MATHEON A1

The Control Reduced Interior Point Method

Anton Schiela

Martin Weiser, Tobias Gänzler, Peter Deuflhard (ZIB)
Fredi Tröltzsch, Uwe Prüfert (TU Berlin)

Graz, 21.06. 2006

Overview

Control Reduced Interior Point Methods...

- Convergence results in function space
- Discretization error estimates
- Algorithmic Concepts

Based on...

- Weiser/Gänzler/Sch. *[ZR 04-38]*, Sch./Weiser *[COAP 2006]*
- Schiela *[PhD-Thesis '06]*, Schiela *[ZR 06-16]*

Related Work...

- M. Weiser et al. (Function Space Oriented IP-Methods)
- M. Hinze (Elimination of the Control)

- S. Ulbrich (Primal-Dual IP-Methods for PDE Opt.)
- M. Ulbrich (Semi-Smooth Newton Methods)
- K. Kunisch, M. Hintermüller (PDAS Strategy, PD Path-Following)

Model Problem in Optimal Control

$$\min \frac{1}{2}\left\|y-y_d\right\|_{L_2(\Omega)}^2+\frac{\alpha}{2}\left\|u\right\|_{L_2(\Omega)}^2$$ objective functional

s.t. $Ly=u$ PDE constraint

$-1 \leqslant u \leqslant 1$ control bounds

- control u, state y
- regular domain in R^d

$u \in L_2 \Rightarrow y \in W \subset L_q \quad q > 2$

L: differential operator

Model Problem in Optimal Control

$$\min \frac{1}{2}\|y - y_d\|^2_{L_2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L_2(\Omega)}$$ objective functional

s.t. $L\,y = u$ PDE constraint

$y \geqslant 0$ control bounds

- control u, state y
- regular domain in R^d

$u \in L_2 \Rightarrow y \in W \subset C(\overline{\Omega})$

L: differential operator

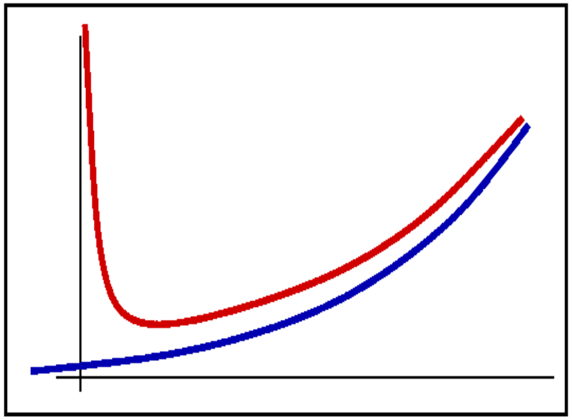
Primal Interior Point Approximation

MATHEON A1

Augment functional by logarithmic **barrier term**

$$\min \frac{1}{2}\|y - y_d\|^2_{L_2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L_2(\Omega)} - \mu \int_\Omega \left(\ln(1-u) + \ln(u+1)\right) dx$$

s.t. $\quad L\,y = u$



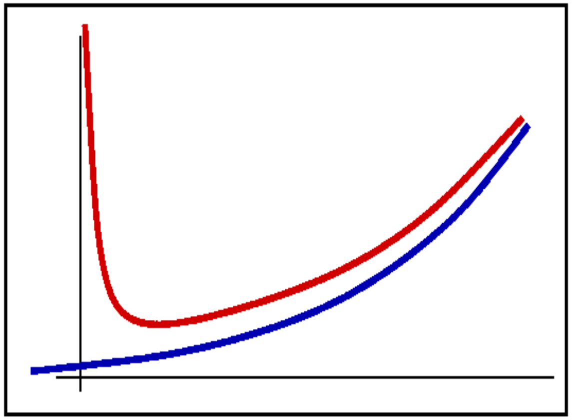
Primal Interior Point Approximation

MATHEON A1

Augment functional by logarithmic **barrier term**

$$\min \frac{1}{2}\left\|y-y_d\right\|^2_{L_2(\Omega)}+\frac{\alpha}{2}\left\|u\right\|^2_{L_2(\Omega)} -\mu\int_\Omega \left(\ln(1-u)+\ln(u+1)\right)\,dx$$

s.t. $\quad Ly=u$



First Order Necessary Conditions

$$y-y_d+L^*\lambda=0$$

$$Ly-u=0$$

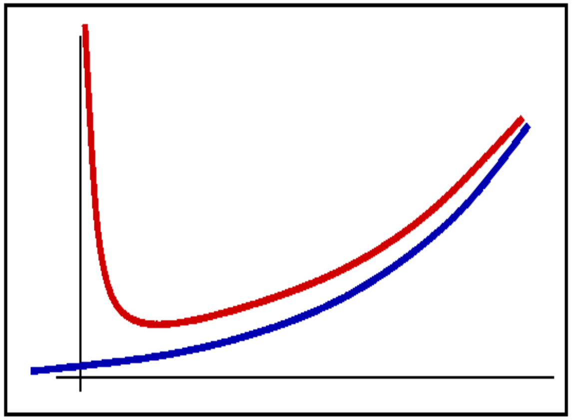
$$\alpha u-\lambda-\frac{\mu}{u+1}+\frac{\mu}{1-u}=0$$

Primal Interior Point Approximation

Augment functional by logarithmic **barrier term**

$$\min \frac{1}{2}\left\|y-y_d\right\|^2_{L_2(\Omega)}+\frac{\alpha}{2}\left\|u\right\|^2_{L_2(\Omega)} -\mu\int_\Omega \left(\ln(1-u)+\ln(u+1)\right)\,dx$$

s.t. $\quad Ly=u$



First Order Necessary Conditions

$$y-y_d+L^*\lambda=0$$
$$Ly-u=0$$
$$\alpha u-\lambda-\frac{\mu}{u+1}+\frac{\mu}{1-u}=0$$

Elimination of the control

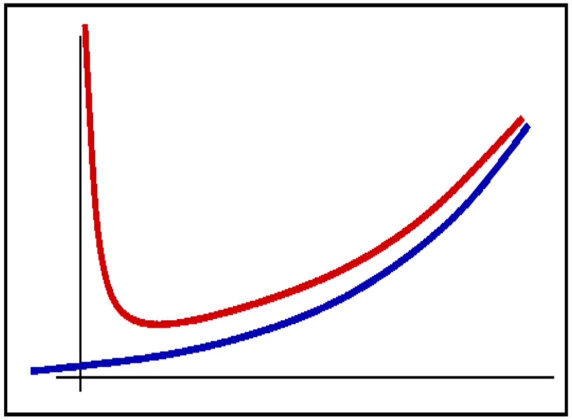
$$y-y_d+L^*\lambda=0$$
$$Ly-u(\lambda;\mu)=0$$

Primal Interior Point Approximation

Augment functional by logarithmic **barrier term**

$$\min \frac{1}{2}\|y-y_d\|^2_{L_2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L_2(\Omega)} - \mu \int_\Omega \left(\ln(1-u)+\ln(u+1)\right)\,dx$$

s.t. $\quad L\,y = u$



First Order Necessary Conditions

$$y - y_d + L^*\lambda = 0$$
$$Ly - u = 0$$
$$\alpha u - \lambda - \frac{\mu}{u+1} + \frac{\mu}{1-u} = 0$$

