U and let M be such that $\|G(x)^{-1}\| \le M$ for all $x \in B(x^*, r)$. We apply (A) with $x = x^*$. Let $\eta \in (0, 1]$ be arbitrary. Then there exists $\rho \in (0, r)$ such that

$$|F(x^* + h) - F(x^*) - G(x^* + h)h| < \frac{\eta}{M}\; |h| \le \frac{1}{M}\; |h| \tag{24}$$

for all $|h| < \rho$. Consequently, if we choose x^0 such that $|x^0 - x^*| < \rho$, then by induction from (23), (24) with $h = x^k - x^*$ we have $|x^{k+1} - x^*| < \rho$ and in particular $x^{k+1} \in B(x^*, \rho)$. It follows that the iterates are well-defined. Moreover, since $\eta \in (0, 1]$ is chosen arbitrarily $x^k \to x^*$ converges superlinearly. □

Here we are especially interested in applications involving the pointwise max operation when X is a function space consisting of elements defined over a bounded domain $\Omega \subset \mathbb{R}^n$ with Lipschitzian boundary $\partial\Omega$. Let $\delta \in \mathbb{R}$ be fixed arbitrarily. We introduce candidates for Newton derivatives G_m of the form

$$G_m(x)(s) = \begin{cases} 1 & \text{if} \quad x(s) > 0\,, \\ 0 & \text{if} \quad x(s) < 0\,, \\ \delta & \text{if} \quad x(s) = 0\,, \end{cases} \tag{25}$$

where $x \in X$.

Proposition 3.1. *(i) G_m can in general not serve as a Newton derivative for* $\max(0,\cdot)\colon\ L^p(\Omega) \to L^p(\Omega)$, *for* $1 \le p \le \infty$.

(ii) The mapping $\max(0,\cdot)\colon L^q(\Omega) \to L^p(\Omega)$ *with* $1 \le p < q \le \infty$ *is Newton differentiable on* $L^q(\Omega)$ *and* G_m *is a Newton derivative.*

For the proof which directly verifies property (A), see [HIK]. Alternatively, if $\psi : \mathbb{R} \to \mathbb{R}$ is semi-smooth in the sense of mappings between finite-dimensional spaces, i.e. ψ is locally Lipschitz continuous and $\lim_{V\in\partial\psi(x+th'),h'\to h,t\to 0^+} Vh'$ exists for all $h \in \mathbb{R}$, then the substitution operator $F : L^q(\Omega) \to L^p(\Omega)$ defined by

$$F(x)(s) = \psi(x(s)) \text{ for } a.e. s \in \Omega$$

is Newton differentiable on $L^q(\Omega)$, if $1 \le p < q \le \infty$, see [U]. In particular this applies to the max operation.

The following chain rule is useful in many applications.

Proposition 3.2. *Let $f : Y \to Z$ and $g :\to Y$ be Newton differentiable in open sets V and U, respectively, with $U \subset X$, $g(U) \subset V \subset Y$. Assume that g is locally Lipschitz continuous and that there exists a Newton map $G_f(\cdot)$ associated to f which is bounded on $g(U)$. Then the superposition $f \circ g : X \to Z$ is Newton differentiable in U with a Newton map $G_f G_g$.*

For the proof we refer to [HK3].

A class of nonlinear complementarity problems: The above concepts are applied to nonlinear complementarity problems of the form

$$g(x) + \lambda = 0, \quad \lambda \ge 0, \;\; x \le \psi \quad \text{and} \quad (\lambda, x - \psi)_{L^2} = 0, \tag{26}$$

where $g : X = L^2(\Omega) \to L^p(\Omega), p > 2$ is Lipschitz continuous and $\psi \in L^p(\Omega)$. If J is a continuously differentiable functional on X then (26) with $g = J'$, is the necessary optimality condition for

$$\min_{x\in L^2(\Omega)} \quad J(x) \quad \text{subject to } x \le \psi. \tag{27}$$

As discussed in the previous section, (26) can equivalently be expressed as

$$g(x) + \lambda = 0, \quad \lambda = \max(0, \lambda + c\,(x - \psi)), \tag{28}$$

for any $c > 0$, where max denotes the pointwise max-operation, with λ the Lagrange multiplier associated to the inequality constraint.

Let us assume that (28) admits a solution $(x^*, \lambda^*) \in L^2(\Omega) \times L^2(\Omega)$. Equation (28) can equivalently be expressed as

$$F(x) = g(x) + \max(0, -g(x) + c\,(x - \psi)) = 0, \tag{29}$$

where F is considered as mapping from X into itself. The semi-smooth Newton iteration for this reduced equation is given by

$$\begin{aligned} g'(x^k)(x^{k+1} - x^k) + G\left(-g(x^k) + c\,(x^k - \psi)\right)\left(-g'(x)(x^{k+1} - x^k) + c\,(x^{k+1} - x^k)\right) \\ +g(x^k) + \max(0, -g(x^k) + c\,(x^k - \psi)) = 0, \end{aligned} \tag{30}$$

where G_m was defined in (25). To investigate local convergence of (30) we denote for any partition $\Omega = \mathcal{A} \cup \mathcal{I}$ into measurable sets $\mathcal{I}$ and $\mathcal{A}$ by $R_{\mathcal{I}} : L^2(\Omega) \to L^2(\mathcal{I})$ the canonical restriction operator and by $R^*_{\mathcal{I}} : L^2(\mathcal{I}) \to L^2(\Omega)$ its adjoint. Further we set

$$g'(x)_{\mathcal{I}} = R_{\mathcal{I}}\, g'(x)\, R^*_{\mathcal{I}}.$$

Proposition 3.3. *Assume that* (28) *admits a solution* x^*, *that* $x \to g(x) - c(x - \psi)$ *is a* C^1 *function from* $L^2(\Omega)$ *to* $L^p(\Omega)$ *in a neighborhood* U *of* x^* *for some* $c > 0$ *and* $p > 2$, *and that*

$$\{g'(x)^{-1}_{\mathcal{I}} \in \mathcal{L}(L^2(\mathcal{I})) : x \in U,\ \Omega = \mathcal{A} \cup \mathcal{I}\} \text{ is uniformly bounded}.$$

Then the iterates x^k *defined by* (30) *converge superlinearly to* x^*, *provided that* $|x^* - x^0|$ *is sufficiently small. Here* x^0 *denotes the initialization of the algorithm.*

Proof. By Propositions 3.1, 3.2 the mapping $x \to \max(0, -g(x) + c(x - \psi))$ is Newton differentiable in U as mapping from $L^2(\Omega)$ into itself and $G_m(-g(x)+c(x-\psi))(-g'(x) + cI)$ is a Newton-derivative. Consequently F is Newton differentiable in U. Moreover $g'(x) + G_m(-g(x) + c(x - \psi))(-g' + cI)$ is invertible in $\mathcal{L}(L^2(\Omega))$ with uniformly bounded inverses for $x \in U$. In fact, setting

$$z = -g(x) + c(x - \psi),\ \mathcal{A} = \{z > 0\},\ \mathcal{I} = \Omega \setminus \mathcal{A},\ h_{\mathcal{I}} = \chi_{\mathcal{I}} h,\ h_{\mathcal{A}} = \chi_{\mathcal{A}} h,$$

this follows from the fact that for given $f \in L^2(\Omega)$ the solution to the equation

$$g'(x)h + G(z)(-g'(x)h + ch) = f$$

is given by

$$ch_{\mathcal{A}} = f_{\mathcal{A}} \text{ and } h_{\mathcal{I}} = g'_{\mathcal{I}}(x)^{-1}(f_{\mathcal{I}} - \frac{1}{c}\chi_{\mathcal{I}}\, g'(x) f_{\mathcal{A}}).$$

From Theorem 3.1 we conclude that $x^k \to x^*$ superlinearly, provided that $|x^* - x^0|$ is sufficiently small. ☐

It can be observed that the semi-smooth Newton step can be equivalently expressed as

$$
\begin{aligned}
&g'(x^k)(x^{k+1}-x^k)+g(x^k)+\lambda^{k+1}=0 \\
&x^{k+1}=\psi \text{ in } \mathcal{A}_k=\{s:-g(x^k)(s)+c\,(x^k(s)-\psi(s))>0\} \\
&\lambda^{k+1}=0 \text{ in } \mathcal{I}_k=\{s:-g(x^k)(s)+c\,(x^k(s)-\psi(s))\le 0\}.
\end{aligned} \tag{31}
$$

Remark 3.1. We refer to (31) as the primal-dual active set strategy for the reduced equation. If the semi-smooth Newton step is applied to (28) rather than to the reduced equation, then the resulting algorithm differs in the update of the active/inactive sets. In fact, in this case the update for the active set is given by $\mathcal{A}_k=\{s:\lambda^k(s)+c\,(x^k(s)-\psi(s))>0\}=\{s:-g(x^{k-1})(s)-g'(x^{k-1})(x^k-x^{k-1})(s)+c\,(x^k(s)-\psi(s))>0\}$. In case g is linear the two updates coincide.

