A Semismooth Newton Method for a Gradient Constrained Minimization Problem

Robert Nürnberg
Department of Mathematics, Imperial College London

In collaboration with

Michael Hintermüller
Humboldt-Universität Berlin

Model Problem

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with boundary $\Gamma = \partial\Omega$.

Assume that either Γ is smooth or that Ω is convex with Lipschitz boundary.

Consider

$$\text{(P)} \qquad \text{minimize} \qquad \frac{1}{2}\int_\Omega |\nabla y|^2 \, \mathrm{d}x - \int_\Omega f\, y \, \mathrm{d}x \qquad \text{over } y \in H_0^1(\Omega)$$
$$\text{subject to } |\nabla y| \le \psi \text{ almost everywhere on } \Omega\,,$$

where $f \in L^p(\Omega)$, $p \ge 2$, and $\psi \in L^r(\Omega)$, $r > 2$, with

$$\psi(x) \ge \delta > 0 \quad \text{for almost all } x \in \Omega.$$

E.g. for simplicity $\psi \equiv 1$. $\Rightarrow$ **Elasto-Plastic Torsion Problem.**

(Studied extensively by Glowinski, Lions, Trémolières, . . .)

Let $(\cdot,\cdot)$ denote the L^2-inner product over Ω.

Let $K := \{\eta \in H_0^1(\Omega) : |\nabla\eta| \le \psi$ almost everywhere on $\Omega\}$.

$$\text{(P)} \qquad \text{minimize} \qquad J(y) = \tfrac{1}{2}(\nabla y, \nabla y) - (f, y) \qquad \text{over } y \in K\,.$$

There exists a unique solution y^* to (P), which is also the unique solution in K of the Variational Inequality

$$\text{(VI)} \qquad (\nabla y, \nabla[v - y]) - (f, v - y) \ge 0 \qquad \forall\ v \in K\,.$$

Moreover, the assumptions on Ω yield that $y^* \in K \cap W^{2,p}(\Omega)$.

Finite Element Approximation

Assume that Ω is polyhedral and let $\{\mathcal{T}^h\}_{h>0}$ be a family of quasi-uniform partitionings of Ω into disjoint open triangles κ with $h_\kappa := \text{diam}(\kappa)$ and $h := \max_{\kappa\in\mathcal{T}^h} h_\kappa$. Hence $\overline{\Omega} = \cup_{\kappa\in\mathcal{T}^h}\overline{\kappa}$.

Let $V^h = \{\eta \in C^0(\overline{\Omega}) : \eta|_\kappa$ is linear $\forall\ \kappa \in \mathcal{T}^h, \eta|_\Gamma = 0\} \subset H^1_0(\Omega)$.

Let $K^h := K \cap V^h$. (Assume from now on that $\psi|_\kappa$ is constant for all $\kappa \in \mathcal{T}^h$.)

Then the standard FE approximation to (P) and (VI) is:

$$(\text{P}^h) \qquad \text{minimize} \qquad J(y_h) = \tfrac{1}{2}(\nabla y_h, \nabla y_h) - (f, y_h) \qquad \text{over } y_h \in K^h .$$

There exists a unique solution y_h^* to (P^h), which is also the unique solution in K^h of the Variational Inequality

$$(\text{VI}^h) \qquad (\nabla y_h, \nabla[v_h - y_h]) - (f, v_h - y_h) \geq 0 \qquad \forall\ v_h \in K^h .$$

Moreover, the assumptions on $\mathcal{T}^h$ yield that $\|y_h^* - y^*\|_1 \leq C\, h^{\frac{1}{2}-\frac{1}{p}}$.

[Suboptimal error bound caused by fact that in general $I_h\,\eta \notin K^h$ for $\eta \in H^2(\Omega) \cap K$.]

Alternative Approximations

There exist other finite element approximations, e.g. based on an equivalent formulation for $\sigma = \nabla y$.

Then (P) is equivalent to

$$\text{minimize} \qquad \tfrac{1}{2}(\sigma,\sigma) - (\varphi,\sigma) \qquad \text{over } \sigma \in \widetilde{K}\,;$$

where $\operatorname{rot}\varphi = -f$ and

$$\widetilde{K} := \{\sigma \in [L^2(\Omega)]^2 : |\sigma| \leq \psi \text{ a.e. on } \Omega \text{ and } (\sigma, \nabla v) = 0 \quad \forall\ v \in H^1(\Omega)\}\,.$$

For the natural finite element approximation

$$\sigma_h \in P_h := \{\eta \in L^1(\Omega) : \eta|_\kappa \text{ is constant } \forall\ \kappa \in \mathcal{T}^h\}$$

one can show that $\|\sigma_h - \sigma\| \leq C\,h$. See Falk, Mercier (77).

However, in this talk we only consider the conforming approximation

$$(\mathrm{P}^h) \qquad \text{minimize} \qquad J(y_h) = \tfrac{1}{2}(\nabla y_h, \nabla y_h) - (f, y_h) \qquad \text{over } y_h \in K^h\,.$$

Let $\mathcal{J}$ be the set of nodes of $\mathcal{T}^h$ and let $\{p_j\}_{j\in\mathcal{J}}$ be the coordinates of these nodes. Let $\mathcal{J}_0 := \{j \in \mathcal{J} : p_j \notin \partial\Omega\} = \{1, \ldots, N\}$. Let $\{\chi_j\}_{j\in\mathcal{J}_0}$ be the standard basis functions for V^h.