Elimination of the control

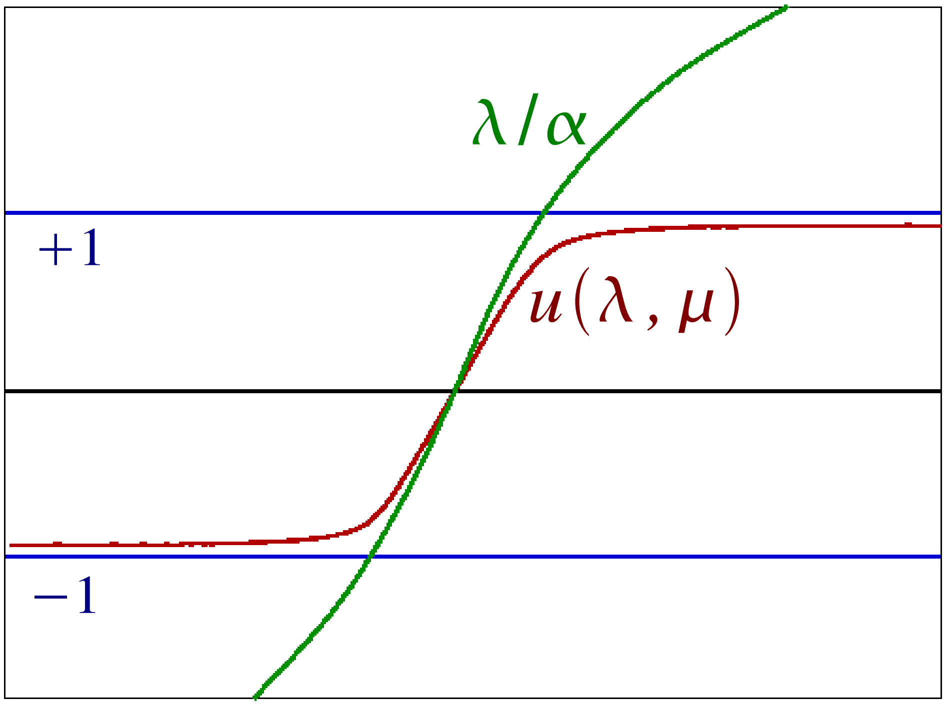
$$y - y_d + L^*\lambda = 0$$
$$L\,y - u(\lambda;\mu) = 0$$

Homotopy: $\mu \to 0$

Elimination of the Control

MATHEON A1

- Use optimality condition to **eliminate** the control **pointwise**
- Idea of *[M. Hinze]* carried over to interior point methods
- Discretization of the **smooth** variables y and λ only
- Control is always feasible



Consequences:

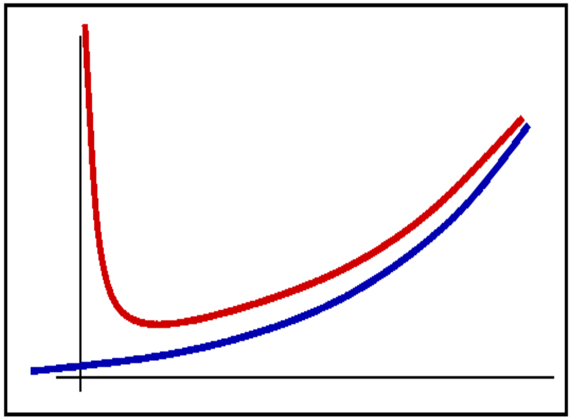
- **fast convergence** of interior point path-following method
- **optimal** discretization error estimates

Primal Interior Point Approximation

Augment functional by logarithmic **barrier term**

$$min \frac{1}{2}\left\|y-y_d\right\|^2_{L_2(\Omega)}+\frac{\alpha}{2}\left\|u\right\|^2_{L_2(\Omega)}$$
$$-\mu\int_\Omega \ln(y)\,dx$$

s.t. $\quad L\,y=u$



First Order Necessary Conditions

$$y-y_d-\frac{\mu}{y}+L^*\lambda=0$$
$$Ly-u=0$$
$$\alpha u-\lambda=0$$

Elimination of the control

$$y-y_d-\frac{\mu}{y}+L^*\lambda=0$$
$$L\,y-u(\lambda;\mu)=0$$

Homotopy: $\mu\rightarrow 0$

Path of Minimizers

Problem in convex analysis:

Original problem: $x_* = argmin(j + \chi_E + \chi_I)(x) = f(x)$

$x \in X$

Barrier problem: $x(\mu) = argmin(j + \chi_E + b_{I,\mu})(x) = f_{b,\mu}(x)$

Apply existence theorem on minimizers of convex problems:

Ingredients:

- Pointwise strictly feasible solution
- Barrier functional is lower semi-continuous

$\Rightarrow$ existence of minimizer $x(\mu)$ of barrier problems

$x(\mu)$ strictly feasible almost everywhere

Path of Minimizers

Convergence:

- $\breve{x}$ pointwise strictly feasible
- x_* minimizer of original problem
- $v(\lambda)=(1-\lambda)x_*+\lambda\breve{x}$ strictly feasible
- compare $x(\mu)$ and $v(\mu)$
- convergence of function values
- convergence of minimizers

$$(f_{b,\mu})(x(\mu))\leq f(x_*)+c\mu(1-\ln\mu)$$

$$\|x(\mu)-x_*\|\leq c\sqrt{\mu(1-\ln\mu)}$$

First Order Necessary Conditions

Subdifferential calculus:

- x minimizes $f_{b,\mu}$

  $\Rightarrow \quad 0 \in \partial f_{b,\mu}(x) = \partial(j + \chi_E + b_{I,\mu})(x) \subset X^*$
- sum-rule (needs regularity (Slater) condition)

  $\Rightarrow \quad 0 \in \partial j(x) + \partial \chi_E(x) + \partial b_{I,\mu}(x)$

-> First order necessary conditions

- control constraints:

  $\partial b_{I,\mu}(x(\mu)) \in L_p(\Omega) \quad \Rightarrow \quad \partial b_{I,\mu} = b'_{I,\mu}$
- state constraints

  $\partial b_{I,\mu}(x(\mu)) \subset M(\overline{\Omega}) \quad \Rightarrow$ difficult to interpret

Application to model problems

Control Constraints:

$$y - y_d + L^* \lambda = 0$$
$$Ly - u = 0$$
$$\alpha u - \lambda - \frac{\mu}{u+1} + \frac{\mu}{1-u} = 0$$

is solvable, even if only

$$u \in L_2 \Rightarrow y \in W \subset L_q \qquad q \geqslant 2$$

State Constraints:

$$y - y_d - \frac{\mu}{y} + L^* \lambda = 0$$
$$Ly - u = 0$$
$$\alpha u - \lambda = 0$$

need not be solvable, even if

$$u \in L_2 \Rightarrow y \in W \subset C(\overline{\Omega})$$

reason: $\partial b_{I,\mu}(x(\mu)) \supset \frac{-\mu}{y}$

Discussion

State Constraints:

- *formal* optimality system may not have a solution
- problems appear only, if $y(\mu)$ is not strictly feasible
- first idea: *force strict feasibility:* $y(\mu) > \varepsilon(\mu) > 0$
- exploit smoothness of $y = L^{-1}u \in W$
- logarithmic barrier function is often to weak
- second idea: *strengthen barrier functional:*

$$\boxed{b_\mu(y) = -\int_\Omega \frac{\mu^q}{y^{q-1}}\,dx \quad for \quad q > 1}$$

„Rational Barrier Functional of order q“

Orientation

Convex analysis:

- constrained minimization problem
- replace characteristic function by barrier function
- existence and convergence of a path of minimizers
- solvability of optimality system for adequate barrier functions

Orientation

Convex analysis:

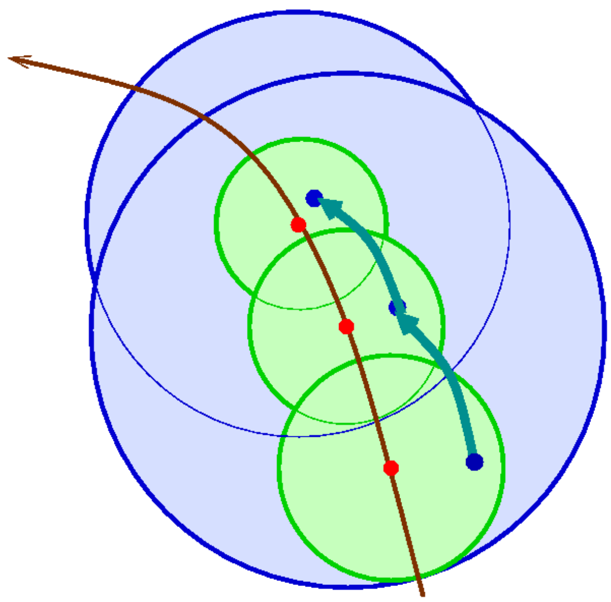
- constrained minimization problem
- replace characteristic function by barrier function
- existence and convergence of a path of minimizers
- solvability of optimality system for adequate barrier functions

