Recent Advances in Optimal Control of Variational Inequalities

Michael Hintermüller

MATHEON – DFG Research Center
&
Department of Mathematics, Humboldt-University of Berlin.
hint@mathematik.hu-berlin.de

Acknowledgment:
FWF under START-grant Y305 "Interfaces and Free Boundaries".

joint work with I. Kopacka and H. Tber.

Motivation I.

- **Box constrained variational inequality.** Let $G : L^2(\Omega) \to L^2(\Omega)$. Problem: Find

$$u \in U_{\text{ad}} : \quad (G(u), v - u)_{L^2} \geq 0 \quad \forall v \in U_{\text{ad}},$$

where

$$U_{\text{ad}} = \{u \in L^2(\Omega) : a \leq u \leq b \text{ a.e. in } \Omega\}.$$

- **Equivalent formulation.** Find

$$u \in U_{\text{ad}} : \quad (u - a)G(u) \leq 0, \quad (u - b)G(u) \leq 0 \quad \text{a.e. in } \Omega.$$

- **Yet another equivalent formulation.** Find $u \in L^2(\Omega)$ such that

$$\tilde{F}(u) := u - P_{U_{\text{ad}}}(u - \sigma G(u)) = 0 \quad \text{a.e. in } \Omega$$

for arbitrarily fixed $\sigma > 0$. Here, $P_{U_{\text{ad}}}$ is the L^2-projection onto U_{ad}.

Motivation I.

- Example structure.
$$G(u) = A(u) + \alpha u$$
with $A : L^2(\Omega) \to L^2(\Omega)$ Fréchet-differentiable and locally Lipschitz from $L^2(\Omega)$ to $L^q(\Omega)$ for some $q > 2$.
- Format of nonsmooth equation ($\sigma = 1/\alpha$).
$$F(u) := \alpha u - P_{\alpha U_{\text{ad}}}(-A(u)) = 0 \quad \text{a.e. in } \Omega.$$
- **Derivative of projection term.**
$$\partial P_{U_{\text{ad}}}(u) = \{ D(u) \cdot A'(u) \}$$
with $D : L^2(\Omega) \to L^\infty(\Omega)$ satisfying
$$D(u)(x) \begin{cases} = 0 & \text{if } -A(u)(x) \notin [\alpha a(x), \alpha b(x)], \\ \in \mathbb{R} & \text{if } -A(u)(x) \in \{\alpha a(x), \alpha b(x)\}, \\ = 1 & \text{if } -A(u)(x) \in (\alpha a(x), \alpha b(x)). \end{cases}$$
- **Derivative of F.**
$$\partial F(u) = \alpha \, \text{id} + \partial P_{U_{\text{ad}}}(u).$$

Motivation I.

- **Semismoothness.**

$$\sup_{G_F \in \partial F(u+d)} \|F(u+d) - F(u) - G_F\, d\|_{L^2} = \mathcal{O}(\|d\|_{L^2}) \quad \text{as } \|d\|_{L^2} \to 0.$$

- Used for analyzing locally superlinear convergence of a generalized version of Newton's method in function space.
- **Note.** Analysis for $S(u^k) \subset \partial F(u^k)$, $S(u^k) \neq \emptyset$, sufficient!
- **Aim.** Establish **mesh independent convergence** for a properly discretized version of the method.
- **Troubles due to nonsmoothness.** Define

$$f(t) = t + \max(t, \omega t), \quad \omega \geq 2,$$

and the perturbed ("discretized") version:

$$f_h(t) = h - h^2 + t^2 + \max(t, \omega t), \quad h \in (0, 1/2].$$

Let $t^* = 0$ and $t_h^* = -h$ denote the respective solution of interest.

Motivation I.

- Then
$$|t^{k+1} - t^*| \leq \ \omega^{-1}|t^k - t^*|^2 \text{ for } t^k \in (-1, 1);$$
BUT: for any $t_h^k \in (0, \zeta h]$, $\zeta \in (0, 1]$
$$|t_h^{k+1} - t_h^*| \geq \ \hat{\omega}|t_h^k - t_h^*|$$
for some $\hat{\omega} \in (0, 1)$ depending only on ω, ζ
- !!! t_h^k **on wrong side of kink**!!!

Motivation I - Mesh independence.

Let $u^* \in L^2(\Omega)$ satisfy $F(u^*) = 0$ and let u_h^* be solution of $F_h(u_h) = 0$.