If we consider regularized least squares problems of the form

$$
\min J(x)=\frac{1}{2}|Tx-z|_Y^2+\frac{\alpha}{2}|x|_{L^2}^2, \text{ subject to } x\le\psi, \tag{32}
$$

where Y is a Hilbert space, $T\in\mathcal{L}(L^2(\Omega),Y)$, $\alpha>0$ and $z\in Y,\psi\in L^p(\Omega)$, then $g(x)=T^*(Tx-z)+\alpha x$ and $g(x)-\alpha(x-\psi)=T^*(Tx-z)+\alpha\psi$. Hence Proposition 3.3 with $c=\alpha$ is applicable if $T^*\in\mathcal{L}(Y,L^p(\Omega))$, for some $p>2$. The optimality condition (29) is given by

$$
\alpha(x-\psi)+\max(0,T^*(Tx-z)+\alpha\psi)=0, \tag{33}
$$

in this case.

So far we addressed local convergence. The following result gives a sufficient condition for global convergence.

Proposition 3.4. *Consider* (32) *and assume that* $\|T\|^2_{\mathcal{L}(L^2(\Omega_1),L^2(\Omega_2))}<\alpha$. *Then the semi-smooth Newton algorithm converges independently of the initialisation to the unique solution of* (32).

The proof is based an argument using

$$
M(x,\lambda)=\alpha^2\int_\Omega|(x-\psi)^+|^2\,ds+\int_{\mathcal{A}(x)}|\lambda^-|^2\,ds
$$

as a merit function, where $\mathcal{A}(x)=\{s:x(s)\ge\psi(s)\}$. It decays along the iterates (x^k,λ^k) of the semi-smooth Newton algorithm. An analogous result can be obtained in case of bilateral constraints and for nonlinear mappings g, if additional requirements are met, [IK3].

Propositions 3.3 and 3.4 are applicable to optimal control problems with control constraints, for example. This is the contents of the following section.

4. Optimal Dirichlet Boundary Control

Let us consider the Dirichlet boundary optimal control problem with point-wise constraints on the boundary, formally given by

$$\begin{cases} \min \frac{1}{2}|y-z|^2_{L^2(Q)} + \frac{\alpha}{2}|u|^2_{L^2(\Sigma)} \\ \text{subject to} \\ \partial_t y - \kappa\Delta y + b\cdot\nabla y = f \quad \text{in} \quad Q \\ y = u, \quad u \le \psi \quad \text{on} \quad \Sigma \\ y(0) = y_0 \quad \text{in } \Omega, \end{cases} \tag{34}$$

where $Q = (0,T]\times\Omega\,,\quad \Sigma = (0,T]\times\partial\Omega$ and Ω a bounded domain in $\mathbb{R}^n$, $n \ge 2$ with C^2 boundary $\partial\Omega$. This guarantees that the Laplacian with homogenous Dirichlet boundary conditions, denoted by Δ, is an isomorphism form $H^2(\Omega)\cap H^1_0(\Omega)$ to $L^2(\Omega)$. Further $\kappa > 0$, $y_0 \in H^{-1}(\Omega)$, $z \in L^2(Q)$, $f \in L^2(H^{-2}(\Omega))$, $u \in L^2(\Sigma)$ and $b \in \mathbb{L}^\infty(Q)$, div $b \in L^\infty(L^{\hat n}(\Omega))$ where $\hat n = \max(n,3)$, and $\mathbb{L}^\infty(Q) = \bigotimes_{i=1}^n L^\infty(Q)$.

Under these conditions there exists a unique very weak solution $y \in L^2(Q)\cap H^1(H^{-2}(\Omega))\cap C(H^{-1}(\Omega))$ satisfying for a.e. $t\in(0,T)$

$$\begin{cases} \langle \partial_t y(t), v\rangle - \kappa(y(t),\Delta v) - (y(t), \operatorname{div}(b(t))\,v) - (y(t), b(t)\nabla v) \\ \qquad\qquad = \langle f(t), v\rangle - \kappa(u(t), \frac{\partial}{\partial n}v)_{\partial\Omega} \text{ for all } v\in H^2(\Omega)\cap H^1_0(\Omega), \\ y(0) = y_0, \end{cases} \tag{35}$$

where $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_{H^{-2}(\Omega),H^2(\Omega)\cap H^1_0(\Omega)}$ denotes the canonical duality pairing, $(\cdot,\cdot)$ and $(\cdot,\cdot)_{\partial\Omega}$ stand for the inner products in $L^2(\Omega)$ and $L^2(\partial\Omega)$ respectively. Moreover

$$|y|_{L^2(Q)\cap H^1(H^{-2}(\Omega))\cap C(H^{-1}(\Omega))} \le C(|y_0|_{H^{-1}(\Omega)} + |f|_{L^2(H^{-2}(\Omega))} + |u|_{L^2(\Sigma)}), \tag{36}$$

where C depends continuously on $\kappa > 0$, $|b|_{L^\infty(Q)}$ and $|\text{div } b|_{L^\infty(L^{\hat n}(\Omega))}$, and is independent of $f \in L^2(H^{-2}(\Omega))$, $u\in L^2(\Sigma)$ and $y_0\in H^{-1}(\Omega)$.

Utilizing the a-priori bound (36) it is straightforward to argue the existence of a unique solution $u^*\in L^2(\Sigma)$ of (34). It can be shown that it is characterized by the optimality system

$$\begin{cases} \partial_t y - \kappa\Delta y + b\cdot\nabla y = f \text{ in } Q, \\ y = u \text{ on } \Sigma,\ y(0) = y_0 \text{ in } \Omega, \\ -\partial_t p - \kappa\Delta p - \operatorname{div} b\, p - b\cdot\nabla p = -(y-z) \text{ in } Q, \\ p = 0 \text{ on } \Sigma\,, \quad p(T) = 0 \text{ in } \Omega, \\ \kappa\frac{\partial p}{\partial n} + \alpha u + \lambda = 0 \text{ on } \Sigma, \\ \lambda = \max(0, \lambda + c(u-\psi)) \text{ on } \Sigma, \end{cases} \tag{37}$$

where the primal must be interpreted in the very weak form. In terms of (32) we have that the operator $T : L^2(\Sigma) \to L^2(Q)$ is given as the control to state operator for (35). Its adjoint $T^* \in \mathcal{L}(L^2(Q), L^2(\Sigma))$ is the solution of the adjoint equation, i.e. the third and forth equations in (37), with right hand side $\varphi \in L^2(Q)$. In [KV] we verified that the adjoint satisfies

$$\left\| \frac{\partial p}{\partial n} \right\|_{L^{q_n}(\Sigma)} \leq C_1 \, \|p\|_{L^2(H^2(\Omega)) \cap H^1(L^2(\Omega))} \leq C_2 \|\varphi\|_{L^2(Q)},$$

with an embedding constant C_1, C_2, where

$$q_n = \begin{cases} \frac{2(n+1)}{n}, & \text{if } n \geq 3\,, \\ 3 - \varepsilon, & \text{if } n = 2, \end{cases}$$

for every $\varepsilon > 0$, so that in particular $q_n > 2$ for every n. Equation (33) is given by

$$\alpha u - \psi + \max(0, \kappa \frac{\partial p}{\partial n} + \alpha \psi) = 0, \tag{38}$$

in this case, and Propositions 3.3 and 3.4 imply that the semi-smooth Newton method applied to (38) converges locally superlinearly, as well as globally, if $\alpha > \|T\|^2_{\mathcal{L}(L^2(\Sigma), L^2(\Omega))}$.

5. Sparse Controls

The control cost in optimal control problems is most frequently chosen to be of the form $\frac{\alpha}{2}|u|^2$, where u denotes the control. In this way the control cost is differentiable, in some applications the term can be interpreted as energy. It is indispensable in the stochastic interpretation of the linear quadratic regulator theory. However, it also has drawbacks, most notably, it does not put proportional weight on the control. The purpose of this section is to sketch a framework for the use of $\alpha|u|$ as control cost. For this choice the cost of the control is proportional to its "size". Moreover it has the feature of being sparse. To get an appreciation for this latter property let us consider the non-differentiable problem in $L^2(\Omega)$ given by

$$\min \frac{1}{2}|u - z|^2_{L^2} + \alpha |u|_{L^1}. \tag{39}$$

The solution to (39) is given in the a.e. sense by

$$u^* = \begin{cases} 0 & \text{if } |z| < \alpha^{-1} \\ z - \alpha^{-1} \operatorname{sgn} z & \text{if } |z| \geq \alpha^{-1}. \end{cases} \tag{40}$$