1. Nonlinear Overrelaxation directly applied to (P^h).

Set $y_h^{(0)} = 0 \in K^h$. For $i \geq 0$:

(i) Let $y_h^{(i+1,0)} = y_h^{(i)}$.

(ii) For $k = 1 \to N$ set $y_h^{(i+1,k)} = \Pi_{K^h} \widehat{y}_h^{(i+1,k)}(\alpha)$, where $\widehat{y}_h^{(i+1,k)}(\alpha) = y_h^{(i+1,k-1)} + \alpha\, \chi_k$ is such that $J(\widehat{y}_h^{(i+1,k)}(\alpha)) \leq J(\widehat{y}_h^{(i+1,k)}(\beta))$ for all $\beta \in \mathbb{R}$.

(iii) Set $y_h^{(i+1)} = y_h^{(i+1,N)}$.

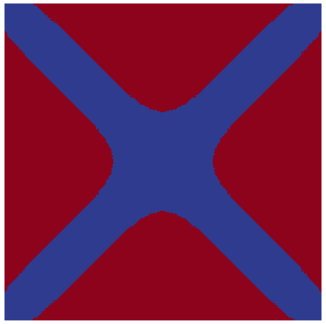
Here $\Pi_{K^h} : V^h \to K^h$ is the orthogonal projection onto K^h.

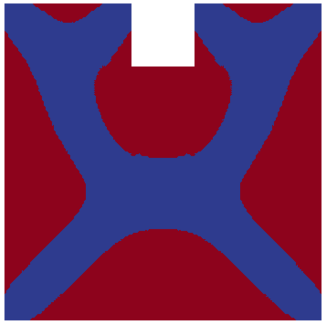
Hence in (ii), at each vertex, we need a local projection onto K^h. The number of constraints to be satisfied there are given by the number of triangles that meet at the vertex.

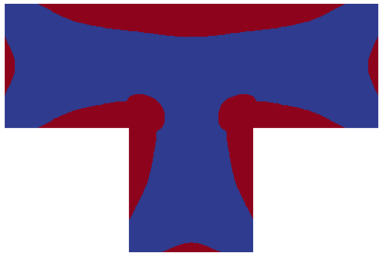
Nonlinear SOR

Difficult to implement. Highly mesh dependent.

Example: Let $f = 10$, $\psi \equiv 1$.







h	# iterations
1/4	16
1/8	28
1/16	139
1/32	520
1/64	1984
1/128	7428
1/256	27608

h	# iterations
1/5	13
1/10	39
1/20	179
1/40	702
1/80	2612
1/160	9731
1/320	36054

h	# iterations
1/6	12
1/12	47
1/24	181
1/48	657
1/96	2414
1/192	8794
1/384	31659

Iteration counts for nonlinear SOR. $\omega = 1$, $tol = 10^{-9}$.

2. Penalization method.

For $\varepsilon \to 0$ consider

$$\text{minimize} \qquad J_\varepsilon(y_h) := J(y_h) + \frac{1}{4\varepsilon}\int_\Omega [|\nabla y_h|^2 - \psi^2]_+^2 \,\mathrm{d}x \qquad \text{over } y_h \in V^h\,;$$

i.e. find $y_{h,\varepsilon} \in V^h$ such that

$$(\nabla y_{h,\varepsilon}, \nabla v_h) + \frac{1}{\varepsilon}([|\nabla y_{h,\varepsilon}|^2 - \psi^2]_+ \nabla y_{h,\varepsilon}, \nabla v_h) = (f, v_h) \qquad \forall\ v_h \in V^h\,.$$

It can be shown that $y_{h,\varepsilon} \to y_h^* \in K^h$ as $\varepsilon \to 0$.

Highly nonlinear equation — difficult to solve.

In practice, our method will be very similar to 2.

3. Duality method.

Consider the Lagrangian

$$\mathcal{L}(y_h, p_h) := J(y_h) + \frac{1}{2}\int_\Omega p_h\,(|\nabla y_h|^2 - \psi^2)\,\mathrm{d}x$$

over $V_h \times P_h^+$, where $P_h^+ := \{\eta \in L^1(\Omega) : 0 \le \eta|_\kappa$ is constant $\forall\ \kappa \in \mathcal{T}^h\}$.

Then $\mathcal{L}$ admits a saddle point $(y_h, p_h) \in V_h \times P_h^+$ with $y_h = y_h^*$ and $p_h\,(|\nabla y_h^*|^2 - \psi^2) = 0$.

Hence (y_h^*, p_h) can be found with a Uzawa iterative algorithm:

Set $p_h^{(0)} = 0 \in P_h^+$. For $n \ge 0$:

(i) Given $p_h^{(n)} \in P_h^+$ let $y_h^{(n)}$ minimize $\mathcal{L}(y_h, p_h^{(n)})$ over V_h. I.e. $y_h^{(n)} \in V^h$ is such that

$$([1 + p_h^{(n)}]\,\nabla y_h^{(n)}, \nabla v_h) = (f, v_h) \qquad \forall\ v_h \in V_h\,.$$

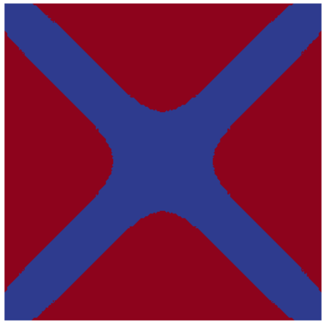
(ii) $p_h^{(n+1)} = [p_h^{(n)} + \rho_n\,(|\nabla y_h^{(n)}|^2 - \psi^2)]_+$ for $\rho_n > 0$.

