Semi-Smooth Newton Methods and their Applications: Part 1

K. Kunisch

Institute for Mathematics and Scientific Computing
University of Graz, Austria
Radon Institute, Austrian Academy of Sciences, Linz, Austria

Woudschoten, October 2017

Collaborators

- M. Bergounioux, Univ. Orleans
- K. Ito, North Carolina State University
- M. Hintermüller, Humboldt University and Weierstraß Institut, Berlin
- C. Clason, University of Duisburg
- F. Kruse, University of Graz

Some preliminary Remarks

- You might not like Newton methods
- You might not like Lagrange multipliers

Collaborators

- M. Bergounioux, Univ. Orleans
- K. Ito, North Carolina State University
- M. Hintermüller, Humboldt University and Weierstraß Institut, Berlin
- C. Clason, University of Duisburg
- F. Kruse, University of Graz

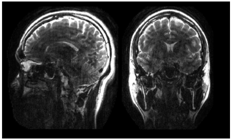
Some preliminary Remarks

- You might not like Newton methods
- You might not like Lagrange multipliers

Optimization of RF pulses (magnetic fields) in Magnetic Resonance Imaging (MRI)

Magnetom Skyra (3T)

[http://www.healthcare.siemens.com]



Goal 1: Reduce the acquisition time in MRI by time-optimal control

Goal 2: Simultaneous multi-slice excitation

Controls: Amplitudes of different magnetic fields

Constraints: Technical limitations of MR hardware, excitation quality

(in cooperation with A. Rund, C. Aigner, R. Stollberger)

Time-optimal control of the Bloch equations

$$\boxed{\min_{u \in U_{\mathrm{ad}}, T>0} T,}$$

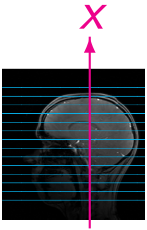
$$\text{s. t.} \quad \dot{M}(t,x) = \gamma \begin{pmatrix} u_1(t) \\ u_2(t) \\ w(t)x \end{pmatrix} \times M(t,x), \qquad \text{in } I \times \Omega,$$

$$M(0,x) = M^0(x), \qquad \text{on } \Omega,$$

$$|M(T,x) - M_d(x)| \leq E(x) \qquad x \in \Omega_{\mathrm{obs}} \subset \Omega,$$

$$\dot{w}(t) = u_3(t), \qquad w(0) = 0, \qquad \text{in } (0,T),$$

$$\int_0^T u_1^2 + u_2^2 \, \mathrm{d}t \leq P_{\max},$$

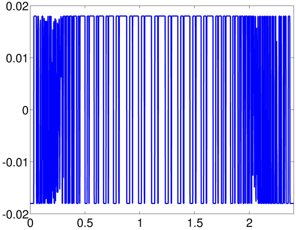


$$u \in U_{\mathrm{ad}} = \{u \in L^2(I)^3 | u_1^2 + u_2^2 \leq R^2, w_a \leq u_3 \leq w_b\}$$

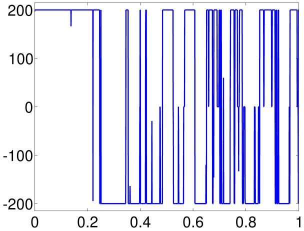
$M(t,x)$ nuclear magnetization

$\Omega = (-L, L),$

Optimal pulses



(a) RF pulse $u_1(t)$



(b) $u_3(t)$

Figure: Results: Time optimal controls $(u_1(t), u_3(t))$ for example "DIFF", $\bar{T} = 2.41,\ u_2(t) \equiv 0$

Revisit: Semi-smooth Newton Method

Definition

$F : D \subset X \to Z$ is called Newton differentiable in $U \subset D$, if there exist $G : U \to \mathcal{L}(X, Z)$:

$$(A)\ lim_{h \to 0} \frac{1}{\mid h \mid} \mid F(x+h) - F(x) - G(x+h)h \mid = 0, \text{ for all } x \in U.$$

Example

$F : L^p(\Omega) \to L^q(\Omega),\ F(\psi) = \max(0, \psi)$ is Newton differentiable if $q < p$, and

$$G_{\max}(\psi)(x) = \begin{cases} 1 & \text{if } \psi(x) > 0 \\ 0 & \text{if } \psi(x) < 0 \\ \delta & \text{if } \psi(x) = 0,\ \delta \in \mathbb{R} \text{ arbitrary.} \end{cases}$$

Example

$p = q$ ... (A) is not satisfied.

Revisit: Semi-smooth Newton Method

Theorem

Let $F(x^*) = 0$, F Newton differentiable in $U(x^*)$, and $\{||G(x)^{-1}||_{\mathcal{L}(X,Z)} : x \in U(x^*)\}$ bounded.

Then the Newton iteration converges locally superlinearly.

Rate of convergence
Mesh independence
Chain Rules

Ref. Kummer, Qi, Chen-Nashed, M. Ulbrich, Hintermüller-Ito-K.