In particular the solution is 0 where z is small relative to $1/\alpha$. The space of $L^1(\Omega)$-controls, however, does not lend itself to weak* compactness arguments which are needed to guarantee existence in the context of optimal control. Consequently the control space is enlarged to measure-valued controls. We consider the model problem

$$\begin{cases} \min\limits_{u \in \mathcal{M}} \dfrac{1}{2}|y - z|_{L^2}^2 + \alpha |u|_{\mathcal{M}} \\ \text{s.t.} \quad Ay = u, \end{cases} \tag{$\mathcal{P}_{\mathcal{M}}$}$$

where $\mathcal{M}$ denotes the vector space of all bounded Borel measures on Ω, that is the space of all bounded σ-additive set functions $\mu : \mathcal{B}(\Omega) \to \mathbb{R}$ defined on the Borel algebra $\mathcal{B}(\Omega)$ satisfying $\mu(\emptyset) = 0$. The total variation of $\mu \in \mathcal{M}$ is defined for all $B \in \mathcal{B}(\Omega)$ by $|\mu|(B) := \sup\{\sum_{i=0}^{\infty} |\mu(B_i)| : \bigcup_{i=0}^{\infty} B_i = B\}$, where the supremum is taken over all partitions of B. Endowed with the norm $|\mu|_{\mathcal{M}} = |\mu|(\Omega)$, $\mathcal{M}$ is a Banach space. By the Riesz representation theorem, $\mathcal{M}$ can be isometrically identified with the topological dual of $C_0(\Omega)$. This leads to the following equivalent characterization of the norm on $\mathcal{M}$:

$$|\mu|_{\mathcal{M}} = \sup_{\substack{\phi \in C_0(\Omega), \\ |\phi|_{C_0} \leq 1}} \int_{\Omega} \phi \, d\mu. \tag{41}$$

Further A is a second order elliptic operator with homogenous Dirichlet boundary conditions in the bounded domain $\Omega \subset \mathbb{R}^n$ with $n \in \{2, 3\}$, and such that

$$\|A \cdot\|_{L^2} \text{ and } \|A^* \cdot\|_{L^2} \text{ are equivalent norms on } H^2(\Omega) \cap H_0^1(\Omega),$$

where A^* denotes the adjoint of A with respect to the inner product in L^2. For $u \in \mathcal{M}$, the equation $Ay = u$ has a unique weak solution $y \in W_0^{1,p}(\Omega)$, for all $1 \leq p < \frac{n}{n-1}$. Furthermore, there exists a constant $C > 0$ such that $|y|_{W_0^{1,p}} \leq C|u|_{\mathcal{M}}$.

Since $W_0^{1,p}(\Omega)$ is compactly embedded in $L^2(\Omega)$, $(\mathcal{P}_{\mathcal{M}})$ is well-defined, and standard arguments imply the existence of a unique solution (y^*, u^*). Next we aim for a formulation of the problem that is appropriate for computational purposes. By Fenchel duality theory the predual to $(\mathcal{P}_{\mathcal{M}})$ is given by

$$\begin{cases} \min\limits_{p \in H^2 \cap H_0^1} \dfrac{1}{2}|A^* p + z|^2 - \dfrac{1}{2}|z|_{L^2}^2 \\ \text{s.t.} \quad |p|_{C_0} \leq \alpha, \end{cases} \tag{$\mathcal{P}_{\mathcal{M}}^*$}$$

which can be considered as a bilaterally constraint problem. Existence of a unique solution p^* can readily be verified and the relationship between solutions to the original and the predual problem are given by:

$$\begin{cases} Ay^* = u^*, \\ A^* p^* = z - y^*, \\ \quad 0 \leq \langle u^*, p^* - p \rangle_{(H^2 \cap H_0^1)^*, H^2 \cap H_0^1} \text{ for all } p \in H^2 \cap H_0^1, \ |p|_{C_0} \leq \alpha. \end{cases} \tag{42}$$

The inequality in (42) can be interpreted as the larger α, the smaller is the support of the control u^*.

While $(\mathcal{P}^*_{\mathcal{M}})$ is of bilateral constraint type, some further consideration is required before Newton methods can be used efficiently. Comparing to (32) and the optimal control problem treated in Section 4, the operator appearing in $(\mathcal{P}^*_{\mathcal{M}})$ is not of smoothing type. Note that if we were to discretize (32) and $(\mathcal{P}^*_{\mathcal{M}})$ then these problems have the same structure. But this is not the case on the continuous level. Computationally this becomes apparent in the context of mesh independence. Applying the semi-smooth Newton method to the discretized form of (32) with T satisfying $T^* \in \mathcal{L}(Y, L^p(\Omega))$ will result in mesh-independent iteration numbers of the semi-smooth Newton method, while this is not the case for $(\mathcal{P}^*_{\mathcal{M}})$. For an analysis of mesh-independence of the semi-smooth Newton method we refer to [HU].

To obtain a formulation which is appropriate for a super-linear and mesh-independent behavior of the semi-smooth Newton method some type of regularization is required. For example an additional regularization term of the form $\frac{\beta}{2}|u|^2_{L^2}$ can be added to the cost in $(\mathcal{P}_{\mathcal{M}})$, see e.g. [St]. Here we go a different way and consider the Moreau-Yosida approximation, see (15), of the inequality constraints leading to

$$\min_{p \in H^2 \cap H^1_0} \frac{1}{2}|A^* p + z|^2_{L^2} - \frac{1}{2}|z|^2_{L^2} + \frac{c}{2}|\max(0, p - \alpha)|^2_{L^2} + \frac{c}{2}|\min(0, p + \alpha)|^2_{L^2}, \quad (\mathcal{P}^*_{\mathcal{M},c})$$

where the max- and min- operations are taken pointwise in Ω. For $c > 0$ let p_c denote the solutions to $(\mathcal{P}^*_{\mathcal{M},c})$. They satisfy the optimality system

$$\begin{cases} AA^* p_c + Az + \lambda_c = 0, \\ \lambda_c = \max(0, c(p_c - \alpha)) + \min(0, c(p_c + \alpha)), \end{cases} \tag{43}$$

where $\lambda_c \in W^{1,\infty}$ approximates the Lagrange multiplier associated to the constraint $|p|_{C_0} \le \alpha$. Let $(p^*, \lambda^*) \in H^2 \cap H^1_0 \times (H^2 \cap H^1_0)^*$ denote the unique solution to the optimality system for $(\mathcal{P}^*_{\mathcal{M}})$:

$$\begin{cases} AA^* p^* + Az + \lambda^* = 0, \\ \langle \lambda^*, p - p^* \rangle \le 0, \end{cases} \tag{44}$$

for all $p \in H^2 \cap H^1_0$ with $|p|_{C_0} \le \alpha$. Then, see [CK], as $c \to \infty$:

$$p_c \quad \to p^* \quad \text{in } H^2 \cap H^1_0, \qquad \lambda_c \rightharpoonup \quad \lambda^* \quad \text{in } (H^2 \cap H^1_0)^*. \tag{45}$$

The regularized optimality system $(\mathcal{P}^*_{\mathcal{M},c})$ can be solved efficiently by the semi-smooth Newton method with G_m as in (25) and appropriately adapted for the min term. For this purpose we express $(\mathcal{P}^*_{\mathcal{M},c})$ as a nonlinear equation $F(p) = 0$ with $F : H^2 \cap H^1_0 \to (H^2 \cap H^1_0)^*$, where

$$F(p) := AA^* p + \max(0, c(p - \alpha)) + \min(0, c(p + \alpha)) + Az. \tag{46}$$

Due to the regularity gap between the domain and the range of F the following result can be obtained quite readily from Theorem 3.1, and Proposition 3.1, [CK].

Theorem 5.1. *If $|p_c - p^0|_{H^2 \cap H^1_0}$ is sufficiently small, the iterates p^k of the semi-smooth Newton algorithm converge superlinearly in $H^2 \cap H^1_0$ to the solution p_c of $(\mathcal{P}^*_{\mathcal{M},c})$ as $k \to \infty$.*

Algorithm 1 Semismooth Newton method for (43)

Set $k = 0$, Choose $p^0 \in H^2 \cap H_0^1$

repeat

  Set

$$\mathcal{A}_{k+1}^{+} = \{x\}\, p^k(x) > \alpha, \quad \mathcal{A}_{k+1}^{-} \quad = \{x\}\, p^k(x) < -\alpha, \quad \mathcal{A}_{k+1} = \mathcal{A}_{k+1}^{+} \cup \mathcal{A}_{k+1}^{-}$$

  Solve for $p^{k+1} \in H^2 \cap H_0^1$:

$$(A^* p^{k+1}, A^* v)_{L^2} + c(p^{k+1} \chi_{\mathcal{A}_k}, v)_{L^2} = -(z, A^* v)_{L^2} + c\alpha(\chi_{\mathcal{A}_k^+} - \chi_{\mathcal{A}_k^-}, v)_{L^2}$$

  for all $v \in H^2 \cap H_0^1$

  Set $k = k + 1$

until $(\mathcal{A}_{k+1}^{+} = \mathcal{A}_k^{+})$ and $(\mathcal{A}_{k+1}^{-} = \mathcal{A}_k^{-})$

For this application, let us give the algorithm in detail in Algorithm 1. The stopping criterion is typically met without any need for globalization. If it applies then the algorithm stops at the solution of (43). For actual computations a discretisation of the infinite dimensional spaces is required. This is not within the scope of this paper.