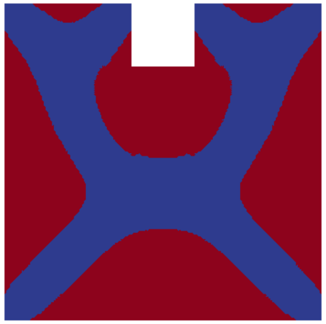
For ρ_n suitably chosen it can be shown that $y_h^{(n)} \to y_h^*$ as $n \to \infty$.

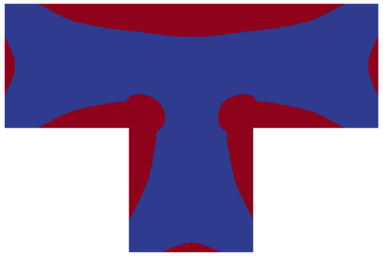
Uzawa Algorithm

Straightforward to implement. In general mesh dependent.

Example: Let $f = 10$, $\psi \equiv 1$.







h	# iterations
1/4	768
1/8	184
1/16	540
1/32	1339
1/64	926
1/128	403
1/256	386

h	# iterations
1/5	222
1/10	622
1/20	4619
1/40	12903
1/80	29916
1/160	41503
1/320	44906

h	# iterations
1/6	154
1/12	173
1/24	416
1/48	5346
1/96	3146
1/192	8082
1/384	17332

Iteration counts for Uzawa. $\rho_n = \rho = 1$, $tol = 10^{-9}$.

We would like to introduce an alternative solution method, which ideally will exhibit mesh independent convergence rates.

Here a natural idea is to reformulate the optimality system for (P) as a nonsmooth system of equations, and then apply a **semismooth Newton method**.

If we can prove that such a method converges in function space, then a conforming finite element approximation will inherit the (mesh-independent) convergence properties.

However, a lack of regularity means that the original problem first needs to be regularized.

In order to motivate our approach for (P), we can first consider the standard obstacle problem.

Obstacle problem

For motivational purposes, consider the standard obstacle problem. See Hintermüller, Ito, Kunisch (2003).

$$\text{(Q)} \qquad \text{minimize} \qquad J(y) = \tfrac{1}{2}(\nabla y, \nabla y) - (f, y) \qquad \text{over } y \in K_Q\,,$$

where $K_Q := \{\eta \in H_0^1(\Omega) : \eta \le \psi$ almost everywhere on $\Omega\}$.

Then the standard FE approximation to (Q) is:

$$(\mathrm{Q}^h) \qquad \text{minimize} \qquad J(y_h) = \tfrac{1}{2}(\nabla y_h, \nabla y_h) - (f, y_h) \qquad \text{over } y_h \in K_Q^h\,,$$

where $K_Q^h := K_Q \cap V^h$. (Assume for simplicity that $\psi \in V^h$.)

The optimality system for (Q^h) is given by

$$\begin{cases} (\nabla y_h, \nabla v_h) + (\lambda_h, v_h)^h = (f, v_h) & \forall\ v_h \in V^h\,, \\ y_h \in K_Q^h,\ \ \lambda_h \in V^h,\ \ \lambda_h \ge 0,\ \ (\lambda_h, y_h - \psi)^h = 0\,. \end{cases}$$

The complementarity condition can be equivalently expressed as

$$\mathcal{C}_c(y_h, \lambda_h) = 0, \quad \text{where} \quad \mathcal{C}_c(y_h, \lambda_h) = \lambda_h - [\lambda_h + c\,(y_h - \psi)]_+\,,$$

for each $c > 0$.

Obstacle problem

A solution to the optimality system

$$(\mathrm{OS}^h) \qquad \begin{cases} (\nabla y_h, \nabla v_h) + (\lambda_h, v_h)^h = (f, v_h) & \forall\ v_h \in V^h\,, \\ \lambda_h - [\lambda_h + c\,(y_h - \psi)]_+ = 0\,, \end{cases}$$

can be found with the following **primal-dual active set algorithm**:

Initialize $y_h^{(0)}$, $\lambda_h^{(0)}$. For $k \geq 0$

(i) Set $\mathcal{I}_k := \{j : [\lambda_h^{(k)} + c\,(y_h^{(k)} - \psi)](p_j) \leq 0\}$, $\mathcal{A}_k := \mathcal{J}_0 \setminus \mathcal{I}_k = \{j : [\lambda_h^{(k)} + c\,(y_h^{(k)} - \psi)](p_j) > 0\}$.

(ii) Solve

$$\begin{cases} (\nabla y_h^{(k+1)}, \nabla v_h) + (\lambda_h^{(k+1)}, v_h)^h = (f, v_h) & \forall\ v_h \in V^h\,, \\ y_h^{(k+1)}(p_j) = \psi(p_j) \quad \forall\ j \in \mathcal{A}_k\,, & \lambda_h^{(k+1)}(p_j) = 0 \quad \forall\ j \in \mathcal{I}_k\,. \end{cases}$$

Hintermüller, Ito, Kunisch (2003) showed that this algorithm is equivalent to a **semismooth Newton method** applied to (OS^h). Moreover, the algorithm converges locally superquadratically since $[\cdot]_+ : \mathbb{R} \to \mathbb{R}$ is *slantly differentiable*. (In fact, here convergence is global.)