Control of Variational Inequalities

Control of Variational Inequalities

$$\min J(y,u)$$
$$y = \mathrm{argmin}_{y \in H_0^1(\Omega)} \frac{1}{2} a(y,y) - (y,u)_{L^2}$$
$$\text{over } u \in L^2(\Omega) \text{ and } y \in K = \{v \in H_0^1(\Omega) : v \le \psi\}.$$

State Constrained Optimal Control

$$\min J(y,u)$$
$$Ay = u$$
$$\text{over } u \in L^2(\Omega) \text{ and } y \in K$$

Inverse Elastohydrodynamic Problem

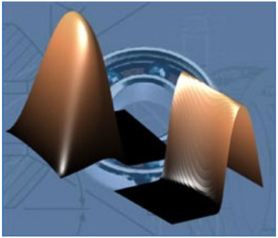
$$(P) \qquad \begin{cases} \min \frac{1}{2}\int_{\tilde{\Omega}} |y - y_{data}|^2\, dx + \frac{\beta}{2}|u|^2 \\ u \in U \\ y = y(u) = argmin\,\{J^u(y)\ :\ y \in K\} \end{cases}$$

$$J^u(y) = \int_\Omega u^3\, |\nabla y|^2\, dx - \int_\Omega \frac{\partial u}{\partial x_2}\, y\, dx$$

$$K = \{y \in H_0^1(\Omega) : y \geq 0\}$$

Reynolds lubrication problem

Inverse Elastohydrodynamic Problem



$$
(P) \qquad \begin{cases} \min \frac{1}{2} \int_{\tilde{\Omega}} |y - y_{data}|^2 \, dx + \frac{\beta}{2} |u|^2 \\ u \in U \\ y = y(u) = argmin \, \{ J^u(y) \, : \, y \in K \} \end{cases}
$$

$$
J^u(y) = \int_\Omega u^3 \, |\nabla y|^2 \, dx - \int_\Omega \frac{\partial u}{\partial x_2} \, y \, dx
$$

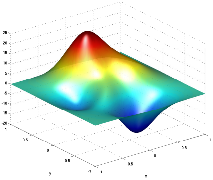
$$
K = \{ y \in H_0^1(\Omega) : y \geq 0 \}
$$

Reynolds lubrication problem

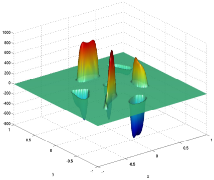
Optimal control with focus on the (non-)smooth cost functional

motivation

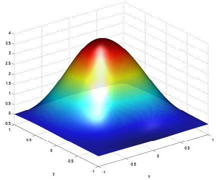
$$\begin{cases} \min\limits_{u \in \mathcal{X}} \|y - z\|^2_{L^2(\Omega)} + \alpha\, \mathcal{N}(\|u\|) \\ \text{s.t.} \quad Ay = u \end{cases} \qquad (\mathcal{P})$$



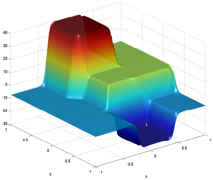
$\int_\Omega |u(x)|^2\, dx$



$\int_\Omega |u(x)|\, dx$



$\int_\Omega |\nabla u(x)|^2\, dx$



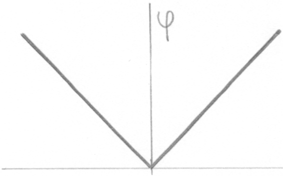
$\int_\Omega |\nabla u(x)|\, dx$

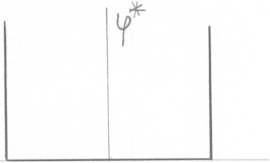
Remark: Sparsity, optimal location problem,...

Dualisation of BV

$$\begin{cases} \text{min} & \frac{1}{2}\int_\Omega |Ku - f|^2 dx + \frac{\alpha}{2}\int_\Omega |u|^2 dx + \beta \int_\Omega |Du| \\ \text{over} & u \in BV, \end{cases}$$

$$\int_\Omega |Du| = \sup\left\{ \int_\Omega u \, div \, \vec{v} : \vec{v} \in (C_0^\infty(\Omega))^2, \, |\vec{v}(x)|_{\ell^\infty} \leq 1 \right\}$$





Portfolio Optimization

$$dX_0(t) = rX_0(t)\,dt - (1+\gamma)dL(t) + (1-\gamma)dM(t)$$

$$dX_1(t) = \mu X_1(t)\,dt + \sigma X_1(t)dW(t) + dL(t) - dM(t)$$

X_0, X_1 wealth processes for bank account /stock

r, μ, σ, γ interest rate, trend, volatility, trading costs

L, M cumulative processes describing purchases/sales of stock

Portfolio Optimization cont.

$$J(t, x_0, x_1)$$
$$= \sup_{L,M} E[\tfrac{1}{\alpha}(X_0(T) + (1-\gamma)X_1(T))^\alpha \,|\, X_0(t) = x_0, X_1(t) = x_1]$$

Theorem

J is concave, continuous, and a viscosity solution of

$$\max \{ \, J_t + \mathcal{A}J, \; -(1+\gamma)J_{x_0} + J_{x_1}, \; (1-\gamma)J_{x_0} - J_{x_1} \, \} = 0$$

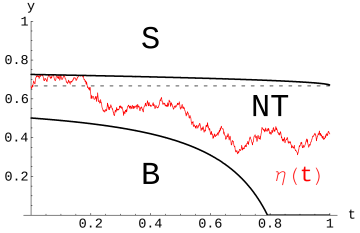
on $[0, T) \times \mathcal{D}$ *with* $J(T, x_0, x_1) = \frac{1}{\alpha}(x_0 + (1-\gamma)x_1)^\alpha$

$$\mathcal{A}\, h(x_0, x_1) = r x_0 h_{x_0}(x_0, x_1) + \mu h_{x_1}(x_0, x_1) + \tfrac{1}{2}\sigma^2 x_1^2 \, h_{x_1, x_1}(x_0, x_1)$$

NO TRADING – BUY – SELL

$y = \frac{x_1}{x_0 + x_1}$

Portfolio Optimization cont.