- **Assumption 1 (Strict complementarity).**

$$\text{meas}\left\{\min(u^* - a, b - u^*) + |G(u^*)| = 0\right\} = 0.$$

- **Assumption 2 (Consistency).**

$$\lim_{h\to 0^+} \|u_h^* - u^*\|_{L^2} = 0, \quad \lim_{h\to 0^+} \|A_h(u_h^*) - A(u^*)\|_{L^q} = 0$$

  for some $q > 2$

- **Assumption 3 (Locally uniform Lipschitz property).**
  There exist $h_0 > 0$, $\delta_0 > 0$, and $L_A > 0$ such that

$$\|A(u^2) - A(u^1)\|_{L^q} \leq L_A \|u^2 - u^1\|_{L^2}, \qquad \|u^i - u^*\|_{L^2} \leq \delta_0,$$

$$\|A_h(u_h^2) - A_h(u_h^1)\|_{L^q} \leq L_A \|u_h^2 - u_h^1\|_{L^2}, \quad \|u_h^i - u_h^*\|_{U_h} \leq \delta_0$$

  for all $0 < h \leq h_0$.

  Here, $U_h \subset L^2(\Omega)$ with $\|\cdot\|_{U_h} = \|\cdot\|_{L^2}$.

Motivation I – Mesh independence.

- **Assumption 4 (Uniform linear approximation property).**
  - A and A_h, $h \leq h_0$, are Fréchet differentiable in a neighborhood of u^* and u_h^*;
  - there exists $\rho : [0, \delta_0) \to [0, \infty)$ with $\lim_{t \to 0^+} \frac{\rho(t)}{t} = 0$ and

$$\|A(u) - A(u^*) - A'(u)(u - u^*)\|_{L^2} \leq \rho(\|u - u^*\|_{L^2})$$
$$\forall u \in L^2(\Omega), \quad \|u - u^*\|_{L^2} \leq \delta_0,$$

$$\|A_h(u_h) - A_h(u_h^*) - A_h'(u_h)(u_h - u_h^*)\|_{L^2} \leq \rho(\|u_h - u_h^*\|_{L^2})$$
$$\forall u_h \in U_h, \quad \|u_h - u_h^*\|_{L^2} \leq \delta_0, \quad h \leq h_0.$$

Motivation I – Mesh independence.

Theorem. Let $\delta_2, \delta_2' > 0$, $\kappa, \kappa' > 0$ and $h_2' \leq h_0$ such that for all $0 < h \leq h_2'$

$$\sup\{\|G^{-1}\|_{L^2,L^2} : G \in S(u^* + s), \|s\|_{L^2} \leq \delta_2\} \leq \kappa,$$
$$\sup\{\|G_h^{-1}\|_{L^2,L^2} : G_h \in S_h(u_h^* + s_h), \|s_h\|_{L^2} \leq \delta_2'\} \leq \kappa'.$$

Then, for $\theta \in (0,1)$, there exist $\bar{\delta} > 0$ and $\bar{h} > 0$ such that

$$\begin{aligned} \|u^{k+1} - u^*\|_{L^2} &\leq \theta \|u^k - u^*\|_{L^2}, \\ \|u_h^{k+1} - u_h^*\|_{L^2} &\leq \theta \|u_h^k - u_h^*\|_{L^2}, \quad \forall 0 < h \leq \bar{h} \end{aligned}$$

if $\max\{\|u^0 - u^*\|_{L^2}, \|u_h^0 - u_h^*\|_{L^2}\} \leq \bar{\delta}$.

[M.H., M. Ulbrich; Math. Prog.]

Motivation II – KKT-theory in Banach space.

Consider the minimization problem

$$\min_{x \in X} f(x) \quad \text{s.t. } x \in C,\ g(x) \in K, \tag{P}$$

- f real functional defined on a real Banach space X (C^1),
- C is a non-empty closed convex subset of X,
- g is a map from X into a real Banach space Y (C^1),
- K is a closed convex cone in Y with vertex at the origin.

For fixed $x \in X$ and $y \in Y$ let $C(x)$ and $K(y)$ denote the conical hulls of $C - \{x\}$ and $K - \{y\}$ respectively, i.e.,

$$C(x) := \{\beta(c - x) \mid c \in C,\ \beta \geq 0\}$$
$$K(y) := \{k - \beta y \mid k \in K,\ \beta \geq 0\}.$$

Motivation II – KKT-theory in Banach space.