The question also arises how to choose c in practice. Large c implies that we can be close to the solution of the unregularized problem at the expense of *possible* ill-conditioning of the regularized one. We have only rarely experienced that ill-conditioning actually occurs. In practice it is certainly advisable to utilize a continuation principle, applying the Algorithm with a moderate value for c, and utilizing the solution thus obtained as initialization for a computation with a larger value for c. This procedure can be put onto solid ground for unilateral constraints by means of path following techniques as detailed in [HK2]. Since the infinite dimensional problem always needs to be discretized a natural stopping criterion for the increase of c is given once the error due to regularization is smaller than that of discretization. For certain obstacle type problems which satisfy a maximum principle the L^∞ error due to regularization can be estimated, see [IK5]. – Concerning regularization let us stress that discretization also has a regularizing effect. In this case staggered grid strategies applied to the original unregularized formulation, i.e. $(\mathcal{P}_{\mathcal{M}}^*)$ in our case, correspond to the increase of the regularisation parameter c and can be very effective in numerical computations. The formal analysis of this procedure has not been carried out yet.

6. Time Optimal Control

This section is devoted to time optimal control problems for a class of nonlinear ordinary differential equations. The techniques are applicable to much wider class of problems, but the detailed analysis yet needs to be carried out. While the computation of time optimal controls and trajectories has a long history, the use of Newton-type methods is a very recent one. We refer to [IK4] for a detailed description of the procedure that we describe here.

$$\begin{cases} \min_{\tau\geq 0}\int_0^\tau dt \quad \text{subject to} \\ \frac{d}{dt}x(t) = Ax(t) + Bu(t),\ |u(t)|_{\ell^\infty} \leq 1,\ x(0) = x_0,\ x(\tau) = x_1, \end{cases} \tag{P_T}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $x_0 \neq x_1$ are given vectors in $\mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, u is measurable, and $|\cdot|_{\ell^\infty}$ denotes the infinity-norm on $\mathbb{R}^m$. It is assumed that x_1 can be reached in finite time by an admissible control. Then (P_T) admits a solution with optimal time denoted by τ^*, and associated optimal state x^* and optimal control u^*.

It is wellknown that under appropriate conditions [HL] the optimal solution is related to the adjoint equation

$$p(t) = \exp\left(A^T(\tau^* - t)\right) q, \ \text{with } q \in \mathbb{R}^n,$$

through

$$u^*(t) = -\sigma(B^T p(t)) = -\sigma\big(B^T \exp\left(-A^T(\tau^* - t)\right) q\big), \tag{47}$$

for $t \in [0, \tau^*]$, where $q \in \mathbb{R}^n$ and σ denotes the coordinate-wise operation

$$\sigma(s) \ \in \ \begin{cases} \text{sgn } s & \text{if } s \neq 0 \\ [-1,\, 1] & \text{if } s = 0. \end{cases} \tag{48}$$

This operation prohibits the use of superlinear Newton-type methods for solving (P_T) numerically. Therefore a family of regularized problems given by

$$\begin{cases} \min_{\tau\geq 0}\int_0^\tau (1 + \frac{\varepsilon}{2}\,|u(t)|^2)\,dt \quad \text{subject to} \\ \frac{d}{dt}x(t) = Ax(t) + Bu(t),\ |u(t)|_{\ell^\infty} \leq 1,\ x(0) = x_0,\ x(\tau) = x_1, \end{cases} \tag{P_ε}$$

with $\varepsilon > 0$ is considered. The norm $|\cdot|$ used in the cost-functional denotes the Euclidean norm. It is straightforward to argue the existence of a solution $(u_\varepsilon, x_\varepsilon, \tau_\varepsilon)$. Convergence of the solutions $(x_\varepsilon, u_\varepsilon, \tau_\varepsilon)$ of (P_ε) to a solution $(x^*,\, u^*,\, \tau^*)$ of (P_T) was analysed in [IK4]. Note that τ^* is unique.

Proposition 6.1. *For every $0 < \varepsilon_0 < \varepsilon_1$ and any solution (τ^*, u^*) of (P) we have*

$$\tau^* \leq \tau_{\varepsilon_0} \leq \tau_{\varepsilon_1} \leq \tau^*(1 + \frac{\varepsilon_1}{2}), \tag{49}$$

$$|u_{\varepsilon_1}|_{L^2(0,\,\tau_{\varepsilon_1})} \leq |u_{\varepsilon_0}|_{L^2(0,\,\tau_{\varepsilon_0})} \leq |u^*|_{L^2(0,\,\tau^*)}. \tag{50}$$

If u^* is a bang-bang solution, then

$$0 \leq |u^*|^2_{L^2(0,\,\tau^*)} - |u_\varepsilon|^2_{L^2(0,\,\tau_\varepsilon)} \leq \ \textit{meas}\ \{t \in [0, \tau^*] : |u_\varepsilon(t)| < 1\} \tag{51}$$

for every $\varepsilon > 0$. Moreover, if (A, B_i) is controllable for each column B_i of B, then the solution u^* is unique, it is bang-bang and $u_\varepsilon \to u^*$ in L^2 as $\varepsilon \to 0^+$.

Recall that a control is called bang-bang if $|u_i(t)| = 1$ for all $t \in [0, \tau^*]$ and $i = 1, \ldots, m$. Concerning a necessary optimality condition for (P_ε) we have the following result:

Theorem 6.1. *Let $(x_\varepsilon, u_\varepsilon, \tau_\varepsilon)$ be a solution of (P_ε). Assume that there exists some $\tilde{i}$ such that*

$$(A, B_{\tilde{i}}) \text{ is controllable,} \tag{H1}$$

and such that exist $\eta > 0$ and an interval $I_{\tilde{i}} \subset (0, 1)$ satisfying

$$|\,(\hat{u}_\varepsilon)_{\tilde{i}}(t)|_{\ell^\infty} \le 1 - \eta \quad \text{for a.e. } t \in I_{\tilde{i}}. \tag{H2}$$

Then there exists an adjoint state p_ε such that

$$\begin{cases} \dot{x}_\varepsilon = Ax_\varepsilon + Bu_\varepsilon,\ x_\varepsilon(0) = x_0,\ x_\varepsilon(\tau_\varepsilon) = x_1, \\ -\dot{p}_\varepsilon = A^T p_\varepsilon, \\ u_\varepsilon = -\sigma_\varepsilon(B^T p_\varepsilon), \\ 1 + \frac{\epsilon}{2}|u_\varepsilon(\tau_\varepsilon)|^2_{\mathbb{R}^m} + p_\varepsilon(\tau_\varepsilon)^T(Ax_\varepsilon(\tau_\varepsilon) + Bu_\varepsilon(\tau_\varepsilon)) = 0, \end{cases} \tag{52}$$

where

$$\sigma_\varepsilon(s) \ \in \ \begin{cases} sgn\ s & \text{if } s \le -\varepsilon \\ \frac{s}{\varepsilon} & \text{if } |s| < \ \varepsilon. \end{cases} \tag{53}$$

System (52) can readily be treated by a semi-smooth Newton method. In a first step the method of mappings is used to transform the system to a fixed time domain. The transformation $t \to \frac{t}{\tau}$ transforms (52) to

$$\begin{cases} \dot{x} = \tau(Ax + Bu),\ x(0) = x_0,\ x(1) = x_1, \\ -\dot{p} = \tau A^T p, \\ u = -\sigma_\varepsilon(B^T p), \\ 1 + \frac{\epsilon}{2}|u(1)|^2 + p(1)^T(Ax(1) + Bu(1)) = 0. \end{cases} \tag{54}$$

To investigate the semi-smooth Newton method we require an additional assumption

$$|B_i^T\, p_\varepsilon(1)| \neq \varepsilon, \text{ for all } i = 1, \dots, m, \tag{H3}$$

where we now fix ε and a solution $(x_\varepsilon,\, u_\varepsilon,\, \tau_\varepsilon) \in W^{1,2}(0,1) \times L^2(0,1) \times \mathbb{R}$ of (P_ε) with associated adjoint $p_\varepsilon \in W^{1,2}(0,\,1)$. With (H2) and (H3) holding there exists a neighborhood $\mathcal{U}_{p_\varepsilon}$ of p_ε in $W^{1,2}(0,1;\mathbb{R}^n)$, $\bar{t} \in (0,1)$, and a nontrivial interval $(\alpha, \alpha+\delta) \subset (0,1)$ such that for $p \in \mathcal{U}_{p_\varepsilon}$ we have

$$|B_i^T\, p(t)| \neq \varepsilon \text{ for all } t \in [\bar{t},\, 1], \text{ and } i = 1, \dots, m, \text{ and } |B_{\tilde{i}}^T\, p(t)| < \varepsilon \text{ for } t \in (\alpha, \alpha+\delta).$$

Equation (54) suggests to introduce

$$F(x,\, p,\, u,\, \tau) = \begin{pmatrix} \dot{x} - \tau Ax - \tau Bu \\ -\dot{p} - \tau A^T p \\ u + \sigma_\varepsilon(B^T p) \\ x(1) - x_1 \\ 1 + \frac{\epsilon}{2}|u(1)|^2 + p(1)^T\,(Ax(1) + Bu(1)) \end{pmatrix}. \tag{55}$$

where

$$F : D_F \subset X \to L^2(0,1;\mathbb{R}^n) \times L^2(0,1;\mathbb{R}^n) \times U \times \mathbb{R}^n \times \mathbb{R},$$

and

$$D_F = W^{1,2}(0,1) \times \mathcal{U}_{p_\varepsilon} \times U \times \mathbb{R}, \quad X = W^{1,2}(0,1;\mathbb{R}^n) \times W^{1,2}(0,1;\mathbb{R}^n) \times U \times \mathbb{R}.$$