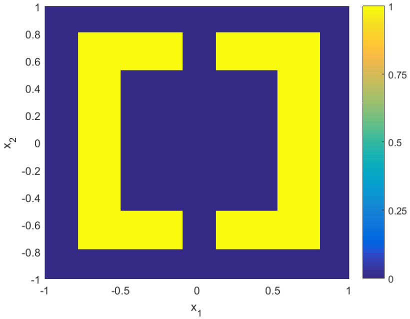


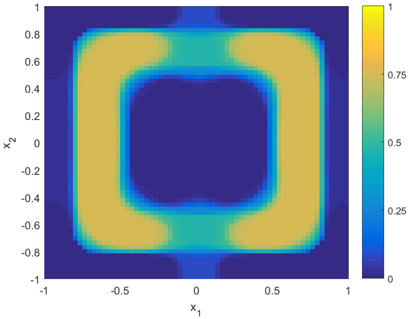
without liquidation costs

optimally controlled risky fraction $\eta = \frac{x_1}{x_0+x_1}$

Multi-topology Optimization

$$
\begin{cases}
\min\limits_{u \in BV(\Omega)} \frac{1}{2}\|y - z\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2} + \beta \int_\Omega \prod_{i=1}^{d} |u(x) - u_i|^0 \, dx + \beta \int_\Omega |Du| \\
\text{s.t.} \quad -div(u\nabla y) = f \text{ in } \Omega, \\
\qquad y = 0 \text{ on } \partial\Omega.
\end{cases}
$$





A model problem: control constrained optimal control

$$(P)\begin{cases} \min\ J(y,u) = \frac{1}{2}\int_\Omega |y-z|^2 + \frac{\beta}{2}\int_\Omega |u|^2 \\ -\Delta y = u \ \text{ in }\ \Omega, \quad y = 0 \ \text{ on }\ \partial\Omega \\ u \le \psi, \qquad u \in L^2(\Omega). \end{cases}$$

$$\mathcal{L}(y,u,p,\lambda) = J(y,u) + \langle p, -\Delta y + u\rangle + \langle \lambda, u - \psi\rangle$$

$$(OS)\begin{cases} -\Delta y = u \ \text{ in }\ \Omega, \quad y = 0 \ \text{ on }\ \partial\Omega \\ -\Delta p = -(y-z) \ \text{ in }\ \Omega, \quad y = 0 \ \text{ on }\ \partial\Omega \\ \beta u + \lambda = p \ \text{ in }\ \Omega \\ \lambda \ge 0, \quad u \le \psi, \quad \int_\Omega \lambda\,(u-\psi)\,dx = 0 \end{cases}$$

A model problem: control constrained optimal control cont

$$\lambda \geq 0, \quad u \leq \psi, \quad \int_\Omega \lambda\,(u-\psi)\,dx = 0$$

complementarity functional: $\mathcal{C}(u,\lambda) = 0$ e.g.,

$$\mathcal{C}(u,\lambda) = \lambda - \max(0, \lambda + c(u-\psi)), \quad c > 0$$

(OS) equivalent to

$$(*) \quad F(u,\lambda) :\,= \begin{pmatrix} \beta\, u + (-\Delta)^{-2} u + \lambda - f \\ \lambda - \max(0, \lambda + c(u-\psi)) \end{pmatrix} = 0$$

where $f = (-\Delta)^{-1} z$.

A model problem: control constrained optimal control cont

Optimal control: $c = \beta$ in

$$(*) \quad F(u,\lambda) := \begin{pmatrix} \beta\, u + (-\Delta)^{-2} u + \lambda - f \\ \lambda - \max(0, \lambda + c(u-\psi)) \end{pmatrix} = 0$$

$$\beta\, u - \beta\, \psi + \max(0, (-\Delta)^{-2}\, u - f + \beta\, \psi) = 0$$

Theorem

If $\|(u^\circ, \lambda^\circ) - (u^*, \lambda^*)\| \ll$, then the semi–smooth Newton-method for

$$F(u,\lambda) = \begin{pmatrix} \beta u + (-\Delta)^{-2} u + \lambda - f \\ \lambda - \max(0, \lambda + \beta(u-\psi)) \end{pmatrix} = 0$$

converges super-linearly, and rate of convergence.

The reduced problem

$$(*) \qquad F(u,\lambda) :=\begin{pmatrix} \beta\, u + (-\Delta)^{-2}u + \lambda - f \\ \lambda - \max(0, \lambda + c(u-\psi)) \end{pmatrix} = 0$$

Reduced problem

$$\begin{cases} \min J(u) = \frac{1}{2}(u, Au)_{L^2(\Omega)} - (f,u)_{L^2(\Omega)} \\ \text{subject to } u \le \psi \end{cases} \qquad \text{(P)}$$

where $A := \beta I + (-\Delta)^{-1}$ compact perturbation of identity

Focus on: $\lambda - \max(0, \lambda + \beta(u-\psi))$

Primal–dual active set strategy

Predictor: $\mathcal{A}_k = \{x\colon (\lambda^k + c(u^k - \psi))(x) > 0\}$

PDA–Algorithm

(i) choose u°, λ°

(ii) Set $\mathcal{A}_k, \mathcal{I}_k$

(iii) Solve $Au_k + \lambda_k = f$

$$u_k = \psi \text{ auf } \mathcal{A}_k, \quad \lambda_k = 0 \text{ on } \mathcal{I}_k.$$

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

for optimal control replace (iii) by

$$-\Delta y_k = \begin{cases} \psi & \text{in } \mathcal{A}_k \\ \dfrac{p_k}{\beta} & \text{in } \mathcal{I}_k \end{cases} \qquad - \Delta p_k = -(y_k - z_d),$$

and set

$$u_k = \begin{cases} \psi & \text{in } \mathcal{A}_k \\ \dfrac{p_k}{\beta} & \text{in } \mathcal{I}_k. \end{cases} \qquad \text{and} \qquad \lambda_k = p_k - \alpha u_k.$$

Primal–dual active set strategy cont.

Theorem

PDA-algorithm is equivalent to semi-smooth Newton algorithm.