$y^* \in Y^*$ is called a Lagrange multiplier for (P) at an optimal point $x^* \in X$, if

$$\begin{aligned} &(i) && y^* \in K^+ \\ &(ii) && \langle y^*, g(x^*)\rangle_{Y^*,Y} = 0 \\ &(iii) && f'(x^*) - y^*(g'(x^*)) \in C(x^*)^+, \end{aligned}$$

where X^* and Y^* denote the topological duals of X and Y and for each subset A of X (or Y respectively), A^+ denotes its polar cone

$$A^+ := \{w \in X^* \mid \langle w, a\rangle_{X^*,X} \geq 0 \quad \text{for all } a \in A\}.$$

Theorem. Let x^* be an optimal solution for problem (P). If

$$g'(x^*)C(x^*) - K(g(x^*)) = Y,$$

then the set $\Lambda(x^*)$ of Lagrange multipliers for problem (P) at x^* is non-empty and bounded.

[J. Zowe, S. Kurcyusz; Math. Prog.]

Motivation III – MPECs / MPCCs in function space.

Elliptic VIs.

- **Control of elVIs.**
- **Parameter identification in elVIs.**

- Reynolds lubrication equation.

$$-\operatorname{div}(u^3 \nabla y) = -\frac{\partial u}{\partial x_2} \text{ in } \Omega, \; y \in H_0^1(\Omega).$$

u ... gap height, y ... pressure in lubricant.

- + contact model (point, line, ...)

Cavitation phenomena require a VI-formulation:

$$y^* \geq 0, \quad \langle -\operatorname{div}(u^3 \nabla y^*) + \frac{\partial u}{\partial x_2}, y - y^* \rangle \geq 0 \quad \forall y \geq 0.$$

Motivation III – MPECs / MPCCs in function space.

Inverse Problem.

Reconstruct the gap height u from measurements $y_d \in L^2(\Omega)$ of the pressure y.

Output-least-squares formulation.

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\delta}{2}\|u\|_{\mathcal{U}}^2 =: J(y,u) \\ \text{over} \quad & (y,u) \in (K \subset H_0^1(\Omega)) \times \mathcal{U} \\ \text{s.t.} \quad & \langle -\operatorname{div}(u^3 \nabla y) + \frac{\partial u}{\partial x_2}, v - y \rangle \geq 0 \quad \forall v \in K. \end{aligned}$$

$K = \{v \in H_0^1(\Omega) | v \geq 0\}$, Hilbert space $\mathcal{U}$.

Parabolic VIs.

- Control of parVIs.
- Parameter id. in elastohydrodynamic lubrication (EHL) problem.

Motivation III – MPECs / MPCCs in function space.

- Calibration in American put options.

Black-Scholes model.

$$\frac{\partial y}{\partial t} - \frac{ux^2}{2}\frac{\partial^2 y}{\partial x^2} - rx\frac{\partial y}{\partial x} + ry \geq 0, \quad y(t,x) \geq y_0(x),$$

$$\left(\frac{\partial y}{\partial t} - \frac{ux^2}{2}\frac{\partial^2 y}{\partial x^2} - rx\frac{\partial y}{\partial x} + ry\right)(y - y_0) = 0, \quad t \in (0,T],\ x > 0,$$

$$y(t=0,x) = y_0(x),\ x > 0,$$

where $y_0(x) = (S - x)_+$ is the payoff.

Notation:

$y = y(t,x)$	... price,	$\sqrt{u} = \sqrt{u(t,x)}$	... volatility
r	... interest rate,	$T > 0$	... maturity
$S \geq 0$	... strike (price, fixed),	x	... spot price

Motivation III – MPECs / MPCCs in function space.

Define the spaces

$$X = \{u \in L^2(\mathbb{R}_+) : (x+1)\frac{\partial u}{\partial x} \in L^2(\mathbb{R}_+)\},$$

$$Y = \{u : u, x\frac{\partial u}{\partial x} \in H^1(0, T; X)\},$$

$$\mathcal{U} = \{u \in Y : 0 < \underline{u} \leq u \leq \overline{u},\ |x\frac{\partial u}{\partial x}| \leq M \text{ in } (0, T) \times \mathbb{R}_+\}$$

Volatility ($= \sqrt{u}$) estimation (calibration problem).

$$\begin{cases} \text{minimize} \quad h_1(y) + h_2(u) \\ \text{s.t.} \ \ u \in \mathcal{U}, \\ \qquad y = y(u) \text{ solves the Black-Scholes model} \end{cases}$$

Available work

Control of $\{$el, par$\}$VIs, MPCCs, MPECs.

- Finite dimensions: MPCCs, MPECs.
  Fletcher, Kocvara, Leyffer, Luo, Pang, Morduchovich, Outrata, Ralph, Scholtes, Zowe, ...
- Function space: Control or parameter id for VIs.
  - Barbu, Bergounioux, Bermudez, H., Ito, Kunisch, Mignot, Puel, Saguez, ...
  - Applications: Bayada, Capriz, Cimatti (EHL), Achdou (Black-Scholes), H. (Parameter id. for EHL, Black-Scholes)...
  - MPEC view: [H., Kopacka], [H., Ralph, Scholtes].
  - Literature on algorithms very scarce.