[F slantly differentiable $\iff$ $\lim_{h\to 0} \frac{1}{\|h\|} \|F(x+h) - F(x) - G(x+h)\,h\| = 0$]

However, the method breaks down in function space. Formally, a solution to the optimality system

$$(\mathrm{OS}) \qquad \begin{cases} (\nabla y, \nabla v) + (\lambda, v) = (f, v) & \forall\ v \in H_0^1(\Omega)\,, \\ \lambda - [\lambda + c\,(y - \psi)]_+ = 0\,, \end{cases}$$

can be found with the following primal-dual active set algorithm:

Initialize $y^{(0)}$, $\lambda^{(0)}$. For $k \geq 0$

(i) Set $\mathcal{I}_k := \{x : [\lambda^{(k)} + c\,(y^{(k)} - \psi)](x) \leq 0\}$, $\mathcal{A}_k := \Omega \setminus \mathcal{I}_k$ $= \{x : [\lambda^{(k)} + c\,(y^{(k)} - \psi)](x) > 0\}$.

(ii) Solve

$$\begin{cases} (\nabla y^{(k+1)}, \nabla v) + \langle \lambda^{(k+1)}, v \rangle_{H^{-1}, H_0^1} = (f, v) & \forall\ v \in H_0^1(\Omega)\,, \\ y^{(k+1)} = \psi \quad \text{on } \mathcal{A}_k\,, \qquad \lambda^{(k+1)} = 0 \quad \text{on } \mathcal{I}_k\,. \end{cases}$$

However, the Lagrange multiplier $\lambda^{(k+1)}$ in (ii) is in general only a distribution in $H^{-1}(\Omega)$, and so step (i) in the above algorithm is no longer meaningful.

Correspondingly, the associated semismooth Newton method cannot be shown to converge, since $[\cdot]_+ : L^q(\Omega) \to L^p(\Omega)$ is *slantly differentiable* iff $p < q$.

In order to obtain a method with potentially mesh-size independent convergence rates, it is desirable to derive a solution algorithm in function space.

For the obstacle problem this was achieved in Hintermüller, Kunisch (2006) by considering the following Moreau-Yosida regularization. (See also Ito, Kunisch (2003).)

For $\gamma \to \infty$ consider

$$\text{minimize} \qquad \hat{J}_\gamma(y) := J(y) + \frac{1}{2\gamma} \int_\Omega [\hat{\lambda} + \gamma\,(y - \psi)]_+^2 \, \mathrm{d}x \qquad \text{over } y \in H_0^1(\Omega),$$

where $\hat{\lambda} \in L^2(\Omega)$, $\hat{\lambda} \geq 0$.

For a fixed $\gamma > 0$, the optimality system is given by

$$(\mathrm{OS}_\gamma) \qquad \begin{cases} (\nabla y, \nabla v) + (\lambda, v) = (f, v) & \forall\ v \in H_0^1(\Omega)\,, \\ \lambda - [\hat{\lambda} + \gamma\,(y - \psi)]_+ = 0\,, \end{cases}$$

where $(y, \lambda) \in H_0^1(\Omega) \times L^2(\Omega)$.

Obstacle problem

$$(\mathrm{OS}_\gamma) \qquad \begin{cases} (\nabla y, \nabla v) + (\lambda, v) = (f, v) & \forall\, v \in H_0^1(\Omega)\,, \\ \lambda - [\hat{\lambda} + \gamma\,(y - \psi)]_+ = 0\,. \end{cases}$$

For fixed γ, (OS$_\gamma$) can be solved with a semismooth Newton method to obtain y_γ. (here globally convergent in function space)

As $\gamma \to \infty$, $y_\gamma \to y^\star$, the solution of (Q).

If $\hat{\lambda} = 0$, then $y^\star \le y_{\gamma_2} \le y_{\gamma_1}$ for $\gamma_1 \le \gamma_2$. (infeasible case)

If $\hat{\lambda} \ge 0$ sufficiently large, then $y_{\gamma_1} \le y_{\gamma_2} \le y^\star \le \psi$ for $\gamma_1 \le \gamma_2$. (feasible case)

Based on a model function, Hintermüller, Kunisch (2006) derive in the infeasible case a γ_k-update strategy that implies q-superlinear convergence of y_{γ_k} to $y^\star$ in $H_0^1(\Omega)$.

Gradient constrained minimization problem

Return to the original problem:

$$(\mathrm{P}) \qquad \text{minimize} \qquad J(y) = \tfrac{1}{2}(\nabla y, \nabla y) - (f, y) \qquad \text{over } y \in K\,,$$

where $K = \{\eta \in H_0^1(\Omega) : |\nabla \eta| \le \psi$ almost everywhere on $\Omega\}$, with standard FE approximation

$$(\mathrm{P}^h) \qquad \text{minimize} \qquad J(y_h) = \tfrac{1}{2}(\nabla y_h, \nabla y_h) - (f, y_h) \qquad \text{over } y_h \in K^h\,.$$

- It is not possible to derive a primal-dual active set algorithm for (P^h), as the boundary conditions for the solutions y_h^k on the inactive set are not known.

- Similarly, applying a semismooth Newton method to the optimality system of (P^h) would in general yield only local convergence, with mesh-dependent estimates.

The aim is to derive a solution method in function space. Thus conforming discretizations would converge with the same rate for sufficiently small mesh-sizes.

$$\text{(P)} \qquad \text{minimize} \qquad J(y) = \tfrac{1}{2}(\nabla y, \nabla y) - (f, y) \qquad \text{over } y \in K\,,$$

where $K = \{\eta \in H_0^1(\Omega) : |\nabla \eta| \leq \psi$ almost everywhere on $\Omega\}$.