Global convergence ?

$$\begin{cases} \min J(u) = \frac{1}{2}(u, Au)_{\mathbb{R}^n} - (f, u)_{\mathbb{R}^n} \\ \text{subject to } u \leq \psi \end{cases} \tag{P}$$

Theorem

Assume that A is an M-matrix. Then $\lim_{k\to\infty}(u_k, \lambda_k) = (u^*, \lambda^*)$ for arbitrary initial data. Moreover, $u^* \leq u_{k+1} \leq u_k$ for all $k \geq 1$ and $u_k \leq \psi$ for all $k \geq 2$.

Theorem

If A is a P*-matrix and for every partitioning of the index set into disjoint subsets $\mathcal{I}$ and $\mathcal{A}$ we have $\|(A_{\mathcal{I}}^{-1} A_{\mathcal{I}\mathcal{A}})_+\|_1 < 1$ and $\sum_{i\in\mathcal{I}}(A_{\mathcal{I}}^{-1} y_{\mathcal{I}})_i \geq 0$ for $u_{\mathcal{I}} \geq 0$, then $\lim_{k\to\infty}(u_k, \lambda_k) = (u^*, \lambda^*)$.*

Theorem

Assume that $A = M + K$ with M an M-matrix and $\|K\|_1$ sufficiently small. Then the primal-dual active set algorithm converges to the unique solution u^* of (P).

Applicable to optimal control ?

Primal–dual active set strategy cont.

Theorem

PDA-algorithm is equivalent to semi-smooth Newton algorithm.

Global convergence ?

$$\begin{cases} \min J(u) = \frac{1}{2}(u, Au)_{\mathbb{R}^n} - (f, u)_{\mathbb{R}^n} \\ \text{subject to } u \leq \psi \end{cases} \tag{P}$$

Theorem

Assume that A is an M-matrix. Then $\lim_{k\to\infty}(u_k, \lambda_k) = (u^*, \lambda^*)$ for arbitrary initial data. Moreover, $u^* \leq u_{k+1} \leq u_k$ for all $k \geq 1$ and $u_k \leq \psi$ for all $k \geq 2$.

Theorem

If A is a P-matrix and for every partitioning of the index set into disjoint subsets $\mathcal{I}$ and $\mathcal{A}$ we have $\|(A_{\mathcal{I}}^{-1} A_{\mathcal{I}\mathcal{A}})_+\|_1 < 1$ and $\sum_{i\in\mathcal{I}}(A_{\mathcal{I}}^{-1} y_{\mathcal{I}})_i \geq 0$ for $u_{\mathcal{I}} \geq 0$, then $\lim_{k\to\infty}(u_k, \lambda_k) = (u^*, \lambda^*)$.

Theorem

Assume that $A = M + K$ with M an M-matrix and $\|K\|_1$ sufficiently small. Then the primal-dual active set algorithm converges to the unique solution u^* of (P). *Applicable to optimal control ?*

PDA algorithm: The continuous bilateral case

$$\min \tfrac{1}{2} \, |Au - z|_Y \; + \; \tfrac{\beta}{2} \, |u|^2_{L^2(\Omega)}$$

$$\text{over } u \text{ subject to } \; \varphi \le u \le \psi,$$

where $A \in \mathcal{L}(L^2(\Omega), Y)$. e.g.: $A = -\Delta^{-1}, Y = L^2(\Omega)$.

$$\begin{cases} \alpha \, u + A^*(Au - z) + \lambda = 0 \\ \lambda = \max(0, \lambda + \beta \, (u - \psi)) + \min(0, \lambda + \beta \, (u - \varphi)). \end{cases} \tag{OS}$$

Primal–dual active set algorithm

(i) Given (u^k, λ^k), determine
$\mathcal{A}^+_{k+1} = \{x \colon (\lambda^k + \beta(u^k - \psi))(x) > 0\}$
$\mathcal{I}_{k+1} = \{x \colon (\lambda^k + \beta(u^k - \psi))(x) \le 0 \le \lambda^k + \beta(u^k - \varphi))(x)$
$\mathcal{A}^-_{k+1} = \{x \colon (\lambda^k + \beta(u^k - \varphi))(x) < 0\}$.

(ii) Determine (u^{k+1}, λ^{k+1}) from

$$u^{k+1} = \psi \;\text{ on }\; \mathcal{A}^+_{k+1}, \;\; u^{k+1} = \varphi \;\text{ on }\; \mathcal{A}^-_{k+1}, \;\; \lambda^{k+1} = 0 \;\text{ on }\; \mathcal{I}_{k+1},$$

$$\beta \, u^{k+1} + A^*(Au^{k+1} - z) + \lambda^{k+1} = 0. \tag{1}$$

Global convergence

Merit function.

$$M(u,\lambda) = \beta^2 \int_\Sigma (|(u-\psi)^+|^2 + |(\varphi-u)^+|^2)dx + \int_{\mathcal{A}^+(u)} |\lambda^-|^2 dx + \int_{\mathcal{A}^-(u)} |\lambda^+|^2 dx,$$

where $\mathcal{A}^+(u) = \{x\colon u \geq \psi\}$ and $\mathcal{A}^-(u) = \{x\colon u \leq \varphi\}$.

Theorem

Assume that $\|A\| < \beta$. Then there exists $\rho \in (0,1)$ such that $M(u^{k+1},\lambda^{k+1}) \leq \rho M(u^k,\lambda^k)$ for every $k = 1, \cdots$.
Moreover there exist $(u^*,\lambda^*) \in L^2(\Omega) \times L^2(\Omega)$, such that $\lim_{k\to\infty}(u^k,p^k,\lambda^k) = (u^*,p^*,\lambda^*)$ and (u^*,y^*,p^*,λ^*) satisfies (OS).

Corollary

Finite step convergence for discretized problem.

FIND DIFFERENT MERIT FUNCTIONS ?