Model problem

$$\text{minimize} \quad h_1(y) + h_2(u) =: J(y,u) \quad \text{over} \quad (y,u) \in V \times \mathcal{U}$$
$$\text{s. t.} \quad y \in K,\ \langle A(u)y, v-y\rangle \geq (f(u), v-y) \quad \forall v \in K,$$

$\Omega \subset \mathbb{R}^d$, $d = 1,2,3$, open, bounded and sufficiently smooth,

$$K = \{v \in H_0^1(\Omega) =: V | v \geq 0\}, \quad \mathcal{U} = L^2(\Omega), H^1(\Omega), H^2(\Omega),$$

- h_1 is sufficiently smooth and non-negative,
- h_2 is C^1, convex lower semi-continuous, and for some constants $C_1 > 0$ and $C_2 \in \mathbb{R}$

$$h_2(u) \geq C_1 \|u\|_{\mathcal{U}} + C_2 \quad \forall u \in \mathcal{U}.$$

- $f(u) = Fu + g, \quad F \in \mathcal{L}(\mathcal{U}, L^2(\Omega)), \quad g \in L^2(\Omega).$

Model problem

VI written as a complementarity system:

$$A(u)y - \lambda = f(u), \quad y \geq 0, \quad \lambda \geq 0, \quad \langle \lambda, y \rangle = 0.$$

Additional regularity yields $\lambda \in L^2(\Omega)$.

$$\begin{aligned} \text{minimize} \quad & h_1(y) + h_2(u) =: J(y,u) \\ \text{over} \quad & (y,u,\lambda) \in H_0^1(\Omega) \times \mathcal{U} \times L^2(\Omega) \\ \text{s. t.} \quad & A(u)y - \lambda = f(u), \\ & y \geq 0, \quad \lambda \geq 0, \quad (\lambda, y)_{L^2} = 0. \end{aligned}$$

- Violates constraint qualifications (multiplier existence???);
- Bi-activities, i.e. $\{y = 0 = \lambda\} =: \mathcal{B}$, are in particular problematic;
- Structure of feasible set – pieces;
- Stationarity concept??? (no longer unique);
- Algorithm???

Relaxation/Regularisation approach.

(P_α).

$$\begin{aligned} \text{minimize} \quad & \tilde{J}(y,u,\lambda) := J(y,u) + \frac{\kappa}{2}\|\lambda\|_{L^2}^2 \\ \text{over} \quad & y \in H^2(\Omega) \cap H_0^1(\Omega) \text{ and } u \in \mathcal{U}, \lambda \in L^2(\Omega), \\ \text{s.t.} \quad & A(u)y = \lambda + f(u), \\ & \lambda \geq 0,\ y \geq 0, \ \text{ a.e. in } \Omega, \quad (y,\lambda)_{L^2} \leq \alpha. \end{aligned}$$

$(P_{\alpha,\gamma})$.

$$\begin{aligned} \text{minimize} \quad & \tilde{J}_\gamma(y,u,\lambda) := J(y,u) + \frac{\kappa}{2}\|\lambda\|_{L^2}^2 + \frac{1}{2\gamma}\|\max(0, \bar{\lambda} - \gamma y)\|_{L^2}^2 \\ \text{over} \quad & y \in H_0^1(\Omega) \text{ and } u \in \mathcal{U}, \lambda \in L^2(\Omega), \\ \text{s.t.} \quad & A(u)y = \lambda + f(u), \\ & \lambda \geq 0, \ \text{ a.e. in } \Omega, \quad (y,\lambda)_{L^2} \leq \alpha \end{aligned}$$

with $\bar{\lambda} \in L^2(\Omega)$ a non-negative (fixed) multiplier approximation; compare augmented Lagrangian approach.

Relaxation/Regularisation approach.