The optimality system for (P) is given by

$$\text{(OS)} \qquad \begin{cases} (\nabla y, \nabla v) + (\lambda \dfrac{\nabla y}{|\nabla y|}, \nabla v) = (f, v) & \forall\ v \in H_0^1(\Omega)\,, \\ \lambda - [\lambda + c\,(|\nabla y| - \psi)]_+ = 0 & \text{in } L^2(\Omega)\,, \end{cases}$$

for any $c > 0$.

On recalling that $[\cdot]_+ : L^s(\Omega) \to L^s(\Omega)$, with $1 \leq s \leq \infty$, is not slantly differentiable, the semismooth Newton iteration for solving (OS) need not converge.

Hence we employ the following Moreau-Yosida regularization, similarly to Hintermüller, Kunisch (2006) for the obstacle problem.

For $\gamma > 0$ consider

$$(\mathsf{P}_\gamma) \quad \text{minimize } J_\gamma(y) := J(y) + \frac{1}{2\gamma}\int_\Omega [\hat{\lambda} + \gamma\,(|\nabla y| - \psi)]_+^2 \,\mathrm{d}x \quad \text{over } y \in H_0^1(\Omega)\,,$$

where $\hat{\lambda} \in L^2(\Omega)$, $\hat{\lambda} \geq 0$.

The first order optimality condition for (P_γ) can be expressed as

$$(\mathsf{OS}_\gamma) \qquad (\nabla y_\gamma, \nabla v) + (\underbrace{[\hat{\lambda} + \gamma\,(|\nabla y_\gamma| - \psi)]_+}_{\lambda_\gamma} \frac{\nabla y_\gamma}{|\nabla y_\gamma|}, \nabla v) = (f, v) \quad \forall\, v \in H_0^1(\Omega)\,.$$

It is straightforward to show that $y_\gamma \to y^*$, the solution of (P), in $H_0^1(\Omega)$ as $\gamma \to \infty$.

Moreover, it can be shown that (OS_γ) is slantly differentiable, and so the semismooth Newton method is guaranteed to converge locally q-superlinearly.

For simplicity, consider the infeasible case ($\hat{\lambda} = 0$) from now on.

Then, in particular, the mapping $q : L^r(\Omega)^n \to L^2(\Omega)^n$ with

$$q(u)(x) = [|u(x)| - \psi(x)]_+ \frac{u(x)}{|u(x)|} = \begin{cases} \left(1 - \frac{\psi(x)}{|u(x)|}\right) u(x) & |u(x)| > \psi(x)\,, \\ 0 & |u(x)| \leq \psi(x)\,, \end{cases}$$

is Newton differentiable for $r > 2$.

A particular Newton-derivative is given by

$$Q(u)(x) = \begin{cases} \left(1 - \frac{\psi(x)}{|u(x)|}\right) \mathrm{id} + \psi(x) \frac{u(x)u(x)^\top}{|u(x)|^3} & |u(x)| > \psi(x)\,, \\ 0 & |u(x)| \leq \psi(x)\,. \end{cases}$$

Hence the semismooth Newton method applied to (OS_γ) for finding the solution y_γ of (P_γ) can be formulated as follows.

Let $y_\gamma^{(0)} \in H_0^1(\Omega) \cap W^{1,r}(\Omega) =: W_0^{1,r}(\Omega)$, $r > 2$. For $k \geq 0$:

(i) Find $\delta_y^{(k)} \in H_0^1(\Omega)$ such that

$$(\nabla \delta_y^{(k)}, \nabla v) + \gamma\,(Q(\nabla y_\gamma^{(k)})\,\nabla \delta_y^{(k)}, \nabla v)$$
$$= -(\nabla y_\gamma^{(k)}, \nabla v) + (f, v) - \gamma\,(q(\nabla y_\gamma^{(k)}), \nabla v) \quad \forall\, v \in H_0^1(\Omega).$$

(ii) $y_\gamma^{(k+1)} := y_\gamma^{(k)} + \delta_y^{(k)}$.

Regularity theory for (i) yields that if $y_\gamma^{(k)} \in W_0^{1,r}(\Omega)$, $r > 2$, then $\delta_y^{(k)} \in W_0^{1,r^*}(\Omega)$ with $r^* \in (2, r]$. Hence $y_\gamma^{(k+1)} \in W_0^{1,r^*}(\Omega)$ and this allows us to show that q is Newton-differentiable at $\nabla y_\gamma^{(k+1)}$, and consequently that the semismooth Newton method converges locally.

Our FE approximation to the Newton iteration is as follows.

Let $\mathcal{I}^h : L^1(\Omega) \to D^h$ be the orthogonal projection onto the space

$$D^h := \{\chi \in L^1(\Omega) : \chi|_\kappa \text{ is constant a.e. in } \kappa \quad \forall\ \kappa \in \mathcal{T}^h\}$$

of piecewise constant functions, i.e.

$$(\mathcal{I}^h \chi)|_\kappa = \frac{1}{|\kappa|} \int_\kappa \chi \, \mathrm{d}x \qquad \forall\ \kappa \in \mathcal{T}^h .$$

We then define the mapping $q^h : [D^h]^n \to [D^h]^n$ such that

$$q^h(U)(x) = \begin{cases} \left(1 - \frac{\mathcal{I}^h \psi(x)}{|U(x)|}\right) U(x) & |U(x)| > \mathcal{I}^h \psi(x)\,, \\ 0 & |U(x)| \leq \mathcal{I}^h \psi(x)\,. \end{cases}$$

Moreover, let $Q^h : [D^h]^n \to [D^h]^{n\times n}$ be defined such that

$$Q^h(U)(x) := \begin{cases} \left(1 - \frac{\mathcal{I}^h \psi(x)}{|U(x)|}\right) \mathrm{id} + \mathcal{I}^h \psi(x) \frac{U(x)U(x)^\top}{|U(x)|^3} & |U(x)| > \mathcal{I}^h \psi(x)\,, \\ 0 & |U(x)| \leq \mathcal{I}^h \psi(x)\,. \end{cases}$$

Then the discrete Newton iteration is defined as follows.