State constraints: towards regularisation

$$(P)\begin{cases}\min\ J(y,u) = \frac{1}{2}\int_\Omega |y-z|^2\,dx + \frac{\beta}{2}\int_\Omega |u|^2\,dx \\ -\Delta y = u \ \text{ in } \ \Omega, \quad y = 0 \ \text{ on } \partial\Omega \\ y \le \psi \ \text{ in } \ \Omega\end{cases}$$

$$\Omega \subset \mathbb{R}^n, \quad n \le 3.$$

$$(y,u) \in (H_0^1(\Omega) \cap H^2(\Omega)) \times L^2(\Omega) = \mathcal{W} \times L^2(\Omega)$$

inequality constraint $G(y) = y - \psi \le 0$,

$G : H_0^1 \cap H^2 \to L^2(\Omega)$ NOT surjective.

ψ sufficiently regular, $\psi \ge 0$ on $\partial\Omega$

there exists an admissible pair $(y(\tilde{u}), \tilde{u}) \Longrightarrow$ existence to (P).

State constraints: towards regularisation

$$(P)\begin{cases}\min\ J(y,u)=\frac{1}{2}\int_\Omega |y-z|^2\,dx+\frac{\beta}{2}\int_\Omega |u|^2\,dx\\ -\Delta y=u\ \text{ in }\ \Omega,\quad y=0\ \text{ on }\partial\Omega\\ y\le\psi\ \text{ in }\ \Omega\end{cases}$$

$$\Omega\subset\mathbb{R}^n,\quad n\le 3.$$

$$(y,u)\in(H_0^1(\Omega)\cap H^2(\Omega))\times L^2(\Omega)=\mathcal{W}\times L^2(\Omega)$$

inequality constraint $G(y)=y-\psi\le 0,$

$G: H_0^1\cap H^2\to L^2(\Omega)$ NOT surjective.

$$\psi \text{ sufficiently regular},\quad \psi\ge 0 \text{ on } \partial\Omega$$

there exists an admissible pair $(y(\tilde u),\tilde u)\Longrightarrow$ existence to (P).

State constraints: towards regularisation

Theorem

(y^*, u^*) is solution to (P) iff $\exists\, p^* \in L^2(\Omega),\ \lambda^* \in \mathcal{W}^* \cap C(\Omega)^*$:

$$\begin{cases} -\Delta y^* = u^* \text{ in } \Omega, \quad y^* = 0 \text{ on } \partial\Omega \\ (p^*, -\Delta y) + \langle \lambda^*, y \rangle_{C^*,C} = -(y^* - z), \qquad \forall\, y \in \mathcal{W}(\Omega) \\ \langle \lambda^*, y - y^* \rangle_{C^*,C} \leq 0 \quad \text{for all } \ y \in C(\Omega), \quad y \leq \psi. \\ p^* = \beta u^*, \quad y^* \leq \psi. \end{cases}$$

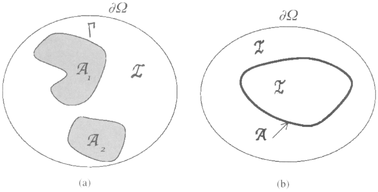


Fig. 1. (a) Active set: case (A1). (b) Active set: case (A2).

Theorem (Case 1)

$$p^* \in H_0^1(\Omega), \quad p^*|_{\mathring{\mathcal{A}}} \in H^2(\mathring{\mathcal{A}}), \quad p^*|_{\mathcal{I}} \in H^2(\mathcal{I})$$

$$\begin{aligned} p^* &= -\beta \Delta \psi & & \text{on } \mathcal{A} \\ -\Delta p^* &= -(y^* - z), & & \text{in } \mathcal{I}, \quad p^* = 0 \text{ on } \partial\Omega \\ \tau_\Gamma p^* &= -\beta\, \tau_\Gamma \Delta \psi\ , & & \Gamma = \partial \mathcal{A} \\ \lambda^* &= \mu^* + \mu_\Gamma^*\ , & & \mu^* \in L^2(\Omega)\ , \quad \mu_\Gamma^* \in H^{1/2}(\Gamma) \end{aligned}$$

$$\begin{aligned} \mu^* &= \begin{cases} 0 & \text{on} \quad \mathcal{I} \\ z - \psi - \alpha \Delta^2 \psi & \text{in} \quad \mathring{\mathcal{A}} \end{cases} \\ \mu_\Gamma^* &= -\frac{\partial p^*|_{\mathcal{I}}}{\partial n|_{\mathcal{I}}} + \frac{\partial p^*|_{\mathcal{A}}}{\partial n|_{\mathcal{A}}}. \end{aligned}$$

Theorem (Case 2)

$$p^* \in W_0^{1,q}(\Omega)\ , \quad \forall q \in (1,2)$$

$$\lambda^*|_{\mathring{\mathcal{A}}} = \frac{\partial p_+^*}{\partial n} - \frac{\partial p_-^*}{\partial n} \quad \text{in} \quad \mathcal{M}(\mathring{\mathcal{A}}).$$

Recall: Primal-Dual Active Set Method

$$\langle \lambda^*, y - y^* \rangle_{C^*,C} \leq 0, \quad \forall\, y \leq \psi$$

$$\langle \lambda^*, \psi - y^* \rangle_{C^*,C} = 0\,, \quad y^* \leq \psi\,, \quad \lambda^* \geq 0$$

formally: $\lambda^* = \max(0, \lambda^* + c(y^* - \psi))\,, \quad c > 0.$

Algorithm

(i) initialize (y_0, u_0, λ_0), choose $c > 0$

(ii) $\mathcal{A}_{k+1} = \{x : (\lambda_k + c(y_k - \psi))(x) > 0\}$

(iii)

$$-\Delta y_{k+1} = u_{k+1}$$

$$-\Delta p_{k+1} + \lambda_{k+1} = -(y_{k+1} - z)$$

$$y_{k+1} = \psi \text{ on } \mathcal{A}_{k+1}\,, \quad \lambda_{k+1} = 0 \quad \text{on} \quad \mathcal{I}_{n+1}$$

$$\beta u_{k+1} = p_{k+1} \text{ in } \Omega$$

(iv) $k = k + 1.$

Convergence for discretized problems

Remarks:

1. (iii) is solution to (P) with $y = \psi$ on $\mathcal{A}_{k+1}$ and free on $\mathcal{I}_{k+1}$
2. solve (iii) only on $\mathcal{I}_{k+1}$
3. termination $\mathcal{A}_{k+1} = \mathcal{A}_k$
4. $\mathcal{A}_{k+1}$ may be very different from $\mathcal{A}_k$

numerical practice: mesh dependent convergence

Mesh size h	1/16	1/32	1/64	1/128	1/256
PDAS	14	27	54	113	226

Regularization of Infinite Dimensional Problems

$$(P_\gamma) \quad \begin{cases} \min \frac{1}{2}\,|y-z|^2 + \frac{\beta}{2}\,|u|^2 + \frac{1}{2\gamma}\int_\Omega |(\bar\lambda + \gamma(y-\psi))^+|^2 dx \\ -\Delta y = u, \ \text{ in } \ \Omega, \quad y = 0 \ \text{ on } \partial\Omega. \end{cases}$$

$$(OS_\gamma) \begin{cases} -\Delta y = u \ \text{ in } \Omega, & y = 0 \ \text{ on } \ \partial\Omega \\ -\Delta p + \lambda = -(y-z) \ \text{ in } \Omega, & p = 0 \ \text{ on } \ \partial\Omega \\ \beta u = p & \\ \lambda = \max(0, \bar\lambda + \gamma(y-\psi)) & \end{cases}$$

optimality condition for reduced regularized problem

$$\lambda = \max\left(0, \bar\lambda + \gamma((I + \beta\Delta^2)^{-1}(z-\lambda) - \psi)\right)$$

compare: optimality condition for reduced regularized problem

$$\lambda = \max\left(0, \lambda + c(y-\psi)\right) = \max(0, \lambda + (I+\beta\Delta^2)^{-1}(z-\lambda) - \psi) \quad \text{ each } c > 0$$

Regularization of Infinite Dimensional Problems cont

Theorem

$$(y_\gamma, p_\gamma, \lambda_\gamma) \underset{\gamma\to\infty}{\longrightarrow} (y^*, p^*, \lambda^*) \;\; in \;\; \mathcal{W} \times L^2(\Omega) \times \mathcal{W}^*_{weak}$$

Algorithm (regularized)

(i) initialize (y_0, u_0, λ_0), $\quad k := 0,$ $\quad$ choose $\gamma > 0$

(ii) $\mathcal{A}_{k+1} = \{x : \bar{\lambda} + \gamma(y_k - \psi) > 0\}$

(iii) $\begin{cases} -\Delta y = \frac{1}{\beta}\, p \\ -\Delta p + (\bar{\lambda} + \gamma(y - \psi))\mathcal{X}_{\mathcal{A}_{k+1}} = -(y - z) \end{cases}$

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

Regularization of Infinite Dimensional Problems cont

Theorem

If $y_0, \psi, \bar{\lambda} \in L^p(\Omega), p > 2$, $|y_0 - y_\gamma|_{L^p} <<,$ *then* $(y_k, p_k, \lambda_k) \underset{k\to\infty}{\longrightarrow} (y_\gamma, p_\gamma, \lambda_\gamma)$ *superlinearly in* $\mathcal{W} \times \mathcal{W} \times L^2(\Omega)$.

Proposition:

If $\mathcal{A}_{k+1} = \mathcal{A}_k \Rightarrow (y_k, u_k) = (y_\gamma, u_\gamma)$.

Theorem

If $\frac{\gamma}{\beta}||\Delta^{-1}||_{\mathcal{L}(L^2(\Omega))} < 1$, *then global convergence.*

γ	10^3	10^4	10^5	10^6	10^8	10^9	10^{10}
iter	10	17	27	30	30	31	31
active	791	667	606	587	577	575	575

Table

γ	10^3	10^6	10^9	
iter	10	6	3	$\Sigma = 19$

Table

How to choose γ: Path Following

$$(P_\gamma) \quad \begin{cases} \min \frac{1}{2}\,|y-z|^2 + \frac{\beta}{2}\,|u|^2 + \frac{1}{2\gamma}\int_\Omega |(\bar{\lambda} + \gamma(y-\psi))^+|^2 dx \\ -\Delta y = u, \quad \text{in} \quad \Omega, \qquad y = 0 \quad \text{on } \partial\Omega. \end{cases}$$

$$\mathcal{P} = \{(y_\gamma, u_\gamma, p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2 \times L^2 \times \mathcal{W}^*\}$$

$(P_{\gamma=0})$ unconstrained, $\qquad (P_{\gamma=\infty})$ constrained.

Theorem

$\mathcal{P}$ *is globally Lipschitz continuous, and* $\gamma \to (p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2$ *is locally Lipschitz continuous.*

$$(H) \quad S_\gamma^0 := \{x \in \Omega : \; y_\gamma - \psi = 0\}, \quad \textit{meas } S_\gamma^0 = 0$$

How to choose γ: Path Following

$$(P_\gamma) \quad \min \frac{1}{2}|Au - z|^2 + \frac{\beta}{2}|u|^2 + \frac{1}{2\gamma}\int_\Omega |(\bar{\lambda} + \gamma(y - \psi))^+|^2 dx$$

$$\mathcal{P} = \{(y_\gamma, u_\gamma, p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2 \times L^2 \times \mathcal{W}^*\}$$

$(P_{\gamma=0})$ unconstrained, $(P_{\gamma=\infty})$ constrained.

Theorem

$\mathcal{P}$ is globally Lipschitz continuous, and $\gamma \to (p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2$ is locally Lipschitz continuous.