Here we have set $U = \{u \in L^2(0,\,1;\mathbb{R}^m) : u|[\bar{t},\,1] \in W^{1,2}(\bar{t},1;\mathbb{R}^m)\}$ endowed with the norm $|u|_U = (|u|^2_{L^2(0,1)} + |\dot{u}|^2_{L^2(\bar{t},1)})^{\frac{1}{2}}$. The only equation that requires special attention in (55) is the third one which contains the operator σ_ε. We use

$$G\sigma_\varepsilon(s) := \begin{cases} \frac{1}{\varepsilon} & \text{if } |s| < \varepsilon \\ 0 & \text{if } |s| \geq \varepsilon \end{cases} \tag{56}$$

as generalized derivative in the sense of Definition 3.1 for σ_ε. It is now straightforward to argue that F is Newton differentiable. To apply Theorem 3.1 it remains to argue that the inverse of the Newton derivative of F is uniformly bounded in a neighborhood of $(x_\varepsilon,\, u_\varepsilon,\, \tau_\varepsilon, p_\varepsilon)$. For this purpose the Newton system is considered for reduced unknowns $(p(1), \tau)^T \in \mathbb{R}^{n+1}$. In terms of these variables the system matrix becomes:

$$\mathcal{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{pmatrix},$$

where

$$A_{11} = \varepsilon^{-1}\tau \int_0^1 e^{\tau A(1-t)} B\chi_I B^T\, e^{\tau A^T(1-t)}\, dt \in \mathbb{R}^{n\times n}, \tag{57}$$

$$\begin{aligned} A_{12} = \varepsilon^{-1}\tau \textstyle\int_0^1 e^{\tau A(1-t)} B\chi_I B^T \int_t^1\, e^{-\tau A^T(t-s)} A^T p(s)\, ds\, dt \\ - \textstyle\int_0^1 e^{\tau A(1-t)}(Ax + Bu)\, dt \in \mathbb{R}^n, \end{aligned} \tag{58}$$

$$A_{21} = (Ax(1) + Bu(1))^T - (p^T(1)B + \varepsilon u^T(1))\, G\sigma_\varepsilon(B^T p(1))B^T \in (\mathbb{R}^n)^T, \tag{59}$$

with $\chi_I = diag(\chi_{I_1}, \ldots, \chi_{I_m})$ and χ_{I_i} the characteristic function of the set

$$I_i = I_i(p) = \{t : |(B^T p)_i| < \varepsilon\}, \quad i = 1, \ldots, m$$

which is nonempty for $p \in \mathcal{U}_{p_\varepsilon}$ and $i = \tilde{i}$. The controllability assumption (H1) together with (H2) imply that the symmetric matrix A_{11} is invertible with bounded inverse uniformly with respect to $p \in \mathcal{U}_{p_\varepsilon}$ and τ in compact subsets of $(0, \infty)$.

To guarantee uniform boundedness of the inverse of the Newton derivative we require that the Schur complement $A_{21}\, A_{11}^{-1}\, A_{12} \in \mathbb{R}$ of $\mathcal{A}$ for $(x,\, p,\, u,\, \tau)$ in a neighborhood of $(x_\varepsilon,\, p_\varepsilon,\, u_\varepsilon,\, \tau_\varepsilon)$ is nontrivial. We therefore assume that

$$\begin{cases} \text{there exists a bounded neighborhood} \\ \mathcal{U} \subset D_F \subset X \text{ of } (x_\varepsilon,\, p_\varepsilon,\, u_\varepsilon,\, \tau_\varepsilon) \text{ and } c > 0 \text{ such that} \\ |A_{21}\, A_{11}^{-1}\, A_{12}| \geq c \text{ for all } (x,\, p,\, u,\, \tau) \in \mathcal{U}. \end{cases} \tag{H4}$$

Theorem 6.2. *If* (H1)–(H4) *hold and* $(x_\varepsilon,\, u_\varepsilon,\, \tau_\varepsilon)$ *denotes a solution to* (P_ε) *with associated adjoint* p_ε*, then the semi-smooth Newton algorithm converges superlinearly, provided that the initialization is sufficiently close to* $(x_\varepsilon, p_\varepsilon,\, u_\varepsilon,\, \tau_\varepsilon)$.

7. L^1-Data Fitting

Here we treat the data fitting problem with robust $L^1(\Omega)$ fit-to-data term and consider

$$\min_{x\in L^2}\left\{\mathcal{J}_\alpha(x)\equiv|Kx-y^\delta|_{L^1}+\frac{\alpha}{2}|x|^2\right\},\tag{$\mathcal{P}_{L^1}$}$$

where $K: L^2(\Omega)\to L^2(\Omega)$ is a compact linear operator, and $y^\delta\in L^2$ are measurements corrupted by noise. For every α there exists a unique minimizer x_α. For the value function

$$F(\alpha)=|Kx_\alpha-y^\delta|_{L^1}+\frac{\alpha}{2}|x_\alpha|^2,$$

it can be shown that

$$F'(\alpha)=\frac{1}{2}|x_\alpha|^2,\tag{60}$$

[CJK, IK6]. Fenchel duality theory implies that the dual to ($\mathcal{P}_{L^1}$) is given by

$$\begin{cases}\min\limits_{p\in L^2}\dfrac{1}{2\alpha}|K^*p|^2_{L^2}-(p,y^\delta)_{L^2}\\ \text{s.t.}\quad |p|_{L^\infty}\le 1.\end{cases}.\tag{$\mathcal{P}^*_{L^1}$}$$

The dual problem has at least one solution p_α and the relationship between x_α and p_α is given by

$$K^*p_\alpha=\alpha x_\alpha,\quad 0\le(Kx_\alpha-y^\delta,p-p_\alpha)_{L^2},\ \text{ for all } p\in L^2 \text{ with } |p|_{L^\infty}\le 1.\tag{61}$$

Problem ($\mathcal{P}^*_{L^1}$) does not lend itself to treatment with a superlinearly convergent semi-smooth Newton algorithm. In fact the optimality system for ($\mathcal{P}^*_{L^1}$) is given by

$$\frac{1}{\alpha}KK^*p_\alpha-y^\delta+\lambda_\alpha=0,\quad(\lambda_\alpha,p-p_\alpha)_{L^2}\le 0,\ \text{ for all } |p|_{L^\infty}\le 1,\tag{62}$$

where λ_α denotes the Lagrange multiplier associated to the inequality constraint. This system does not admit a reformulation such that Theorem 3.1(ii) is applicable. We therefore introduce the family of regularized problems

$$\begin{cases}\min\limits_{p\in H^1}\dfrac{1}{2\alpha}|K^*p|^2_{L^2}+\dfrac{\beta}{2}|\nabla p|^2_{L^2}-(p,y^\delta)\\ \text{s.t.}\quad |p|_{L^\infty}\le 1,\end{cases}\tag{$\mathcal{P}^*_\beta$}$$

for $\beta>0$, and finally for the numerical realisation the Moreau-Yosida regularization of the box constraints:

$$\min_{p\in H^1}\frac{1}{2\alpha}|K^*p|^2_{L^2}+\frac{\beta}{2}|\nabla p|^2_{L^2}-(p,y^\delta)\\ +\frac{1}{2c}|\max(0,c(p-1))|^2_{L^2}+\frac{1}{2c}|\min(0,c(p+1))|^2_{L^2},\tag{$\mathcal{P}^*_{\beta,c}$}$$

for $c>0$. It is assumed that $\ker K^*\cap\ker\nabla=\emptyset$. Then ($\mathcal{P}^*_\beta$) and ($\mathcal{P}^*_{\beta,c}$) admit unique solutions in H^1 denoted by p_β and p_c respectively. At the end of this section we comment on the choice of the regularization parameters.

The optimality system for ($\mathcal{P}^*_{\beta,c}$) is given by

$$\begin{cases}\dfrac{1}{\alpha}KK^*p_c-\beta\Delta p_c+\lambda_c=y^\delta,\\ \lambda_c=\max(0,c(p_c-1))+\min(0,c(p_c+1)),\end{cases}\tag{63}$$

where $\lambda_c \in H^1(\Omega)$. It can be shown by techniques which are by now quite standard [IK5, CJK] that for each fixed $\beta > 0$ we have

$$(p_c, \lambda_c) \to (p_\beta, \lambda_\beta) \text{ in } H^1(\Omega) \times H^1(\Omega)^*,$$

where $\lambda_\beta \in H^1(\Omega)^*$ is the Lagrange multiplier associated to the inequality constraint in $(\mathcal{P}_\beta^*)$. Moreover, for every sequence $\beta_n \to 0$ there exists a subsequence such that $p_{\beta_k} \rightharpoonup p_\alpha$ in $L^2(\Omega)$, where p_α is a solution of $(\mathcal{P}_{L^1}^*)$. Analogously, if c is fixed then the solutions to $(\mathcal{P}_{\beta,c}^*)$, now denoted by $p_{\beta_n,c}$ converge to a solution of $p_{0,c}$ of $(\mathcal{P}_{\beta,c}^*)$ with $\beta = 0$.