- $(P_{\alpha,\gamma})$ treated by standard KKT-theory in Banach space: Stationarity conditions.

$$\begin{aligned}
A(u_\gamma)^\star p_\gamma - \max(\bar{\lambda} - \gamma y_\gamma, 0) + r_\gamma \lambda_\gamma + J_y(y_\gamma, u_\gamma) &= 0,\\
J_u(y_\gamma, u_\gamma) + \langle A'(u_\gamma)[\cdot] y_\gamma, p_\gamma \rangle - F^\star p_\gamma &= 0,\\
\kappa \lambda_\gamma - p_\gamma + r_\gamma y_\gamma - \xi_\gamma &= 0,\\
\lambda_\gamma \geq 0, \quad \xi_\gamma \geq 0, \quad (\lambda_\gamma, \xi_\gamma)_{L^2} &= 0,\\
(y_\gamma, \lambda_\gamma)_{L^2} \leq \alpha, \quad r_\gamma \geq 0, \quad r_\gamma((y_\gamma, \lambda_\gamma)_{L^2} - \alpha) &= 0,\\
A(u_\gamma) y_\gamma - \lambda_\gamma - f(u_\gamma) &= 0.
\end{aligned}$$

- Study convergence behavior of stationary points of $(P_{\alpha,\gamma})$ for $\gamma \to \infty$ and $\kappa, \alpha \to 0$ with

$$\max\left(\frac{1}{\alpha\sqrt{\gamma}}, \kappa\sqrt{\gamma}\right) \leq C.$$

- .... yields ϵ-almost C-stationarity of limit points.

New stationarity characterizations

ϵ-almost C-stationarity.

The point $(y, u, \lambda) \in H_0^1(\Omega) \times \mathcal{U} \times L^2(\Omega)$ is called ϵ-almost C-stationary if there exist $p \in H_0^1(\Omega)$ and $\mu \in H^{-1}(\Omega)$ and if for every $\epsilon > 0$ there exists $E_\epsilon \subset \Omega^+$ with $\text{meas}(\Omega^+ \setminus E_\epsilon) < \epsilon$ such that

$$
\begin{aligned}
A(u)^\star p - \mu + J_y(y, u) &= 0, \\
J_u(y, u) + \langle A'(u)[\cdot]y, p\rangle - F^\star p &= 0, \\
\langle \mu, p\rangle &\leq 0, \\
p &= 0 \text{ a.e. in } \{\lambda > 0\}, \\
\langle \mu, y\rangle &= 0, \\
\langle \mu, \phi\rangle &= 0 \; \forall \phi \in H_0^1(\Omega), \phi = 0 \text{ a.e. in } \Omega \setminus E_\epsilon, \\
A(u)y - \lambda - f(u) &= 0, \\
\lambda - \max(0, \lambda - \sigma y) &= 0
\end{aligned}
$$

with $\Omega^+ := \{y > 0\}$ and for some arbitrarily fixed real $\sigma > 0$.

New stationarity characterizations

- If $\langle \max(\bar{\lambda} - r_\gamma \lambda_\gamma, 0), y \rangle = 0$ or alternatively $\|y_\gamma - y\|_{L^2} = \mathcal{O}(\gamma^{-1/2})$ as $\gamma \to \infty$ we obtain almost C-stationarity.

The point $(y, u, \lambda) \in H_0^1(\Omega) \times \mathcal{U} \times L^2(\Omega)$ is called almost C-stationary if there exist $p \in H_0^1(\Omega)$ and $\mu \in H^{-1}(\Omega)$ such that

$$
\begin{aligned}
A(u)^\star p - \mu + J_y(y,u) &= 0, \\
J_u(y,u) + \langle A'(u)[\cdot]y, p \rangle - F^\star p &= 0, \\
\langle \mu, p \rangle &\leq 0, \\
p &= 0 \text{ a.e. in } \{\lambda > 0\}, \\
\langle \mu, y \rangle &= 0, \\
\langle \mu, \phi \rangle &= 0 \ \forall \phi \in H_0^1(\Omega), \phi = 0 \text{ a.e. in } \Omega \setminus \Omega^+, \\
&\qquad \phi|_{\Omega^+} \in H_0^1(\Omega^+), \\
A(u)y - \lambda - f(u) &= 0, \\
\lambda - \max(0, \lambda - \sigma y) &= 0
\end{aligned}
$$

with $\Omega^+ := \{y > 0\}$ and for some arbitrarily fixed real $\sigma > 0$.

New stationarity characterizations

- If, moreover, Ω^+ is Lipschitz we obtain C-stationarity.

The point $(y, u, \lambda) \in H_0^1(\Omega) \times \mathcal{U} \times L^2(\Omega)$ is called C-stationary if there exist $p \in H_0^1(\Omega)$ and $\mu \in H^{-1}(\Omega)$ such that

$$
\begin{aligned}
A(u)^\star p - \mu + J_y(y, u) &= 0, \\
J_u(y, u) + \langle A'(u)[\cdot]y, p\rangle - F^\star p &= 0, \\
\langle \mu, p\rangle &\leq 0, \\
p &= 0 \text{ a.e. in } \{\lambda > 0\}, \\
\langle \mu, \phi\rangle &= 0 \ \forall \phi \in H_0^1(\Omega),\ \phi = 0 \text{ a.e. in } \Omega \setminus \Omega^+, \\
A(u)y - \lambda - f(u) &= 0, \\
\lambda - \max(0, \lambda - \sigma y) &= 0
\end{aligned}
$$

with $\Omega^+ := \{y > 0\}$ and for some arbitrarily fixed real $\sigma > 0$.