Nonlinear functional analysis:

- differentiability of the central path
- improved error estimates $\|x(\mu)-x_*\|=O(\mu^\gamma) \quad \gamma>1/2$
- Newton methods for finding barrier solutions
- overall convergence theory

Affine Covariant Path–Following

Path–Following in normed function spaces



Affine covariant existence and convergence theory

[Deuflhard '04],[Sch. PhD '06]

Parameterized Families of Equations

Family of nonlinear operators: $F(x;\mu) : X \times M \to Y$

Family of nonlinear equations: $F(x;\mu) = 0$

Family of solutions: $x(\mu)$

$$y - y_d + L^* \lambda = 0$$
$$L y - u(\lambda;\mu) = 0$$

Main assumptions: $x(\mu_0)$ exists + “smoothness” of F

Questions:

- local existence of $x(\mu)$
- differentiability of $\mu \to x(\mu)$
- computation of $x(\mu)$

Parameterized Families of Equations

Family of nonlinear operators: $F(x;\mu):X\times M\to Y$

Family of nonlinear equations: $F(x;\mu)=0$

Family of solutions: $x(\mu)$

Main assumptions: $x(\mu_0)$ exists + “smoothness” of F

Questions:

- local existence of $x(\mu)$
- differentiability of $\mu\to x(\mu)$
- computation of $x(\mu)$

Implicit Function Theorem

Newton's Method

Parameterized Families of Equations

Family of nonlinear operators: $F(x;\mu): X \times M \to Y$

Family of nonlinear equations: $F(x;\mu) = 0$

Family of solutions: $x(\mu)$

Main assumptions: $x(\mu_0)$ exists + “smoothness” of F

Questions:

- local existence of $x(\mu)$
- differentiability of $\mu \to x(\mu)$
- computation of $x(\mu)$

Implicit Function Theorem

Newton's Method

Fixed Point Iterations in X

Parameterized Families of Equations

Family of nonlinear operators: $F(x;\mu):X\times M\rightarrow Y$

Family of nonlinear equations: $F(x;\mu)=0$

Family of solutions: $x(\mu)$

Main assumptions: $x(\mu_0)$ exists + “smoothness” of F

Questions:

- local existence of $x(\mu)$
- differentiability of $\mu\rightarrow x(\mu)$
- computation of $x(\mu)$

Implicit Function Theorem

Newton's Method

Fixed Point Iterations in X

Inverse Differentiability

Aim: characterize smoothness of F independently of the choice of Y

Classical assumptions: $F(\cdot\,;\mu) : X \to Y$ cont. Fréchet diff.

$F'(x;\mu)^{-1} : Y \to X$ exists, continuous

$\boxed{\text{result depends on choice of } Y}$

Inverse Differentiability

Aim: characterize smoothness of F independently of the choice of Y

Classical assumptions: $F(\cdot\,;\mu) : X \to Y$ cont. Fréchet diff.

$F'(x;\mu)^{-1} : Y \to X$ exists, continuous

result depends on choice of Y

Affine covariant version: $F'(x)^{-1}F(\cdot) : X \to X$ Fréchet differentiable

Inverse Fréchet Differentiability

MATHEON A1

Aim: characterize smoothness of F independently of the choice of Y

Classical assumptions: $F(\cdot\,;\mu): X \to Y$ cont. Fréchet diff.

$F'(x;\mu)^{-1}: Y \to X$ exists, continuous

result depends on choice of Y

Affine covariant version: $F'(x)^{-1}F(\cdot): X \to X$ Fréchet differentiable

Inverse *Fréchet* differentiability at x:

$$\|(x-\tilde{x})-F^{-}(x)(F(x)-F(\tilde{x}))\|_X = o(\|x-\tilde{x}\|_X)$$

Differentiability of a Homotopy Path

Affine Covariant Implicit Function Theorem:

Inverse *Fréchet* differentiability at x, "extending" into M

$$\|(x-\tilde{x})-F^{-}(x;\mu)(F(x;\tilde{\mu})-F(\tilde{x};\tilde{\mu}))\|_X=\omega(\Delta x,\Delta\mu)\|x-\tilde{x}\|_X$$

Differentiability of $F^{-}(x;\mu)F(x;\cdot)$ w.r.t. μ

Self-Mapping Lemma

Local existence criterion (contractivity, compactness, convexity,...)

Differentiability of the local mapping $\mu\rightarrow x(\mu)$

Inverse Newton Differentiability

Inverse *Newton* differentiability at x:

$$\boxed{\|(x-\tilde{x})-F^{-}(\tilde{x})(F(x)-F(\tilde{x}))\|_X = o(\|x-\tilde{x}\|_X)}$$

$F^{-}(\tilde{x})$ not at all unique: $F^{-}(\tilde{x})+o(\|x-\tilde{x}\|)$ is an inverse derivative, too!

Sufficient for superlinear convergence of Newton's method, if $F(x)=0$

Affine covariant version of Newton differentiability: *[Kunisch et al.]*

$$\begin{aligned}
&\|(x-\tilde{x})-F^{-}(\tilde{x})(F(x)-F(\tilde{x}))\|_X \\
&\quad = \|F'(\tilde{x})^{-1}(F'(\tilde{x})(x-\tilde{x})-(F(x)-F(\tilde{x})))\|_X \\
&\quad \leqslant \|F'(\tilde{x})^{-1}\|_{Y\to X}\|F'(\tilde{x})(x-\tilde{x})-(F(x)-F(\tilde{x}))\|_Y
\end{aligned}$$

Matheon A1

Quantification

Strengthened *Newton* differentiability at x :

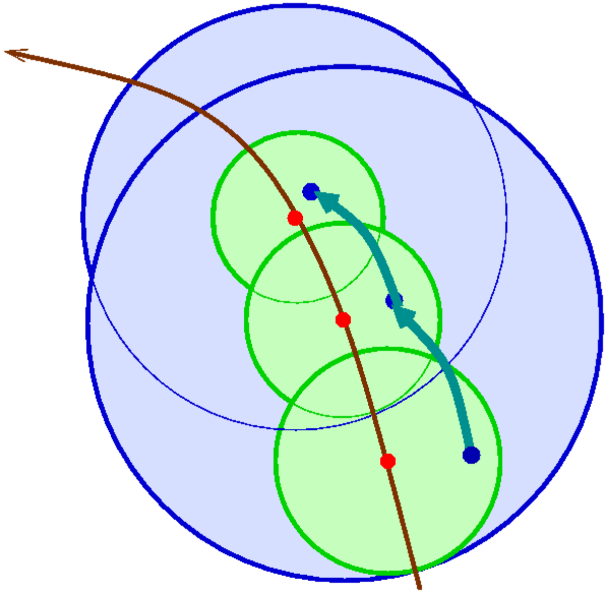
$$\|(x-\tilde{x})-F^{-}(\tilde{x})(F(x)-F(\tilde{x}))\|_X=(\omega\|x-\tilde{x}\|_X)^{\beta}\|x-\tilde{x}\|_X$$

Radius of contraction:

$$\rho_{\Theta}(\mu)=\frac{\Theta^{1/\beta}}{\omega(\mu)}$$

Slope of Homotopy Path:

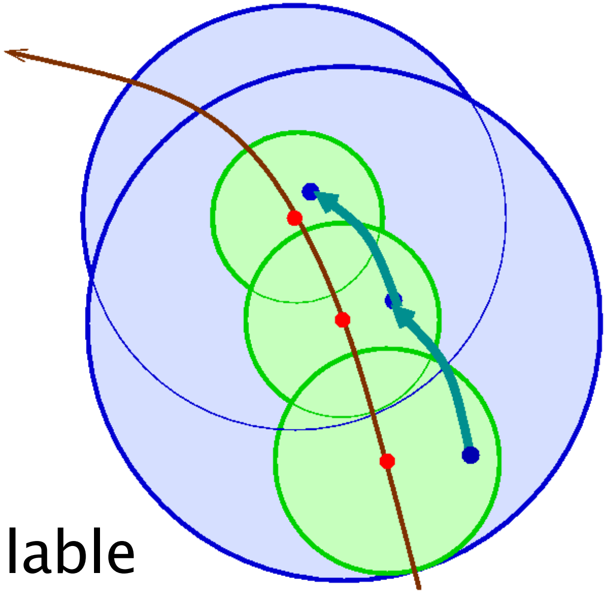
$$\eta(\mu)=\left\|\frac{d}{d\mu}x(\mu)\right\|_X \quad \Rightarrow \quad \|x(\mu)-x(\tilde{\mu})\|_X\leqslant\|\eta\|_{L_1([\mu,\tilde{\mu}])}$$



Key quantities $\eta(\mu)$ and $\omega(\mu)$ characterize progress of Newton path-following methods in terms of X

Role of the Key Quantities

Key quantities $\eta(\mu)$ and $\omega(\mu)$ characterize progress of Newton path-following methods in terms of X



Strategy of proof:

- *a-priori* estimates for $\eta(\mu)$ and $\omega(\mu)$
- worst-case bounds for step-size selection
- truncation error: $\|x(\mu)-x(0)\|_X \leqslant \|\eta\|_{L_1([\mu,0])}$

Adaptive path-following methods:

- *a-posteriori* estimates for $\eta(\mu)$ and $\omega(\mu)$ available
- adaptive step-size selection scheme
- termination criterion

Control Reduced IP Method

Nonlinear System

$$F(x;\mu)=\begin{pmatrix} y-y_d+L^*\lambda \\ Ly-u(\lambda;\mu)\end{pmatrix}=0$$

Jacobian

$$F'(x;\mu)=\begin{pmatrix} I & L^* \\ L & -u_\lambda(\lambda;\mu)\end{pmatrix}$$

$$F^-(x;\mu):=F'(x;\mu)^{-1}:L_{q'}\subset W^*\to W\subset L_q$$

Central Path

$$\eta(\mu)=\|F'(x;\mu)^{-1}F_\mu(\lambda;\mu)\|_{L_q}$$

$$F_\mu(x;\mu)=-u_\mu(\lambda;\mu)$$

$$\|u_\mu(\lambda;\mu)\|_{L_{q'}}\leqslant C\|u_\mu(\lambda;\mu)\|_{L_\infty}=O(\mu^{-1/2})$$

Newton Corrector

$$\|F'(x)^{-1}R(x)\|_{L_q}=\omega(\mu)^\beta\|\Delta\lambda\|_{L_q}^{1+\beta}$$

$$\begin{aligned} R(x)&:=F'(x)(x-\tilde{x})-(F(x)-F(\tilde{x})) \\ &=-u_\lambda(\lambda)(\lambda-\tilde{\lambda})+u(\lambda)-u(\tilde{\lambda})\end{aligned}$$

$$\|R(x)\|_{L_{q'}}\leqslant\|\psi\|_{L_s}\|\Delta\lambda\|_{L_q}^{1+\beta}$$

$$\|\psi\|_{L_s}\leqslant C\|\psi\|_{L_\infty}=O(\mu^{-1/2})$$

Linear Convergence

Central Path

$$\eta(\mu) = O(\mu^{-1/2})$$

Newton Corrector

$$\omega(\mu) = O(\mu^{-1/2})$$

$$\sigma \sim (\mu \eta \omega)^{-1}$$

linear convergence:

$$\mu_{k+1} = \sigma \mu_k$$

$$\|x_k - x_*\| = O(\sqrt{\mu_k})$$

Complementarity Condition

MATHEON A1

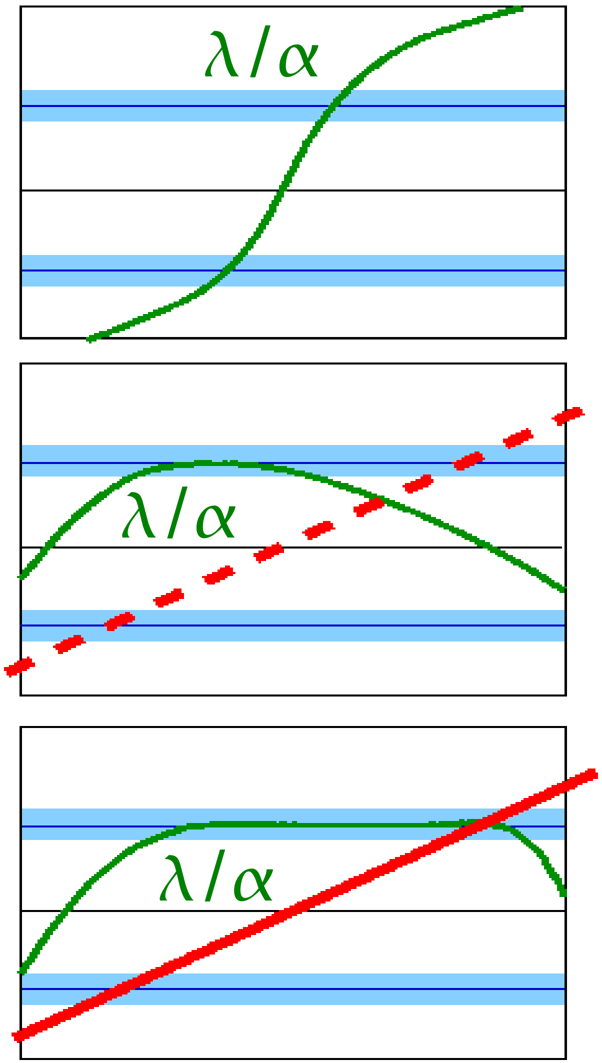
- Finite dimensional IP-Methods: „Strict complementarity implies superlinear convergence“
- Analogue in function space:

$$\left|\left\{x\in\Omega : \left|\lambda(x)\pm\alpha\right|\leqslant e\right\}\right|\leqslant\Gamma\, e$$

- Similar to *[M. & S. Ulbrich]*: „Strong strict complementarity“

$$\|u_\mu(\lambda;\mu)\|_{L_{q'}} \ll C\,\|u_\mu(\lambda;\mu)\|_{L_\infty}$$

$$\|\psi\|_{L_s} \ll C\,\|\psi\|_{L_\infty}$$



Superlinear Convergence

Central Path

$$\eta(\mu) = O(1 - \ln \mu)$$

$$F'(x)^{-1} : L_1 \subset W^* \to W \subset L_\infty$$

Newton Corrector

$\omega(\mu) \leqslant c$ uniformly in $\mu \to 0$

strong strict complementarity

$$\sigma \sim (\mu \eta \omega)^{-1}$$

convergence of *r-order 2*:

$$\lim_{k\to\infty} \frac{\ln \mu_{k+1}}{\ln \mu_k} = 2$$

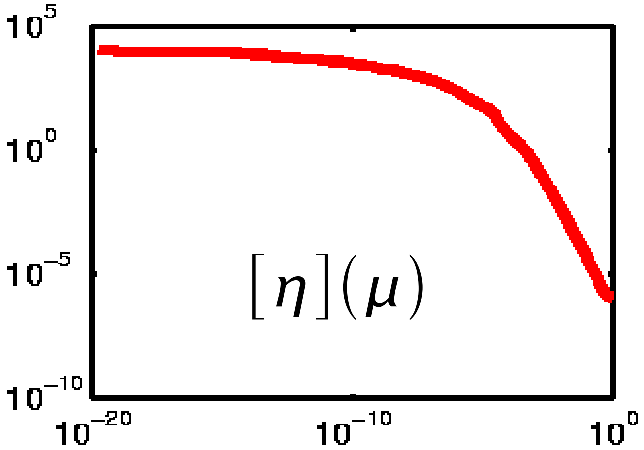
$$\|x_k - x_*\| = O(\mu_k \ln(\mu_k))$$

[Sch., Weiser COAP '06],
[Sch. PhD '06]