To solve the optimality system (63) for the regularized problem we consider the nonlinear operator equation $F(p) = 0$ for $F : H^1(\Omega) \to H^1(\Omega)^*$, where

$$F(p) := \frac{1}{\alpha} KK^* p - \beta \Delta p + \max(0, c(p-1)) + \min(0, c(p+1)) - y^\delta. \tag{64}$$

In view of Section 3 we use as Newton derivative for the projection operator $P(p) := \max(0, (p-1)) + \min(0, (p+1))$ the mapping

$$G_P(p)h := h\chi_{\{|p|>1\}} = \begin{cases} h(x) & \text{if } |p(x)| > 1, \\ 0 & \text{if } |p(x)| \le 1. \end{cases}$$

It can readily be verified that the update $p^{k+1} \in H^1(\Omega)$ of the Newton equation $G_P(p^k)(p^{k+1} - p^k) = -F(p^k)$ is the solution to the equation

$$\frac{1}{\alpha} KK^* p^{k+1} - \beta \Delta p^{k+1} + c\chi_{\mathcal{A}_k} p^{k+1} = y^\delta + c(\chi_{\mathcal{A}_k^+} - \chi_{\mathcal{A}_k^-}), \tag{65}$$

where the active sets are given by

$$\mathcal{A}_k^+ := \{x\}\, p^k(x) > 1, \quad \mathcal{A}_k^- := \{x\}\, p^k(x) < -1, \quad \mathcal{A}_k := \mathcal{A}_k^+ \cup \mathcal{A}_k^-.$$

Moreover we can use the techniques of Section 3 to establish the following result.

Theorem 7.1. *If $|p_c - p^0|_{H^1}$ is sufficiently small, then the iterates p^k of the semi-smooth Newton algorithm converge superlinearly in $H^1(\Omega)$ to the solution p_c of $(\mathcal{P}_{\beta,c}^*)$ as $k \to \infty$.*

We turn to a discussion of the choice of the parameters α, β and c in problem $(\mathcal{P}_{\beta,c}^*)$. Clearly β and c, which are used in the inner loop of an iterative procedure, should be taken close to 0 and ∞, respectively. The choice of α, which is different from 0 in general, is the most delicate one and we turn to it first.

Choice of α by model function approach: The model function approach proposed in [IK6] approximates the value function $F(\alpha)$ by rational polynomials. Here we consider a model function of the form

$$m(\alpha) = b + \frac{d}{t + \alpha}. \tag{66}$$

Noting that $x_\alpha \to 0$ for $\alpha \to \infty$ and $\alpha |x_\alpha|^2 \to 0$ by (61), we fix $b = |y^\delta|_{L^1}$. The parameters d and t are determined by interpolation conditions according to

$$m(\alpha) = F(\alpha), \quad m'(\alpha) = F'(\alpha), \tag{67}$$

which together with the definition of $m(\alpha)$ gives

$$b + \frac{d}{t+\alpha} = F(\alpha), \quad -\frac{d}{(t+\alpha)^2} = F'(\alpha). \tag{68}$$

Algorithm 2 Fixed-point algorithm for adaptively determining α

1: Set $k = 0$, choose $\alpha_0 > 0$, $b \geq |y|_{L^1}$ and $\sigma > 1$
2: **repeat**
3: Compute x_{α_k} by a path-following semismooth Newton method
4: Compute $F(\alpha_k)$ and $F'(\alpha_k)$
5: Construct the model function $m_k(\alpha) = b + \frac{d_k}{t_k+\alpha}$ by solving the interpolation condition at α_k

$$d_k = -\frac{(b - F(\alpha_k))^2}{F'(\alpha_k)}, \quad t_k = \frac{b - F(\alpha_k)}{F'(\alpha_k)} - \alpha_k.$$

6: Calculate the m-intercept $\hat{m}$ of the tangent of $m_k(\alpha)$ at $(\alpha_k, F(\alpha_k))$ by

$$\hat{m} = F(\alpha_k) - \alpha_k F'(\alpha_k),$$

7: Solve for α_{k+1} by setting $m_k(\alpha_{k+1}) = \sigma\hat{m}$, i.e. $\alpha_{k+1} = \frac{c_k}{\sigma\hat{-}b} - t_k$
8: Set $k = k + 1$
9: **until** the stopping criterion is satisfied.

We recall from (60) that $F'(\alpha) = -\frac{1}{2}|x_\alpha|^2_{L^2}$, and this expression can be calculated without any extra computational effort. Note that $F(\alpha)$, just like $m(\alpha)$, is monotonically increasing. In case the L^1 fit-to-data term is replaced by an L^2 term, then $F''(\alpha) = -(x_\alpha, (\alpha I + K^*K)^{-1}x_\alpha) \leq 0$. In particular, in this case, F is concave, just as m. One of the important features of our approach lies in not requiring knowledge of the noise level. The rationale for noise level estimation is that $F(0)$ represents a lower bound on the noise level and consequently, if m approximates well F, then $m(0)$ can be taken as an approximation of the noise level.

To analyse the sequence $\{\alpha_k\}$ determined by Algorithm (2) one can argue [CJK] that if this sequence converges then its limit α^* satisfies

$$(\sigma - 1)\phi(\alpha^*) - \alpha^* F'(\alpha^*) = 0, \tag{69}$$

where $\varphi(\alpha) = |Kx_\alpha - x^\delta|_{L^1}$. The intuitive interpretation of the iteration is clear: it balances the weighted data-fitting term $(\sigma - 1)\phi(\alpha) = (\sigma - 1)|Kx_\alpha - y^\delta|_{L^1}$ and the penalty term $\alpha F'(\alpha) = \frac{\alpha}{2}|x_\alpha|^2_{L^2}$. The scalar σ controls the relative weighting between the two terms.

From [CJK] we now quote the following result.

Theorem 7.2. *(a) If σ is sufficiently close to* 1 *and $y^\delta \neq 0$, then* (69) *has at least one solution. (b) If in addition $\alpha_0 F'(\alpha_0) - (\sigma - 1)\varphi(\alpha_0) > 0$, then the iterates $\{\alpha_k\}$ converge monotonically from above to a solution of* (69).

Choice of β within a path-following semismooth Newton method: The introduction of the H^1 smoothing alters the structure of the problem and therefore the value of β should be as small as possible. However, the regularized dual problem $(\mathcal{P}^*_{\beta,c})$ becomes increasingly ill-conditioned as β decreases to zero due to the ill-conditioning of discretized KK^* and rank-deficiency of the diagonal matrix corresponding to the (discrete) active set, see (65). Therefore, the respective system matrix will eventually become numerically singular for vanishing β.

One remedy is a continuation strategy: Starting with a large β, e.g. β_0=1, we reduce its value, e.g. geometrically, as long as the system is still solvable, and take the solution corresponding to the smallest such value. The question remains how to automatically select the stopping index without a priori knowledge or expensive computations for estimating the condition number or smallest singular value by e.g. singular value decomposition. To select an appropriate stopping index, we exploit the structure of the (infinite-dimensional) bound constraint problem: the correct solution should be nearly feasible for c sufficiently large, i.e. $\|p\|_{L^\infty} \leq \tau$ for some $\tau \approx 1$. Recall that for the linear system (65), the right hand side f satisfies $\|f\|_{L^\infty} \approx c \gg 1$, which should be balanced by the diagonal matrix $c\chi_{\mathcal{A}}$ in order to verify the feasibility condition. If the matrix is nearly singular, this will no longer be the case, and the solution p blows up and violates the feasibility condition, i.e. $\|p\|_{L^\infty} \gg 1$. Once this happens, we take the last iterate which is still (close to) feasible and return it as the solution. This procedure provides an efficient and simple strategy to achieve the conflicting goals of minimizing the effect on the primal problem and maintaining the numerical stability of the dual problem $(\mathcal{P}^*_{\beta,c})$ for sufficient accuracy.

For the choice of c it appears to be worthwhile to also investigate path-following techniques as introduced in [HK2] but this remains to be done in future work.

8. Mathematical Programming

In this section we discuss a nonsmooth mathematical programming problem, which only in part relies on convexity assumptions. Let X be a Banach space, Y a Hilbert space and Z a Hilbert lattice with an ordering induced by a cone K with vertex at 0, i.e. $x \leq y$ if $x - y \in K$. Consider the minimization

$$\min \quad F(y) \quad \text{subject to } G_1(y) = 0, \;\; G_2(y) \leq 0, \quad y \in \mathcal{C}, \tag{70}$$

where $G_1 : X \to Y$ is C^1, $G_2 : X \to Z$ is convex, and $\mathcal{C} \subset X$ is a closed convex set. We assume that $F = F_0(y) + F_1(y)$ where F_0 is C^1 and $F_1(y)$ is convex. Then we have the following necessary optimality condition.