New stationarity characterizations

If there holds

- $r_\gamma(y_\gamma, v)_{L^2(\mathcal{B})} \to 0 \quad \forall v \in L^2(\mathcal{B})$,
- $r_\gamma(\lambda_\gamma, \phi)_{L^2(\Omega^+ \cup \mathcal{B})} \to 0 \quad \forall \phi \in H_0^1(\Omega)$,

then we obtain strong stationarity.

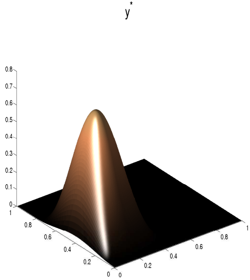
The point $(y, u, \lambda) \in H_0^1(\Omega) \times \mathcal{U} \times L^2(\Omega)$ is called strongly stationary if there exist $p \in H_0^1(\Omega)$ and $\mu \in H^{-1}(\Omega)$ such that

$$\begin{aligned}
A(u)^\star p - \mu + J_y(y, u) &= 0, \\
J_u(y, u) + \langle A'(u) \cdot y, p \rangle - F^\star p &= 0, \\
\langle \mu, p \rangle &\leq 0, \\
p &= 0 \text{ a.e. in } \{\lambda > 0\}, \\
p &\leq 0 \text{ a.e. in } \mathcal{B}, \\
\langle \mu, \phi \rangle &\geq 0 \; \forall \phi \in H_0^1(\Omega),\ \phi \geq 0 \text{ a.e. in } \mathcal{B}, \\
&\qquad \phi = 0 \text{ a.e. in } \Omega \setminus (\Omega^+ \cup \mathcal{B}), \\
A(u)y - \lambda - f(u) &= 0, \\
\lambda - \max(0, \lambda - \sigma y) &= 0.
\end{aligned}$$

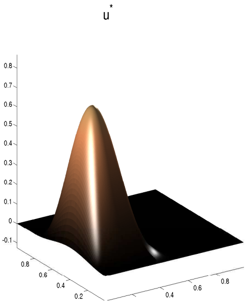
Results I.

Optimal control of the obstacle problem.

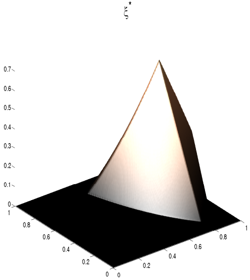
$J(y,u) = \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\delta}{2}\|u\|_{L^2}^2,\ A(u) = -\Delta,\ Fu + g = u + g.$



State y

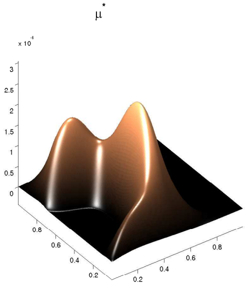


Control u

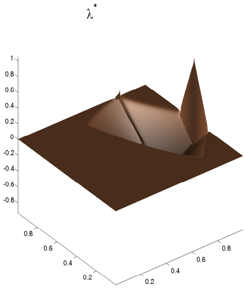


Upper level multiplier λ.

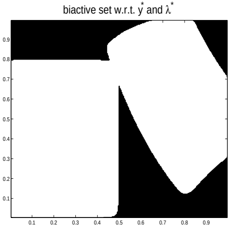
Results I.



MPEC-mult. ($\lambda \geq 0$)



MPEC-mult. μ



$\mathcal{B}$ (black).

Nested iteration

mesh-size $1/h$:	16	32	64	128	256
no. it. :	28	10	11	11	3

Results II.

Lubrication problem.