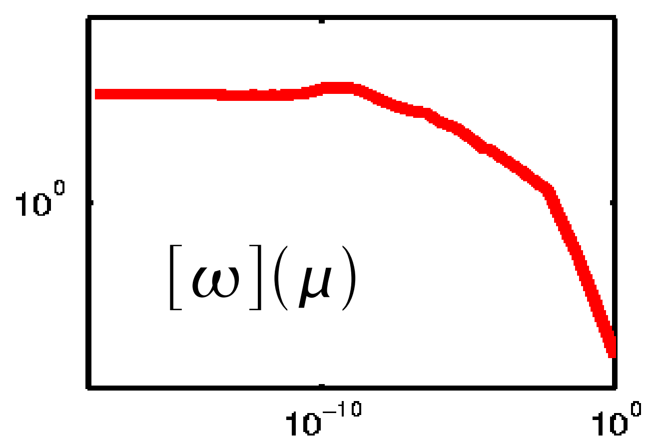
Adaptive Path-Following

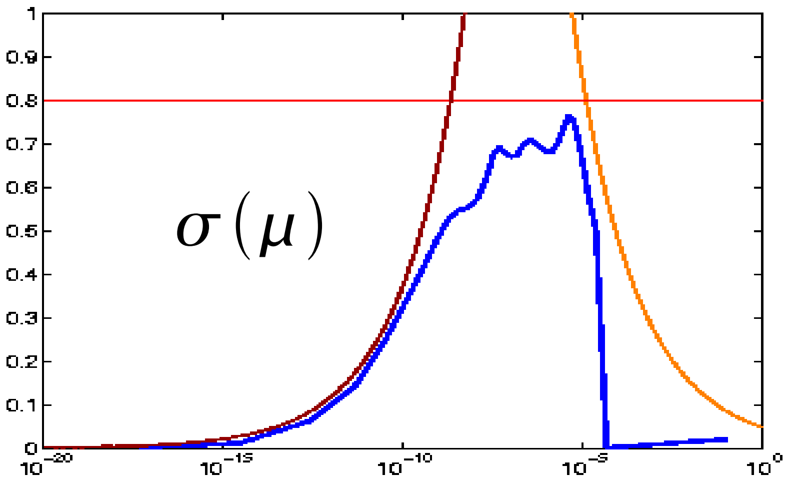
Matheon A1

Central Path



Newton Corrector





[Deuflhard '04]

[Sch. PhD '06]

MATHEON A1

Affine Covariant Estimates

Central Path

$$[\eta](\mu)=\frac{\|x(\mu)-x(\tilde{\mu})\|}{\mu-\tilde{\mu}}$$

Finite differences

Newton Corrector

$$[\omega](\mu)=\frac{\|\overline{\Delta x}\|}{\|\Delta x\|^2}$$

$\overline{\Delta x}$: simplified Newton step

$$[\omega](\mu)=\frac{\|\Delta x-F^{-}(x)(F(x)-F(x_{+}))\|}{\|\Delta x\|^2}$$

·Take into account inexactness
·Scaled norms favourable
·All quantities are mesh independent
·Accuracy matching

[Deuflhard '04]
[Sch. PhD '06]

State Constraints

Nonlinear System

$$F(x;\mu)=\begin{pmatrix} y-y_d-\dfrac{\mu^q}{y^q}+L^*\lambda \\ Ly-u(\lambda;\mu)\end{pmatrix}=0$$

Jacobian

$$F'(x;\mu)=\begin{pmatrix} I+\dfrac{\mu^q}{y^{q+1}} & L^* \\ L & -\alpha^{-1}\end{pmatrix}$$

$$F^-(x;\mu):=F'(x;\mu)^{-1}:L_1\subset W^*\to W\subset L_\infty$$

Central Path

$$\eta(\mu)=\|F'(x;\mu)^{-1}F_\mu(\lambda;\mu)\|$$

$$F_\mu(x;\mu)=\frac{-\mu^{q-1}}{y^q}$$

- Appropriate choice of scaled norms

$$\eta(\mu)=O(\mu^{-1/2})$$

Newton Corrector

- Similar analysis as in the control constrained case
- Qualitative argument: nonlinear system is smooth due to strict feasibility
- Quantitative argument: $\omega(\mu)$ is bounded from above in $[\underline{\mu};\mu_0]$ for each $\underline{\mu}>0$

MATHEON A1

Convergence for State Constraints

Central Path

$\eta(\mu) = O(\mu^{-1/2})$

Newton Corrector

$\omega(\mu)$ bounded in each $[\underline{\mu}; \mu_0]$

$\sigma \sim (\mu \eta \omega)^{-1}$

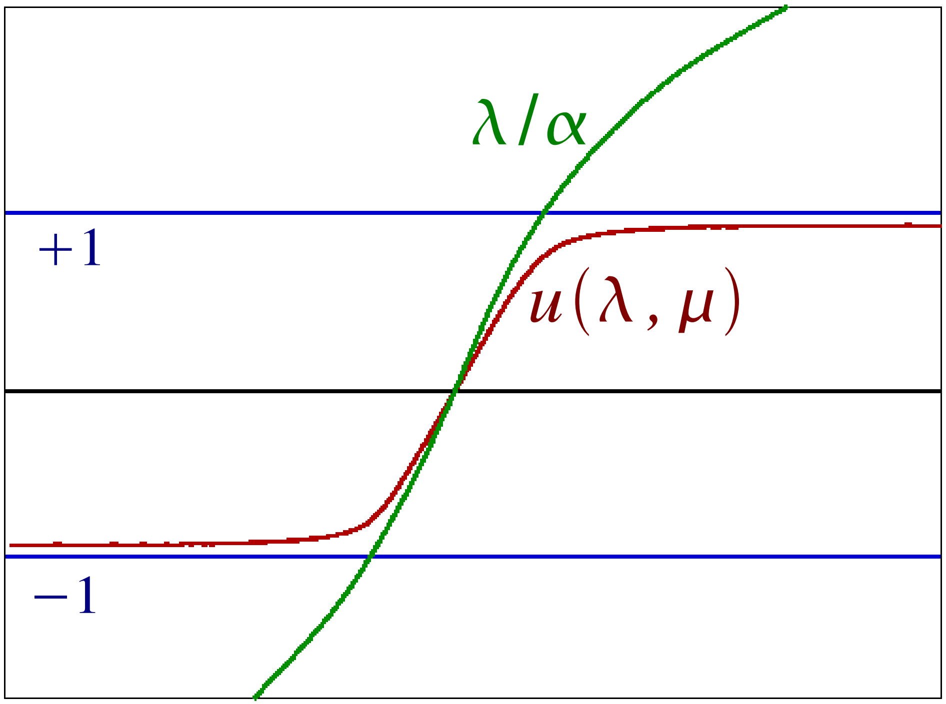
each $\mu > 0$ reached after finitely many steps

$$\|x_k - x_*\| = O(\sqrt{\mu_k})$$

Orientation

MATHEON A1

- Use optimality condition to **eliminate** the control **pointwise**
- Idea of *[M. Hinze]* carried over to interior point methods
- Discretization of the **smooth** variables y and λ only