Let $Y_\gamma^{(0)} \in V^h$. For $k \geq 0$:

(i) Find $\delta_Y \in V^h$ such that

$$(\nabla \delta_Y, \nabla \chi) + \gamma\, (Q^h(\nabla Y_\gamma^{(k)})\, \nabla \delta_Y, \nabla \chi)$$
$$= -(\nabla Y_\gamma^{(k)}, \nabla \chi) + (f, \chi) - \gamma\, (q^h(\nabla Y_\gamma^{(k)}), \nabla \chi) \quad \forall\, \chi \in V^h\,.$$

(ii) $Y_\gamma^{(k+1)} := Y_\gamma^{(k)} + \delta_Y\,.$

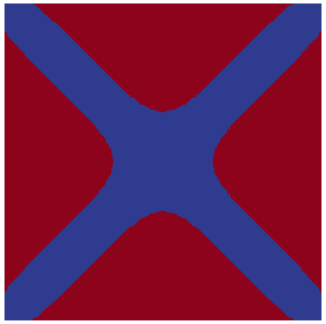
On the discrete level, the above solves $F^h(Y) = 0$, where for $Y \in V^h$ the functional $F^h : V^h \to \mathbb{R}^N$ is defined by

$$[F^h(Y)]_j = (\nabla Y, \nabla \chi_j) - (f, \chi_j) + \gamma\, (q^h(\nabla Y), \nabla \chi_j) \qquad \forall\, j \in \mathcal{J}_0\,.$$

Study number of Newton iterations for varying γ and h. As initial guess $Y_\gamma^{(0)} \in V^h$ we choose the solution of

$$(\nabla Y_\gamma^{(0)}, \nabla \chi) = (f, \chi) \qquad \forall\ \chi \in V^h\,.$$

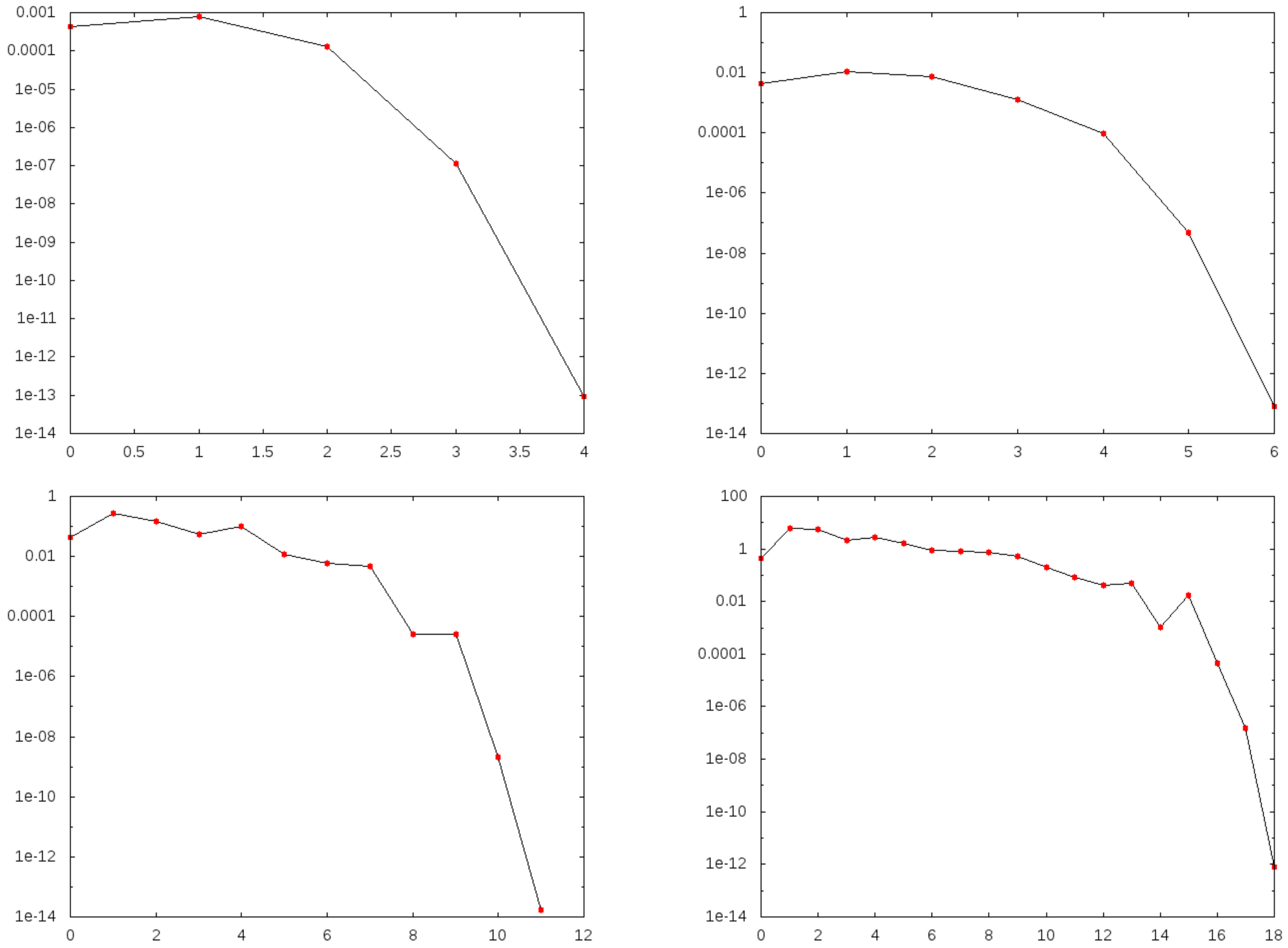
Example: Let $f = 10$, $\psi \equiv 1$.