- Objective functional: tracking type

$$J(y,u) = \tfrac{1}{2}\|y - y_d\|_{L^2}^2 + \tfrac{\delta}{2}\|\nabla u\|_{L^2}^2$$

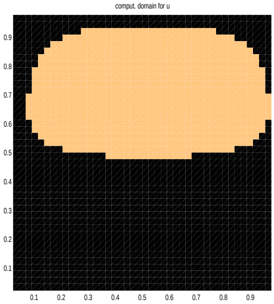
- Parameters:

$$\Omega = (0,1)^2, \quad e(z) = z^3, \quad F(u) = \frac{\partial u}{\partial x_2}, \quad g \equiv 0, \quad \delta = 7.5\text{E-}3.$$

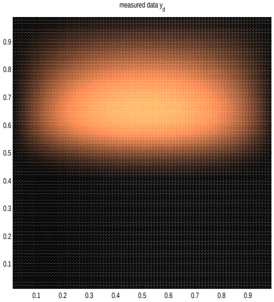
- Initialization: $y^0 \equiv 1$, $u^0 \equiv 10$, $q^0 \equiv 10$, $\lambda^0 \equiv \mu^0 \equiv p^0 \equiv 0$.
- Fixed (exact) parameter u_f:

$$u_f(x_1, x_2) = 1 + 0.5 \cdot \cos(2\pi \cdot x_2)$$

Results II.

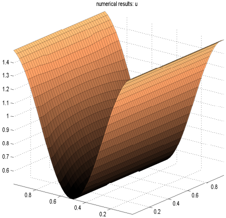


fixed (black) & comput. (orange) domain for u

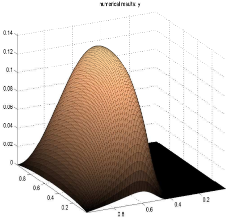


y_d

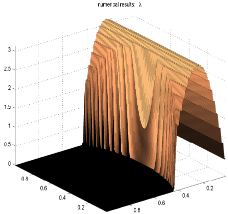
Results II.



u



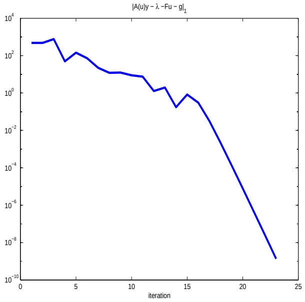
y



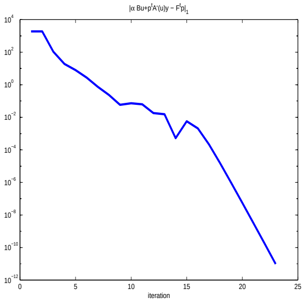
λ

Results II.

Convergence of residuals on finest grid.



primal sys.



adjoint eq.

Smoothing + multigrid approach.

Consider the control constrained version of the MPCC:

$$\begin{aligned} \text{minimize} \quad & J(y,u) = \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\delta}{2}\|u\|_{L^2}^2 \quad \text{over } (y,u) \in H_0^1(\Omega) \times L^2(\Omega) \\ \text{s. t.} \quad & Ay - \lambda = u + g, \\ & y \geq 0, \quad \lambda \geq 0, \quad (y,\lambda) = 0, \\ & a \leq u \leq b \text{ a.e. in } \Omega. \end{aligned}$$

Smoothing: Replace VI by

$$Ay - \gamma \max_{\epsilon}^{\text{loc,glob}}(0, -y) = u + g,$$

where $\max_{\epsilon}^{\bullet}(0, \cdot)$ is (at least) C^1, $\gamma > 0$.

Smoothing + multigrid approach.

→ Apply standard KKT-theory in Banach space.

Let α, $\epsilon > 0$ and $(y, u) \in H_0^1(\Omega) \times L^2(\Omega)$ be an optimal solution of the smoothed MPCC. Then there exists an adjoint state $p \in H_0^1(\Omega)$, and multipliers $\phi^a, \phi^b \in L^2(\Omega)$ such that

$$\begin{aligned} y + A^* p + \gamma \max{}'_{\epsilon}(0, -y)p &= y_d, \\ \delta u - p + \phi^b - \phi^a &= 0, \\ Ay - \gamma \max{}_{\varepsilon}(0, -y) &= u + g, \\ u - a \geq 0 \text{ a.e.}, \ \phi^a \geq 0 \text{ a.e.}, (u - a, \phi^a) &= 0, \\ b - u \geq 0 \text{ a.e.}, \ \phi^b \geq 0 \text{ a.e.}, (b - u, \phi^b) &= 0. \end{aligned}$$

Complementarity system for control is equivalent to

$$\phi = \max(0, \phi + \sigma(u - b)) + \min(0, \phi + \sigma(u - a))$$

for arbitrarily fixed $\sigma > 0$ and $\phi = \phi^b - \phi^a$.

Smoothing + multigrid approach.

- $\frac{\epsilon(\gamma)}{\gamma} \to 0$: Stationary points of smoothed MPCC yield ϵ-almost C stationary points for original MPCC.
- Solution by nonlinear multigrid methods (FAS or MG of second kind):
  - Full Approximation Scheme.
  - MultiGrid of second kind.

Smoothing + multigrid approach.