Consequences:

- **fast convergence** of interior point path-following method
- **optimal** discretization error estimates

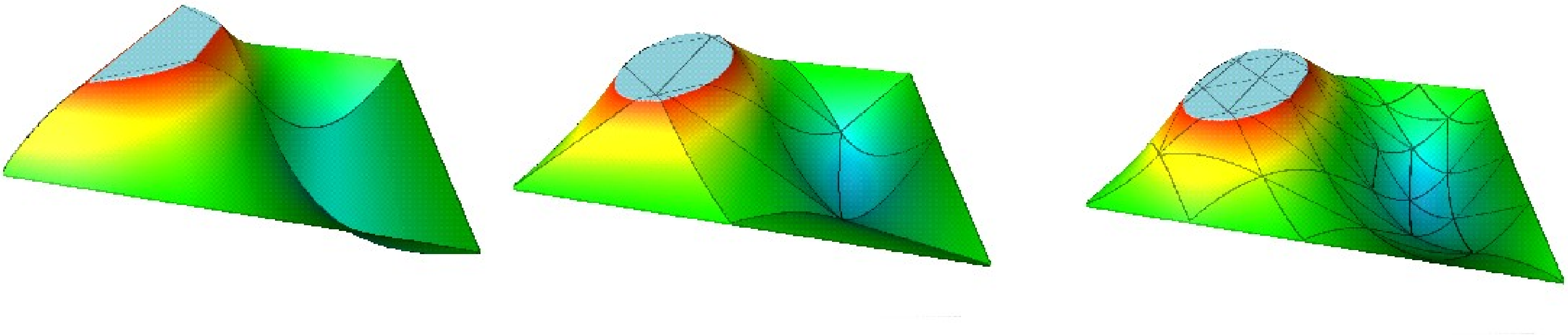
Discretization Error Estimates

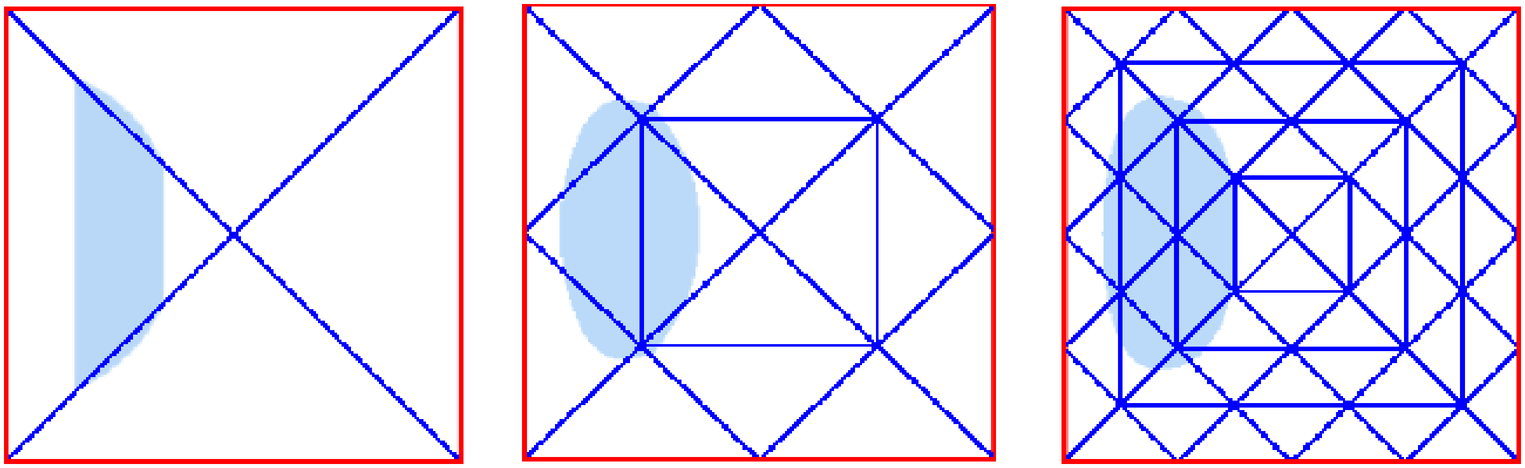
Finite element discretization of y and λ

Exact assembly of non-grid function u_h *[M. Hinze]* :

$$\|y_h - y(\mu)\|_{L_2} + \|\lambda_h - \lambda(\mu)\|_{L_2} + \|u_h - u(\mu)\|_{L_2} = O(h^{p+1})$$

Illustration





MATHEON A1

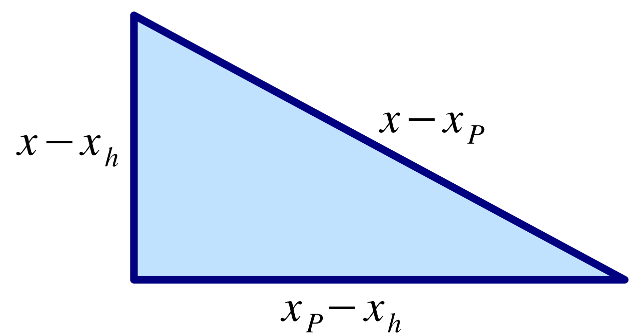
Sketch of Proof

$$\langle y-y_h, v_h\rangle + \langle \nabla(\lambda-\lambda_h), \nabla v_h\rangle = 0$$
$$\langle \nabla(y-y_h), \nabla w_h\rangle - \langle E^2(\lambda-\lambda_h), w_h\rangle = 0$$

Galerkin approximation of optimality system

$$\langle \nabla(\lambda-\lambda_P), \nabla v_h\rangle = 0$$
$$\langle \nabla(y-y_P), \nabla w_h\rangle = 0$$

Galerkin projections w.r.t. PDE



$$\langle y-y_h, y_P-y_h\rangle + \langle E^2(\lambda-\lambda_h), \lambda_P-\lambda_h\rangle \overset{?}{=} 0$$

Pythagoras' Theorem:

$$\|y-y_h\|^2 + \|\lambda-\lambda_h\|_E^2 + \|y_P-y_h\|^2 + \|\lambda_P-\lambda_h\|_E^2 = \|y-y_P\|^2 + \|\lambda-\lambda_P\|_E^2$$

Sketch of Proof

$$\langle y-y_h, v_h\rangle + \langle \nabla(\lambda-\lambda_h), \nabla v_h\rangle = 0$$
$$\langle \nabla(y-y_h), \nabla w_h\rangle - \langle E^2(\lambda-\lambda_h), w_h\rangle = 0$$

Galerkin approximation of optimality system

$$\langle \nabla(\lambda-\lambda_P), \nabla v_h\rangle = 0$$
$$\langle \nabla(y-y_P), \nabla w_h\rangle = 0$$

Galerkin projections w.r.t. PDE

$$\langle y-y_h, y_P-y_h\rangle + \langle E^2(\lambda-\lambda_h), \lambda_P-\lambda_h\rangle =$$
$$\langle y-y_h, y_P-y_h\rangle + \langle E^2(\lambda-\lambda_h), \lambda_P-\lambda_h\rangle$$
$$-\langle \nabla(y_P-y_h), \nabla(\lambda_P-\lambda_h)\rangle + \langle \nabla(\lambda_P-\lambda_h), \nabla(y_P-y_h)\rangle$$

Sketch of Proof

$$\langle y-y_h, v_h\rangle+\langle\nabla(\lambda-\lambda_h), \nabla v_h\rangle=0$$
$$\langle\nabla(y-y_h), \nabla w_h\rangle-\langle E^2(\lambda-\lambda_h), w_h\rangle=0$$

Galerkin approximation of optimality system

$$\langle\nabla(\lambda-\lambda_P), \nabla v_h\rangle=0$$
$$\langle\nabla(y-y_P), \nabla w_h\rangle=0$$