We observe locally superlinear convergence in practice. Example plots of $\|F^h(Y_\gamma^{(k)})\|_\infty$ for $h = 1/128$ and $\gamma = 10^l$, $l = 0 \to 3$, follow.

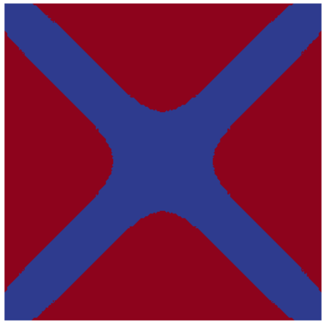
Local Convergence of Newton Iteration

$\|F^h(Y_\gamma^{(k)})\|_\infty$ for $h = 1/128$ and $\gamma = 10^l$, $l = 0 \to 3$. (tol $= 10^{-10}$)



Number of Newton iterations for varying γ and h.

Example: Let $f = 10$, $\psi \equiv 1$.



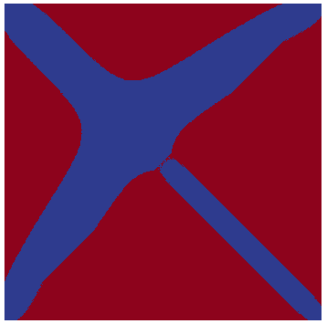
	# iterations			
h	$\gamma = 1$	$\gamma = 10$	$\gamma = 1000$	$\gamma = 10000$
1/4	3	5	10	25
1/8	3	5	7	11
1/16	4	5	13	26
1/32	4	6	13	31
1/64	4	6	15	–
1/128	4	6	18	40
1/256	4	7	20	–

tol $= 10^{-10}$

Local Convergence of Newton Iteration

Same behaviour for more complicated situations.

Example: Let $f = 10$, $\psi(x) = \begin{cases} 1 & x_1 < x_2\,, \\ 0.5 & x_1 \geq x_2\,. \end{cases}$

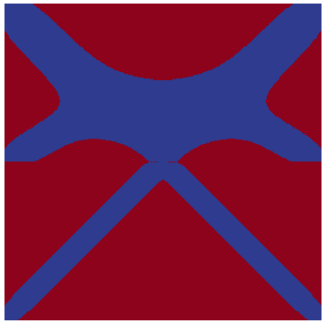


	# iterations			
h	$\gamma = 1$	$\gamma = 10$	$\gamma = 1000$	$\gamma = 10000$
1/4	3	6	12	16
1/8	4	6	14	–
1/16	4	6	15	57
1/32	4	7	22	–
1/64	4	7	24	–
1/128	4	7	22	–
1/256	4	7	26	–

tol $= 10^{-10}$

Same behaviour for more complicated situations.

Example: Let $f = 10$, $\psi(x) = \begin{cases} 1 & x_2 > 0\,, \\ 0.5 & x_2 \le 0\,. \end{cases}$

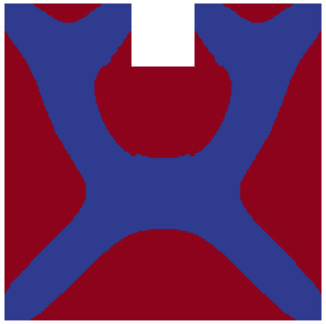


	# iterations			
h	$\gamma = 1$	$\gamma = 10$	$\gamma = 1000$	$\gamma = 10000$
1/4	3	6	19	–
1/8	4	6	15	56
1/16	4	6	16	164
1/32	4	7	21	–
1/64	4	8	20	–
1/128	4	8	22	–
1/256	4	8	–	–

tol $= 10^{-10}$

Similar behaviour for nonconvex domains.

Example: Let $f = 10$, $\psi \equiv 1$.



	# iterations			
h	$\gamma = 1$	$\gamma = 10$	$\gamma = 1000$	$\gamma = 10000$
1/5	3	5	8	12
1/10	4	6	15	43
1/20	4	6	15	–
1/40	4	7	–	–
1/80	4	7	–	–
1/160	5	8	–	–
1/320	5	8	–	–
1/640	5	9	–	–

tol $= 10^{-10}$

Path-following Algorithm

The very local convergence behaviour of the semismooth Newton method motivates a path-following algorithm, similarly to Hintermüller, Kunisch (2006).

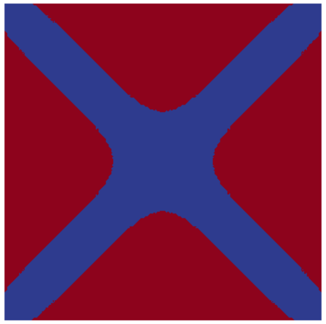
The idea is to compute a sequence of solutions $\{Y_{\gamma_l}\}_{l\geq 0}$ for a sequence $\{\gamma_l\}_{l\geq 0}$ with $\gamma_l \to \infty$ as $l \to \infty$.