Remark: $\bar{\lambda} >>$ can guarantee feasibility of y_γ.
For the obstacle problem, choose $\bar{\lambda} = \max(0, f + \Delta\psi)$.

$(H) \quad S^0_\gamma := \{x \in \Omega : \; y_\gamma - \psi = 0\}, \quad meas\, S^0_\gamma = 0$

How to choose γ: Path Following

$$(P_\gamma) \quad \min \tfrac{1}{2}\,|Au-z|^2 + \tfrac{\beta}{2}\,|u|^2 + \tfrac{1}{2\gamma}\int_\Omega |(\bar\lambda + \gamma(y-\psi))^+|^2 dx$$

$$\mathcal{P} = \{(y_\gamma, u_\gamma, p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2 \times L^2 \times \mathcal{W}^*\}$$

$(P_{\gamma=0})$ unconstrained, $\qquad (P_{\gamma=\infty})$ constrained.

Theorem

$\mathcal{P}$ is globally Lipschitz continuous, and $\gamma \to (p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2$ is locally Lipschitz continuous.

Remark: $\bar\lambda >>$ can guarantee feasibility of y_γ.
For the obstacle problem, choose $\bar\lambda = \max(0, f + \Delta\psi)$.

$(H) \quad S^0_\gamma := \{x \in \Omega : \; y_\gamma - \psi = 0\}, \quad meas\ S^0_\gamma = 0$

How to choose γ: Path Following

$$(P_\gamma) \quad \min \frac{1}{2}|Au - z|^2 + \frac{\beta}{2}|u|^2 + \frac{1}{2\gamma}\int_\Omega |(\bar\lambda + \gamma(y-\psi))^+|^2 dx$$

$$\mathcal{P} = \{(y_\gamma, u_\gamma, p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2 \times L^2 \times \mathcal{W}^*\}$$

$(P_{\gamma=0})$ unconstrained, $(P_{\gamma=\infty})$ constrained.

Theorem

$\mathcal{P}$ is globally Lipschitz continuous, and $\gamma \to (p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2$ is locally Lipschitz continuous.

Remark: $\bar\lambda >>$ can guarantee feasibility of $\underline{y}_\gamma$.
For the obstacle problem, choose $\bar\lambda = \max(0, f + \Delta\psi)$.

$$(H) \quad S_\gamma^0 := \{x \in \Omega : \ y_\gamma - \psi = 0\}, \quad meas\, S_\gamma^0 = 0$$

Sensitivities for Path Following

Theorem

$\gamma \to (y_\gamma, u_\gamma, p_\gamma) \in \mathcal{W} \times L^2 \times L^2_{weak}$ *is differentiable and*

$$(OS_\gamma) \quad -\Delta \dot{y} = \dot{u}, \quad -\Delta \dot{p} + (y_\gamma - \psi + \gamma \dot{y})\mathcal{X}_{S_\gamma} = -\dot{y}, \quad \beta \dot{u} = \dot{p}$$

where $S_\gamma = \{x : y_\gamma - \psi > 0\}$.

$$V(\gamma) = \min\ J(y_\gamma, u_\gamma) + \frac{\gamma}{2} \int_\Omega |(y_\gamma - \psi)^+|^2$$

Theorem

$$\dot{V}(\gamma) = \frac{1}{2} \int_\Omega |(y_\gamma - \psi)^+|^2\,, \qquad \ddot{V}(\gamma) = \int_\Omega (y_\gamma - \psi)^+ \dot{y}$$

Corollary

$$\dot{V}(\gamma) > 0\,, \quad \ddot{V}(\gamma) < 0\,, \quad V(0) \triangleq (P_{unconstr.})\,, \quad V(\infty) \triangleq (P)$$

Sensitivities for Path Following

Theorem

$\gamma \to (y_\gamma, u_\gamma, p_\gamma) \in \mathcal{W} \times L^2 \times L^2_{weak}$ *is differentiable and*

$(OS_\gamma) \quad -\Delta \dot{y} = \dot{u}, \quad -\Delta \dot{p} + (y_\gamma - \psi + \gamma \dot{y})\mathcal{X}_{S_\gamma} = -\dot{y}, \quad \beta \dot{u} = \dot{p}$

where $S_\gamma = \{x : y_\gamma - \psi > 0\}$.

$$V(\gamma) = \min \, J(y_\gamma, u_\gamma) + \frac{\gamma}{2} \int_\Omega |(y_\gamma - \psi)^+|^2$$

Theorem

$$\dot{V}(\gamma) = \frac{1}{2} \int_\Omega |(y_\gamma - \psi)^+|^2 \, , \qquad \ddot{V}(\gamma) = \int_\Omega (y_\gamma - \psi)^+ \dot{y}$$

Corollary

$$\dot{V}(\gamma) > 0 \, , \quad \ddot{V}(\gamma) < 0 \, , \quad V(0) \triangleq (P_{unconstr.}) \, , \quad V(\infty) \triangleq (P)$$

Model Function.

$$m(\gamma) = C_1 - \frac{C_2}{E+\gamma}\ .$$

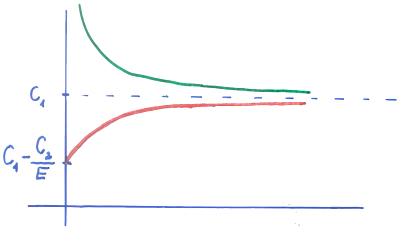
recall: $\dot{m}(\gamma) \sim \dot{V}(\gamma) = \frac{1}{2}\int_\Omega |(y_\gamma - \psi)^+|^2$

test (OS_γ) with $y_\gamma - \psi)^+$ and "approximate" ∞ dimensional quantities by constants

$$(E+\gamma)\ddot{m}(\gamma) + 2\,\dot{m}(\gamma) = 0$$

$\rightarrow$ model function

$\dot{m} \geq 0,\ \ddot{m} \leq 0,\ \gamma^2\dot{m}(\gamma)$ bounded for $\gamma \to \infty$.