Galerkin projections w.r.t. PDE

$$\langle y-y_h, y_P-y_h\rangle+\langle E^2(\lambda-\lambda_h), \lambda_P-\lambda_h\rangle=$$
$$\langle y-y_h, y_P-y_h\rangle+\langle E^2(\lambda-\lambda_h), \lambda_P-\lambda_h\rangle$$
$$-\langle\nabla(y_P-y_h), \nabla(\lambda_P-\lambda_h)\rangle+\langle\nabla(\lambda_P-\lambda_h), \nabla(y_P-y_h)\rangle$$

MATHEON A1

Sketch of Proof

$$\langle y - y_h, v_h \rangle + \langle \nabla(\lambda - \lambda_h), \nabla v_h \rangle = 0$$
$$\langle \nabla(y - y_h), \nabla w_h \rangle - \langle E^2(\lambda - \lambda_h), w_h \rangle = 0$$

Galerkin approximation of optimality system

$$\langle \nabla(\lambda - \lambda_P), \nabla v_h \rangle = 0$$
$$\langle \nabla(y - y_P), \nabla w_h \rangle = 0$$

Galerkin projections w.r.t. PDE

$$\langle y - y_h, y_P - y_h \rangle + \langle E^2(\lambda - \lambda_h), \lambda_P - \lambda_h \rangle =$$
$$\langle y - y_h, y_P - y_h \rangle + \langle E^2(\lambda - \lambda_h), \lambda_P - \lambda_h \rangle$$
$$- \langle \nabla(y - y_h), \nabla(\lambda_P - \lambda_h) \rangle + \langle \nabla(\lambda - \lambda_h), \nabla(y_P - y_h) \rangle$$

Sketch of Proof

$$\langle y - y_h, v_h \rangle + \langle \nabla(\lambda - \lambda_h), \nabla v_h \rangle = 0$$
$$\langle \nabla(y - y_h), \nabla w_h \rangle - \langle E^2(\lambda - \lambda_h), w_h \rangle = 0$$

Galerkin approximation of optimality system

$$\langle \nabla(\lambda - \lambda_P), \nabla v_h \rangle = 0$$
$$\langle \nabla(y - y_P), \nabla w_h \rangle = 0$$

Galerkin projections w.r.t. PDE

$$\langle y - y_h, y_P - y_h \rangle + \langle E^2(\lambda - \lambda_h), \lambda_P - \lambda_h \rangle =$$
$$\langle y - y_h, y_P - y_h \rangle + \langle E^2(\lambda - \lambda_h), \lambda_P - \lambda_h \rangle$$
$$- \langle \nabla(y - y_h), \nabla \lambda_P - \lambda_h \rangle + \langle \nabla(\lambda - \lambda_h), \nabla y_P - y_h \rangle = 0$$

MATHEON A1

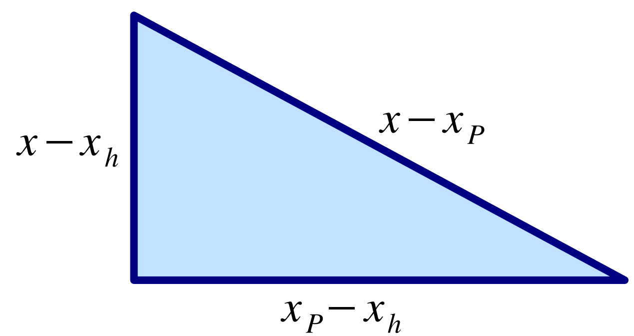
Sketch of Proof

$$\langle y-y_h, v_h\rangle + \langle \nabla(\lambda-\lambda_h), \nabla v_h\rangle = 0$$
$$\langle \nabla(y-y_h), \nabla w_h\rangle - \langle E^2(\lambda-\lambda_h), w_h\rangle = 0$$

Galerkin approximation of optimality system

$$\langle \nabla(\lambda-\lambda_P), \nabla v_h\rangle = 0$$
$$\langle \nabla(y-y_P), \nabla w_h\rangle = 0$$

Galerkin projections w.r.t. PDE



$$\langle y-y_h, y_P-y_h\rangle + \langle E^2(\lambda-\lambda_h), \lambda_P-\lambda_h\rangle = 0$$

Pythagoras' Theorem:

$$\|y-y_h\|^2 + \|\lambda-\lambda_h\|_E^2 + \|y_P-y_h\|^2 + \|\lambda_P-\lambda_h\|_E^2 = \|y-y_P\|^2 + \|\lambda-\lambda_P\|_E^2$$
$$= O(h^{p+1})^2$$

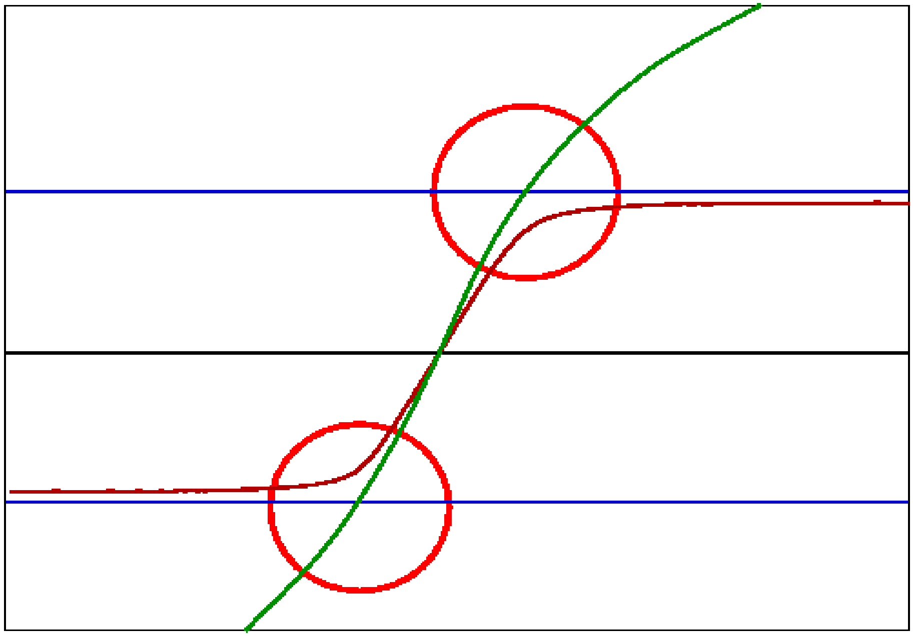
Inexact Assembly

Finite element discretization of y and λ $(p=1,2)$

Inexact assembly of non-grid function u_h :

$$\|y_h - y(\mu)\|_{L_2} + \|\lambda_h - \lambda(\mu)\|_{L_2} + \|u_h - u(\mu)\|_{L_2} = O(h^{p+1}) + e_{\Pi}$$