The initial guess for the Newton iterations is then $Y^{(0)}_{\gamma_{l+1}} = Y_{\gamma_l}$.

We use the following heuristic path following algorithm.

1. Let $\gamma_0 = 0$ and fix $\overline{\gamma} \in (0, \infty)$.
2. For $l \geq 0$, let Y_{γ_l} denote the converged solution of the semismooth Newton iteration for $\gamma = \gamma_l$, if the iteration converged.
3. If iteration did not converge, then set $\gamma_l := \frac{1}{2}(\gamma_l + \gamma_{l-1})$ and go to 2.
4. Set $\gamma_{l+1} := \min(\overline{\gamma}, a\,\gamma_l + b)$ and let $Y^{(0)}_{\gamma_{l+1}} = Y_{\gamma_l}$.
5. Repeat until $\gamma_l = \overline{\gamma}$.

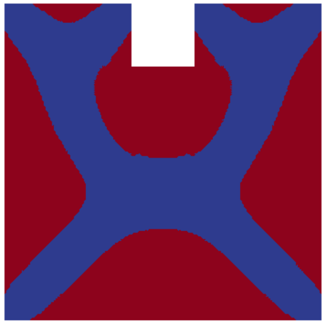
Example: Let $f = 10$, $\psi \equiv 1$.



h	# outer iterations	# inner iterations [max ∥ avg]	(Uzawa)
1/4	14	41 [5 ∥ 2.9]	(768)
1/8	14	35 [5 ∥ 2.5]	(184)
1/16	14	38 [6 ∥ 2.7]	(540)
1/32	14	50 [6 ∥ 3.6]	(1339)
1/64	14	55 [8 ∥ 3.9]	(926)
1/128	14	52 [7 ∥ 3.7]	(403)
1/256	15	55 [7 ∥ 3.7]	(386)

$\overline{\gamma} = 10^8$, $a = 5$, $b = 1$, tol $= 10^{-10}$

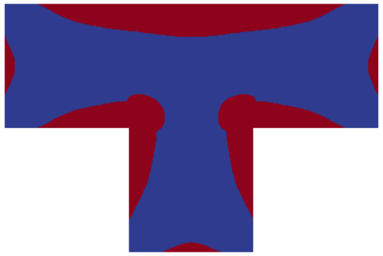
Example: Let $f = 10$, $\psi \equiv 1$.



h	# outer iterations	# inner iterations [max ‖ avg]	(Uzawa)
1/5	94	223 [4 ‖ 2.4]	(222)
1/10	94	248 [4 ‖ 2.6]	(622)
1/20	94	261 [6 ‖ 2.8]	(4619)
1/40	96	300 [6 ‖ 3.1]	(12903)
1/80	109	384 [13 ‖ 3.5]	(29916)
1/160	135	520 [11 ‖ 3.9]	(41503)
1/320	244	942 [18 ‖ 3.9]	(44906)

$\overline{\gamma} = 10^8$, $a = 1.2$, $b = 1$, tol $= 10^{-10}$

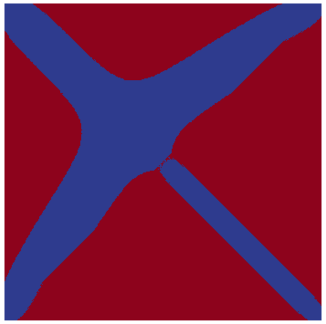
Example: Let $f = 10$, $\psi \equiv 1$.



h	# outer iterations	# inner iterations [max ‖ avg]	(Uzawa)
1/6	94	215 [4 ‖ 2.3]	(154)
1/12	94	226 [4 ‖ 2.4]	(173)
1/24	94	234 [5 ‖ 2.5]	(416)
1/48	94	247 [5 ‖ 2.6]	(5346)
1/96	94	283 [6 ‖ 3.0]	(3146)
1/192	99	340 [8 ‖ 3.4]	(8082)
1/384	114	430 [9 ‖ 3.8]	(17332)

$\overline{\gamma} = 10^8$, $a = 1.2$, $b = 1$, tol $= 10^{-10}$

Example: Let $f = 10$, $\psi(x) = \begin{cases} 1 & x_1 < x_2\,, \\ 0.5 & x_1 \geq x_2\,. \end{cases}$



h	# outer iterations	# inner iterations [max ‖ avg]	(Uzawa)
1/4	28	75 [4 ‖ 2.7]	(707)
1/8	28	94 [5 ‖ 3.4]	(24374)
1/16	28	90 [9 ‖ 3.2]	(17757)
1/32	28	92 [6 ‖ 3.3]	(9526)
1/64	28	97 [7 ‖ 3.5]	(10376)
1/128	29	110 [7 ‖ 3.8]	(9308)
1/256	31	132 [7 ‖ 4.3]	(9210)

$\overline{\gamma} = 10^8$, $a = 2$, $b = 1$, tol $= 10^{-10}$

Open problems:

- Estimate for $\|y^* - y_\gamma\|$ in terms of γ.
- Estimate for $\|y_\gamma - Y_\gamma\|$ in terms of h.
- Then balance both errors and use mesh adaptation.
- Model function for value functional $J_\gamma(y)$, leading to a path following strategy $\{\gamma_l\}_{l\geq 0}$ that exhibits a superlinear convergence rate in $\|y^* - y_{\gamma_l}\|$.