Path-following Algorithms

Model parameters

$$m(0) = V(0)\,, \quad m(\gamma) = V(\gamma)\,, \quad \dot{m}(\gamma) = \dot{V}(\gamma) = \frac{1}{2}\int_\Omega |(y_\gamma - \psi)^+|^2$$

determine $E > 0\,,\ C_1 > 0\,,\ C_2 > 0$.

Update Strategy

$$|V^* - V(\gamma_{k+1})| \leq \tau_k |V^* - V(\gamma_k)|$$

$$|C_{1,k} - m_k(\gamma_{k+1})| \leq \tau_k |C_{1,k} - V(\gamma_k)| =: \beta_k$$

$$\gamma_{k+1} = \frac{C_{2,k}}{\beta_k} - E_k.$$

Theorem (exact path following)

$$\lim_{k\to\infty} (y_{\gamma_k}, u_{\gamma_k}, \lambda_{\gamma_k}) \to (y^*, u^*, \lambda^*).$$

Inexact Path-Following

$$\mathcal{N}(\gamma) = \{(y,\lambda) : |(r_\gamma^1(y,\lambda),\ r_\gamma^2(y,\lambda))|_{\mathbb{R}^2} < \frac{\tau}{\sqrt{\gamma}}\}$$

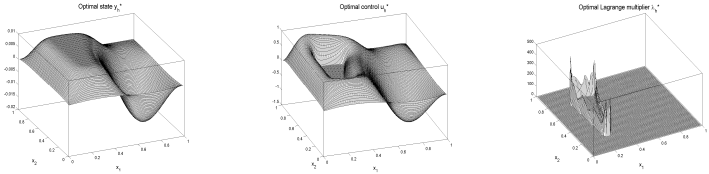
$$r_\gamma^1(y,\lambda) = |\Delta y + \frac{1}{\beta}\Delta^{-1}(\lambda + y - z)|_{H^{-1}}\ ,\ r_\gamma^2(y,\lambda) = |\lambda - \max(0, \bar{\lambda} + \gamma(y-\psi))|_{L^2}.$$

$$\gamma_{k+1} \geq \max\left(\gamma_k \max(\tau_1, \frac{\rho_{k+1}^F}{\rho_{k+1}^C}),\ \frac{1}{\max(\rho_{k+1}^F, \rho_{k+1}^C)^q}\right),$$

where $q \geq 1$, $\quad \tau_1 > 1$

$$\rho_{k+1}^F := \int_\Omega (y_{k+1} - \psi)^+ dx\ ,\ \rho_{k+1}^C := \int_{\mathcal{I}_{k+1}} (y_{k+1} - \psi)^+ dx + \int_{\mathcal{A}_{k+1}} (y_{k+1} - \psi)^- dx.$$

Inexact pathfollowing



Optimal state (left), optimal control (middle), and optimal multiplier (right) for problem 1 with $h = 1/128$.

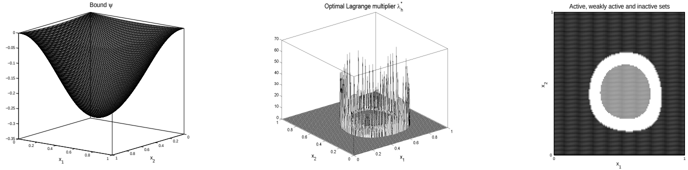
Mesh size h	1/16	1/32	1/64	1/128	1/256
PDAS	14	27	54	113	226
PDIP	12	14	15	19	19
IPF	11	15	14	13	15

Comparison of iteration numbers for different mesh sizes and methods.

Mesh size h	1/4	1/8	1/16	1/32	1/64	1/128	1/256	total
PDAS	3	4	4	5	6	6	6	34
PDIP	3	2	4	4	5	6	7	31
IPF	4	3	3	4	5	5	5	29

Comparison of iteration numbers for different mesh sizes and methods based on nested iteration.

Problem with lack of strict complementarity



bound ψ (left), optimal multiplier (middle), active/inactive sets (right), for $h = 1/128$.

Mesh size h	1/4	1/8	1/16	1/32	1/64	1/128	1/256	total
PDAS	2	4	5	9	10	21	40	91
PDIP	3	2	3	3	6	12	11	40
IPF	7	2	4	4	6	8	15	46

Table: Comparison of iteration numbers for different mesh sizes and methods based on nested iteration.

BUT: CPU-time for PDIP 20 percent higher than for IPF

Some references

- Primal-dual Strategy for Constrained Optimal Control Problems M. BERGOUNIOUX, K. ITO and K. KUNISCH, SIAM J. Control and Optimization, 37(1999), 1176-1194 doi: 10.1137/S0363012997328609
- The primal-dual active set strategy as a semi-smooth Newton method M. HINTERMÜLLER, K. ITO and K. KUNISCH, SIAM Journal on Optimization, 13(2002), 865-888 doi: 10.1137/S1052623401383558
- The primal-dual active set method for nonlinear optimal control problems with bilateral constraints K. ITO and K. KUNISCH, SIAM J. on Control and Optimization, 43(2004), 357-376 doi: 10.1137/S0363012902411015
- Generalized Newton Methods for the 2D-Signorini Contact Problem with Friction K. KUNISCH, G. STADLER, ESAIM: M2AN, 39(2005), 827-854 doi: 10.1051/m2an:2005036
- Feasible and Non-Interior Path-Following in Constrained Minimization with Low Multiplier Regularity M. HINTERMÜLLER and K. KUNISCH, SIAM J. Control and Optim., 45(2006), 1198-1221 doi: 10.1137/050637480