- FAS: Stationarity system for smoothed MPCC reduces to

$$\begin{aligned} y + A^* p + \gamma \max_\epsilon'(0,-y)p &= y_d, \\ Ay - \gamma \max_\epsilon(0,-y) - \delta^{-1}p + (\delta^{-1}p - b)^+ - (-\delta^{-1}p + a)^+ &= g, \end{aligned}$$

  where $(\cdot)^+ = \max(0,\cdot)$.
- Restriction: Full weighting for residuals; straight injection for states and adjoints.
- Smoothing: Collective nonlinear Gauss-Seidel

$$\begin{aligned} y_{ij} + \tfrac{4}{h^2} p_{ij} + \tfrac{1}{\alpha} \max{}'_{\varepsilon}(0,-y_{ij}) p_{ij} &= B_{ij}, \\ \tfrac{4}{h^2} y_{ij} - \tfrac{1}{\alpha} \max{}_{\varepsilon}(0,-y_{ij}) - \tfrac{1}{\nu} p_{ij} + \max(0, \tfrac{1}{\nu} p_{ij} - b_{ij}) + \max(0, -\tfrac{1}{\nu} p_{ij} + a_{ij}) &= A_{ij} \end{aligned}$$

  with

$$\begin{aligned} B_{ij} &= (y_d)_{ij} + \frac{1}{h^2}(p_{i+1,j} + p_{i-1,j} + p_{i,j+1} + p_{i,j-1}), \\ A_{ij} &= f_{ij} + \frac{1}{h^2}(y_{i+1,j} + y_{i-1,j} + y_{i,j+1} + y_{i,j-1}) \end{aligned}$$

  for all inner grid-points (i,j).
- FAS incorporated in "nested iteration" (coarse-to-fine gird refinement).

Smoothing + multigrid approach.

- MG of second kind: Choosing $\sigma = \delta$ yields

$$u = P_{[a,b]}(\delta^{-1} p(u)) =: F(u),$$

where $p(u)$ solves the adjoint equation for given $y(u)$, the solution of the primal equation for given u.

MG2: Recursive multi-grid algorithm of second kind

DATA: perturbation d^h, initial value u^h.

IF $h = h_{max}$ then solve $u^h = F^h(u^h)$ exactly, ELSE

1. $u^h := F^h(u^h) + d^h$ (smoothing).
2. $d^H := I_h^H \left(u^h - F^h(u^h) - d^h\right)$.
3. Apply MG2$(d^H, \tilde{u}^H)$ on coarser grid twice to obtain u^H. Here $\tilde{u}^H$ is the solution of $u^H = F^H(u^H)$
4. $u^h := u^h - I_H^h \left(u^H - \tilde{u}^H\right)$ (coarse grid correction).

ENDIF

Smoothing + multigrid approach.

Theorem. Let u^* satisfy $u^* = F(u^*)$. If

$$\text{meas}\{|\delta^{-1}p(u^*) - a| \leq \tau\} \to 0 \quad \text{as } \tau \to 0,$$
$$\text{meas}\{|\delta^{-1}p(u^*) - b| \leq \tau\} \to 0 \quad \text{as } \tau \to 0$$

and analogously for the discrete problems for all sufficiently small mesh sizes h, then MG2 converges for sufficiently small $h_{\max}$.

Smoothing + multigrid approach.

Table: Numerical performance of the FAS-method with coupling of α and h.

	Example LS		Example Deg.	
#inner gridpts.	max$_{\varepsilon}^{g}$	max$_{\varepsilon}^{l}$	max$_{\varepsilon}^{g}$	max$_{\varepsilon}^{l}$
15	6	9	5	7
31	3	3	3	2
63	3	3	3	3
127	3	3	5	3
255	2	2	2	2
total	17	20	18	17
CPU ratio FAS/MG2	0.4311	0.3738	0.2324	0.2713

Smoothing + multigrid approach.

	constr.		unconstr.	
δ	V-cyc.	W-cyc.	V-cyc.	W-cyc.
1	0.0433	0.0219	0.1217	0.0220
1e-1	0.0370	0.0238	0.1615	0.0251
1e-2	0.2777	0.0234	0.7792	0.0724

Table: Estimated FAS convergence factors for lack of strict complementarity.

	constr.		unconstr.	
δ	V-cyc.	W-cyc.	V-cyc.	W-cyc.
1	0.0387	0.0219	0.0389	0.0218
1e-1	0.0396	0.0217	0.0377	0.0218
1e-2	0.1312	0.0220	0.2425	0.0228

Table: Estimated FAS convergence factors for degenerate problem.

$\gamma = 1E3$ and $\epsilon = 1E-3$ fixed respectively.