Theorem 8.1. *Let $y^* \in \mathcal{C}$ is a minimizer of (70). Then there exists a nontrivial $(\lambda_0, \mu_1, \mu_2) \in \mathbb{R}^+ \times Y^* \times Z^*$ such that*

$$\begin{aligned} &\lambda_0 \left(F_0'(y - y^*) + F_1(y) - F_1(y^*)\right) + \langle \mu_1, G_1'(y^*)(y - y^*)\rangle + \langle \mu_2, G_2(y) - G_2(y^*)\rangle \geq 0 \\ &\mu_2 \geq 0, \quad \langle \mu_2, G_2(y^*)\rangle = 0, \;\; \textit{for all admissible } y \in \mathcal{C}. \end{aligned} \tag{71}$$

Proof. For $\epsilon > 0$ define the functional

$$J_\epsilon(u, \tau) = \left(((F(y) - F(y^*) + \epsilon)^+)^2 + |G_1(y)|_Y^2 + |\max(0, G_2(y))|_Z^2\right)^{\frac{1}{2}}.$$

Then, $J_\epsilon(y^*) = \epsilon$ and $J_\epsilon(y^*) \leq \inf\, J_\epsilon + \epsilon$. For any $y \in \mathcal{C}$ define the metric

$$d(y, y^*) = |y - y^*|_X.$$

By the Ekeland variational principle there exists a y^ϵ such that

$$\begin{aligned} &J_\epsilon(y^\epsilon) \le J_\epsilon(y^*) \\ &J_\epsilon(y) - J_\epsilon(y^\epsilon) \ge -\sqrt{\epsilon}\, d(y, y^\epsilon) \quad \text{ for all } y \in \mathcal{C} \\ &d(y^\epsilon, y^*) \le \sqrt{\epsilon}. \end{aligned} \tag{72}$$

Let

$$\mu_1^\epsilon = 2\, G_1(y^\epsilon), \quad \mu_2^\epsilon = 2\, \max(0, G_2(y^\epsilon)).$$

Letting $y = y_t = y^\epsilon + t\,(\hat{y} - y^\epsilon),\ t \in (0,1)$ with $\hat{y} \in \mathcal{C}$ in (72), we have

$$\begin{aligned} -\sqrt{\epsilon}\, d(y_t, y^\epsilon) \le J_\epsilon(y_t) - J_\epsilon(y^\epsilon) \le \frac{1}{J_\epsilon(y^\epsilon) + J_\epsilon(y_t)}\,(\alpha^{\epsilon,t}\,(F(y_t) - F(y^\epsilon)) \\ +\langle \mu_1^\epsilon, G_1'(y^\epsilon)(y_t - y^\epsilon)\rangle + \langle \mu_2^\epsilon, G_2(y_t) - G_2(y^\epsilon)\rangle + |\mu^\epsilon| o(|y_t - y^\epsilon|)), \end{aligned} \tag{73}$$

where

$$\alpha^{\epsilon,t} = ((F(y_t) - F(y^*) + \epsilon)^+ + (F(y^\epsilon) - F(y^*) + \epsilon)^+).$$

and we used for $t > 0$ sufficiently small

$$(F(y_t) - F(y^*) + \epsilon)(F(y^\epsilon) - F(y^*) + \epsilon) \ge 0.$$

Since F_1 and G_2 are convex,

$$\begin{aligned} &F(y_t) - F(y^\epsilon) \le t\, F_0'(y^\epsilon)(\hat{y} - y^\epsilon) + t\,(F_2(\hat{y}) - F_2(y^\epsilon)) + o(|y_t - y^\epsilon|) \\ &G_2(y_t) - G_2(y^\epsilon) \le t\,(G_2(\hat{y}) - G_2(y^\epsilon)). \end{aligned}$$

Let

$$\tilde{\mu}^{t,\epsilon} = \frac{\mu^\epsilon}{J_\epsilon(y_t) + J_\epsilon(y^\epsilon)}, \quad \tilde{\alpha}^{\epsilon,t} = \frac{\alpha^{\epsilon,t}}{J_\epsilon(y_t) + J_\epsilon(y^\epsilon)}.$$

Since $(\tilde{\alpha}^{\epsilon,t}, \tilde{\mu}^{\epsilon,t})$ is bounded, there exits a subsequence such that $\tilde{\mu}^{\epsilon,t} \rightharpoonup \mu \in (Y \times Z)^*$ (weakly star) and $\tilde{\alpha}^{\epsilon,t} \to \lambda^0 \ge 0$ as $\epsilon \to 0^+,\ t \to 0^+$. Dividing (73) by t and letting $t \to 0^+$ and subsequently $\epsilon \to 0^+$, we obtain

$$\lambda_0\,(F_0'(\hat{y} - y^*) + F_1(\hat{y}) - F_1(y^*)) + \langle \mu_1, G_1'(y^*)(\hat{y} - y^*)\rangle + \langle \mu_2, G(\hat{y}) - G(y^*)\rangle \ge 0,$$

for all $\hat{y} \in \mathcal{C}$. Since $\tilde{\mu}_2^\epsilon \ge 0$ and $\langle \tilde{\mu}_2^\epsilon, G_2(y^\epsilon)\rangle \ge 0$, it follows that $\mu_2 \ge 0$ and $\langle \mu_2, G_2(y^*)\rangle \ge 0$ and since $G_2(y^*) \le 0$, thus $\langle \mu_2, G_2(y^*)\rangle = 0$. ☐

Corollary 8.1. *Assume there exists a nontrivial* $(\lambda_0, \mu_1, \mu_2) \in \mathbb{R}^+ \times Y^* \times Z^*$ *such that* (71) *holds. If the regular point condition:*

$$0 \in \mathit{int} \left\{ \begin{array}{l} G_1'(y^*)(\mathcal{C} - y^*) \\ G_2(y) - G_2(y^*) - K + G_2(y^*) \end{array} \right\}. \tag{74}$$

is satisfied at y^**, then one can take* $\lambda_0 = 1$.

Proof. As a consequence of the regular point condition, there exists for all $(\tilde{\mu}_1, \tilde{\mu}_2)$ belonging to a neighborhood of 0 in $Y \times Z$, elements $y \in \mathcal{C},\ k \in K$ such that

$$(\tilde{\mu}_1, \tilde{\mu}_2) = (G_1'(y^*)(y - y^*), G_2(y) - G_2(y^*) - k + G_2(y^*)).$$

Consequently $\langle \mu_1, \tilde{\mu}_1\rangle + \langle \mu_2, \tilde{\mu}_2\rangle = \langle \mu_1, G_1'(y^*)(y-y^*)\rangle + \langle \mu_2, G_2(y) - G_2(y^*) - k + G_2(y^*)\rangle$. Note that $\langle \mu_2, k - G_2(y^*)\rangle = \langle \mu_2, k\rangle \le 0$. If $\lambda_0 = 0$ then the first equation in (71) implies that $\langle \mu_1, \tilde{\mu}_1\rangle + \langle \mu_2, \tilde{\mu}_2\rangle \ge 0$ for all $(\tilde{\mu}_1, \tilde{\mu}_2)$ in a neighborhood of 0 and thus $\mu_1 = \mu_2 = 0$, which is a contradiction. That is, $\lambda_0 \ne 0$ and thus the problem is strictly normal and one can set $\lambda^0 = 1$. □

L^1-minimum norm control: Consider the optimal exit problem with minimum L^1 norm

$$\left\{\begin{array}{l} \min_{u,\tau} \quad \int_0^\tau (f(x(t)) + \delta |u(t)|)\, dt \quad \text{ subject to} \\[1ex] \quad \frac{d}{dt} x = b(x(t), u(t)), \quad x(0) = x, \\[1ex] \quad g(x(\tau) = 0, \quad |u(t)|_{\mathbb{R}^m} \le \gamma \text{ for a.e. } t, \end{array}\right. \tag{75}$$

where $\delta > 0$, $f : \mathbb{R}^n \to \mathbb{R}$, $b : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^k$ are smooth functions. We have two motivations to consider (75). In the context of sparse controls, compare Section 5, the pointwise norm constraints, allow us avoid controls in measure space. In the context of time optimal controls the term $\delta \int_0^T |u|\, dt$ can be considered as regularisation term. We shall see from the optimality condition (82) below that this determines the control as a function of the adjoint by mean of an equation rather than an inclusion as in (47), where no such regularisation was used.

One can transform (75) to the fixed interval $s \in [0,1]$ via the change of variable $t = \tau\, s$

$$\left\{\begin{array}{l} \min_{u,\tau} \quad \int_0^1 \tau\, (f(x(t)) + \delta |u(t)|)\, dt \quad \text{ subject to} \\[1ex] \quad \frac{d}{dt} x = \tau\, b(x(t), u(t)), \quad x(0) = x, \\[1ex] \quad g(x(1)) = 0, \quad u \in U_{ad} = \{u \in L^\infty(0,1;\mathbb{R}^m) : |u(t)| \le \gamma\}. \end{array}\right. \tag{76}$$

Let $y = (u, \tau)$ and define

$$\begin{array}{l} F_0(y) = \tau \displaystyle\int_0^1 f(x(t))\, dt, \quad F_1(u) = \delta \int_0^1 |u(t)|\, dt, \\ F(y) = F_0(y) + \tau\, F_1(u), \quad G(y) = g(x(1)), \end{array}$$

where $x = x(\cdot; u, \tau)$ is the solution to the initial value problem in (76), given $u \in U_{ad}$ and $\tau \ge 0$. Then the control problem can equivalently formulated as

$$\min_{(u,\tau) \in U_{ad} \times \mathbb{R}^+} \quad F(y) \quad \text{subject to } G(y) \in K. \tag{77}$$

Assume that $y^* = (u^*, \tau^*)$ is an optimal solution to (77) and suppose that the regular point condition

$$0 \in int\, \{G_u(y^*)(v - u^*) + G_\tau(\tau - \tau^*) : v \in U_{ad}, \tau > 0\} \tag{78}$$

holds. Since $\tau_1\, F_1(u_1) - \tau_2 F_1(u_2) = (\tau_1 - \tau_2)\, F_1(u_1) + \tau_2(F_1(u_1) - F_1(u_2))$, it is easy to modify the proof of Theorem 8.1 to obtain the necessary optimality: there exist a Lagrange multiplier $\mu \in \mathbb{R}^k$ such that

$$\begin{array}{l} \tau^*\, (F_1(u) - F_1(u^*)) + (\tau - \tau^*)\, F_1(u^*) \\[1ex] \qquad + ((F_0)_u + G_u^* \mu)(u - u^*) + ((F_0)_\tau + G_\tau^* \mu)(\tau - \tau^*) \ge 0 \end{array} \tag{79}$$

for all $u \in U_{ad}$ and $\tau \geq 0$, where $F_0 = F_0(y^*), G_0 = G_0(y^*)$. Note that for $v \in L^\infty(0,1;\mathbb{R}^m))$

$$G_u(v) = (g_x(x^*(1)), h(1))_{\mathbb{R}^n}, \quad G_\tau = (g_x(x^*(1)), \xi(1))_{\mathbb{R}^n},$$

$$(F_0)_u(v) = \tau \int_0^1 (f'(x(t)), h(t))_{\mathbb{R}^n}\, dt, \quad (F_0)_\tau(v) = \int_0 \big((\tau\, f'(x(t)), \xi(t))_{\mathbb{R}^n} + f(x(t))\big)\, dt,$$

where (h, ξ) satisfies

$$\begin{aligned} &\frac{d}{dt}h(t) = \tau^*\, (b_x(x^*(t), u^*(t))h(t) + b_u(x^*(t), u^*(t))v(t)), \quad h(0) = 0 \\ &\frac{d}{dt}\xi(t) = \tau^*\, b_x(x^*(t), u^*(t))\xi(t) + b(x^*(t), u^*(t), \quad \xi(0) = 0. \end{aligned} \tag{80}$$

Let $p \in H^1(0,1;\mathbb{R}^n)$ satisfy the adjoint equation

$$-\frac{d}{dt}p(t) = \tau^*\, (b_x(x^*(t), u^*(t))^t p(t) + f_x(x^*(t))), \quad p(1) = \mu g_x(x^*(1)), \tag{81}$$

then

$$\begin{aligned} (h(1), p(1))_{\mathbb{R}^n} &= \tau^* \int_0^1 \big(b_u(x^*(t), u^*(t))v(t) - (f'(x^*(t)), h(t))\big)\, dt \\ (\xi(1), p(1))_{\mathbb{R}^n} &= \int_0^1 \big(b(x^*(t), u^*(t), p(t))_{\mathbb{R}^n}\big)\, dt. \end{aligned}$$

From (79) therefore for all $u \in U_{ad}$ and $\tau \geq 0$

$$\begin{aligned} &(\tau - \tau^*) \int_0^1 (f(x^*(t)) + \delta|u^*(t)| + (b(x^*(t), u^*(t)), p(t)))\, dt \\ &\quad + \int_0^1 (b_u(x^*(t), u^*(t))^t p(t), u(t) - u^*(t)) + \delta|u(t)| - \delta|u^*(t)|)\, dt \geq 0. \end{aligned}$$

Hence we obtain the optimality condition

$$u^*(t) = \begin{cases} 0 & \text{if } |b_u(x^*(t), u^*(t))^t p(t)| \leq \delta \\ -\gamma\, \dfrac{b_u(x^*(t), u^*(t))^t p(t)}{|b_u(x^*(t), u^*(t))^t p(t)|} & \text{if } |b_u(x^*(t), u^*(t))^t p(t)| \geq \delta, \end{cases} \tag{82}$$

and

$$\int_0^1 \big(f(x^*(t)) + \delta|u^*(t)| + (b(x^*(t), u^*(t)), p(t))_{\mathbb{R}^n}\big)\, dt = 0.$$

This, together with the fact that the Hamiltonian $\mathcal{H}$ is constant along (u^*, x^*, p), implies the transversality condition

$$\mathcal{H}(t) := f(x^*(t)) + \delta|u^*(t)| + (b(x^*(t), u^*(t)), p(t))_{\mathbb{R}^n} = 0 \text{ on } [0,1]. \tag{83}$$

Acknowledgement Many colleagues have joined us over the last years in the development of semi-smooth Newton methods in function spaces for a wide range of different application. We appreciate specifically the collaboration with M. Hintermüller, M. Bergounioux, C. Clason, C. DeLosReyes, R. Herzog, B. Jin, V. Kovtunenko, G. Stadler, and B. Vexler.

References

[Be] D.P. Bertsekas, *Constraint Optimization and Lagrange Multiplier Methods*, Academic Press, Paris, 1982.

[CJK] C. Clason, B. Jin and K. Kunisch: A semismooth Newton method for L^1 data fitting with automatic choice of regularization parameters and noise calibration, to appear in SIAM Journal on Imaging Sciences.

[CK] C. Clason and K. Kunisch: A duality-based approach to elliptic control problems in non-reflexive Banach spaces, ESAIM: COCV, to appear.

[CNQ] X. Chen, Z. Nashed and L. Qi: Smoothing methods and semi-smooth methods for nondifferentiable operator equations, SIAM J. on Numerical Analysis, 38(2000), 1200–1216.

[ET] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*, North Holland, Amsterdam, 1976.

[ETu] I. Ekeland and T. Turnbull, *Infinite Dimensional Optimization and Convexity*, The University of Chicago Press, Chicago, 1983.

[GLT] R. Glowinski, J.L. Lions and R. Tremolieres, *Numerical Analysis of Variational Inequalities*, North Holland, Amsterdam (1981).

[HIK] M. Hintermüller, K. Ito and K. Kunisch: The primal–dual active set strategy as a semi–smooth Newton method, SIAM Journal on Optimization, 13(2002), 865-888. doi:10.1137/S1052623401383558

[HK1] M. Hintermüller and K. Kunisch: Total bounded variation regularization as bilaterally constrained optimization problem, SIAM J. Appl. Mathematics, 64(2004), 1311-1333. doi:10.1137/S0036139903422784

[HK2] 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. doi:10.1137/050637480

[HK3] M. Hintermüller and K. Kunisch: PDE-Constrained optimization subject to pointwise control and zero- or first-order state constraints, SIAM Journal on Optimization, 20 (2009), 1133-1156. doi:10.1137/080737265

[HL] H. Hermes and J. LaSalle: *Functional Analysis and Time Optimal Control*, Academic Presse, 1969.

[HU] M. Hintermüller and M. Ulbrich: Semismooth Newton methods for operator equations in function spaces. SIAM J. Optim. 13(2002), 805–842

[IK1] K. Ito and K.Kunisch *On the Lagrange Multiplier Approach to Variational Problems and Applications*, SIAM, Philadelphia, 2008.

[IK2] K. Ito and K. Kunisch: Augmented Lagrangian methods for nonsmooth convex optimization in Hilbert spaces, Nonlinear Analysis, Theory, Methods and Applications, 41(2000), 591–616. doi:10.1016/S0362-546X(98)00299-5

[IK3] K. Ito and K. Kunisch: The primal-dual active set method for nonlinear optimal control problems with bilateral constraints, SIAM J. on Control and Optimization, 43(2004), 357-376. doi:10.1137/S0363012902411015

[IK4] K. Ito and K. Kunisch: Semi-smooth Newton Methods for time-optimal control for a class of ODEs, to appear in SIAM Journal on Control and Optimization.

[IK5] 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.

[IK6] K. Ito and K. Kunisch: On the choice of the regularization parameter in nonlinear inverse problems, SIAM J. on Optimization, 2(1992), 1–29. doi:10.1137/0802019

[K] B. Kummer: Generalized Newton and NCP methods: convergence, regularity, actions, Discuss. Math. Differ. Incl. Control Optim., 20, (2000), 209-244.

[KV] K. Kunisch and B. Vexler: Constrained Dirichlet boundary control in L^2 for a class of evolution equations, SIAM J. Control and Optimization, 46 (2007), 1726–1753. doi:10.1137/060670110

[St] G. Stadler: Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices, Computational Optimization and Applications, 2007. doi = 10.1007/s10589-007-9150-9

[U] M. Ulbrich: Semi-smooth Newton methods for operator equations in function spaces, SIAM J. on Optimization, 13(2003), 805-841.

K. Ito, Department of Mathematics, North Carolina State University, Raleigh, North Carolina, 27695-8205, USA.
E-mail: kito@unity.ncsu.edu

K. Kunisch, Institute of Mathematics and Scientific Computing, University of Graz, Austria.
E-mail: karl.kunisch@uni-graz.at