Use **adaptive** quadrature



Inexact Assembly

Finite element discretization of y and λ $(p=1,2)$

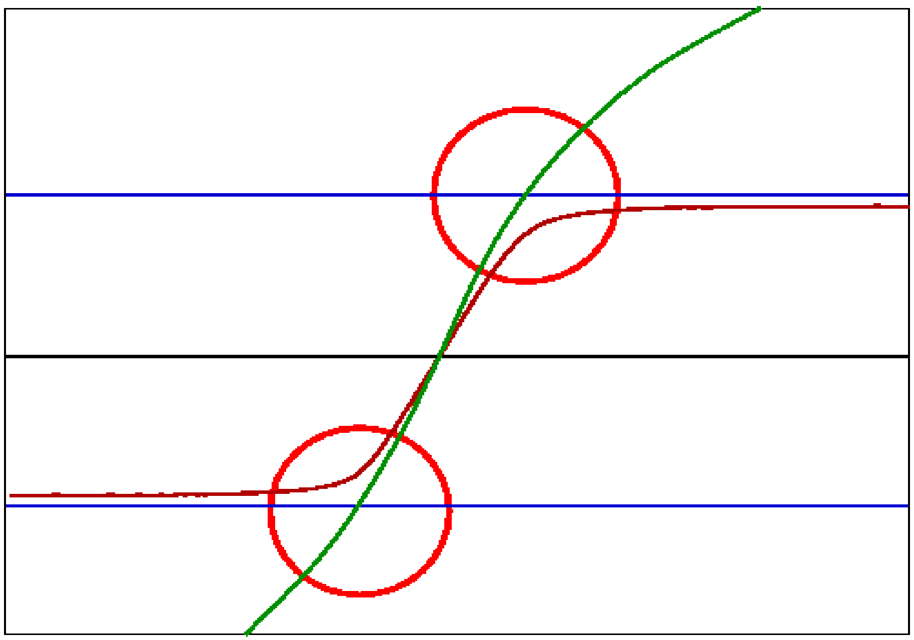
Inexact assembly of non-grid function u_h :

$$\|y_h - y(\mu)\|_{L_2} + \|\lambda_h - \lambda(\mu)\|_{L_2} + \|u_h - u(\mu)\|_{L_2} = O(h^{p+1}) + e_\Pi$$

Use **adaptive** quadrature

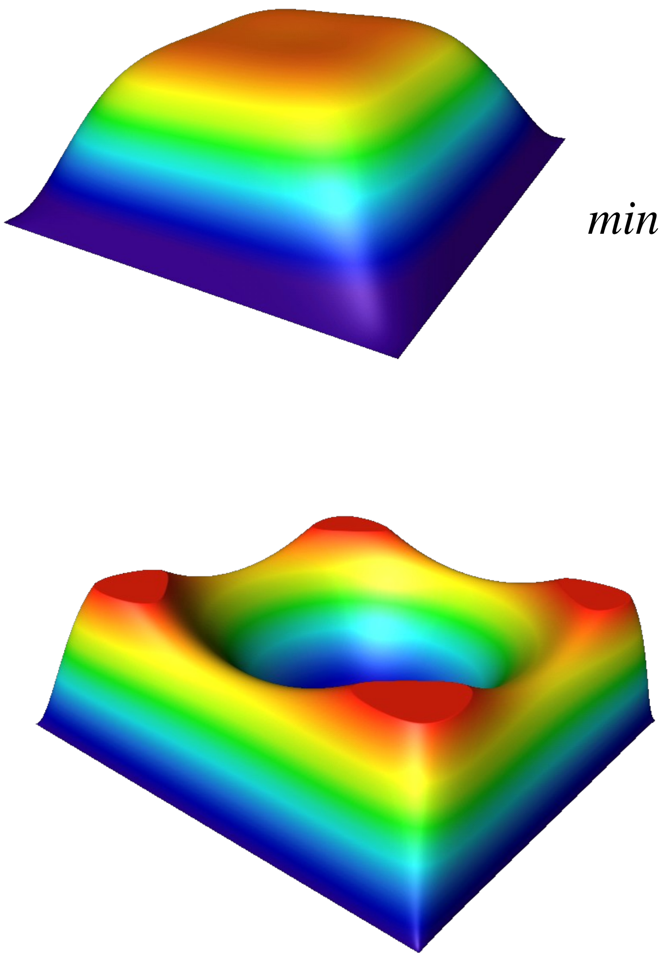
Computational amount:
1d,2d : negligible
3d : tolerable

[Weiser, Gänzler, Sch. ZR 04-38], [Sch. PhD '06]



Matheon A1

Discretization Errors



$$\min \tfrac{1}{2}\|y-y_d\|^2_{L_2(\Omega)} + \tfrac{\alpha}{2}\|u\|^2_{L_2(\Omega)}$$

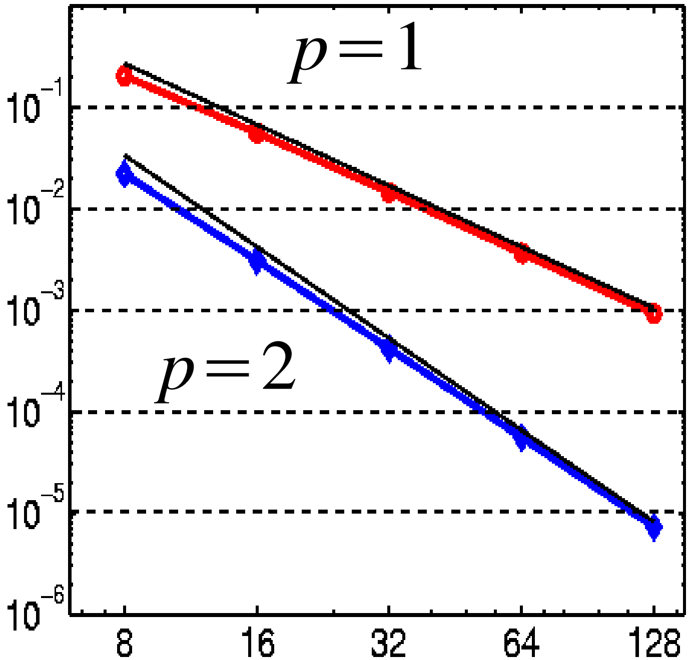
$$L\,y = u \quad on\,\Omega$$

$$y = 0 \quad on\,\partial\Omega$$

$$-40 \leqslant u \leqslant 40$$

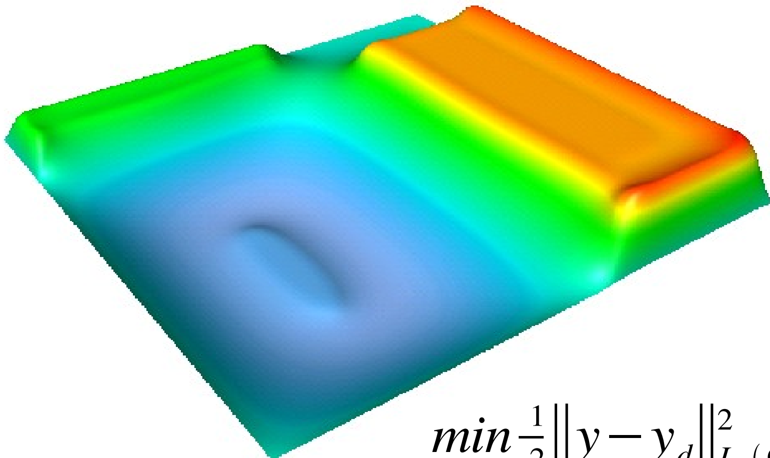
$$y_d = 0.1$$

$$\alpha = 10^{-6}$$



relative L_2–error in control

Discretization Errors

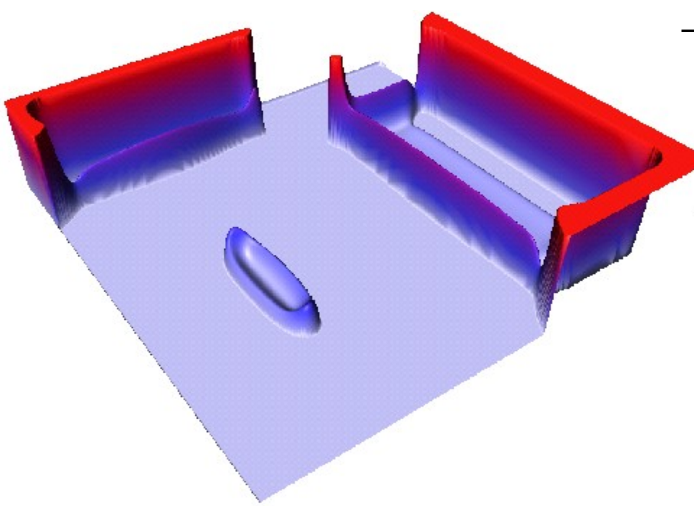


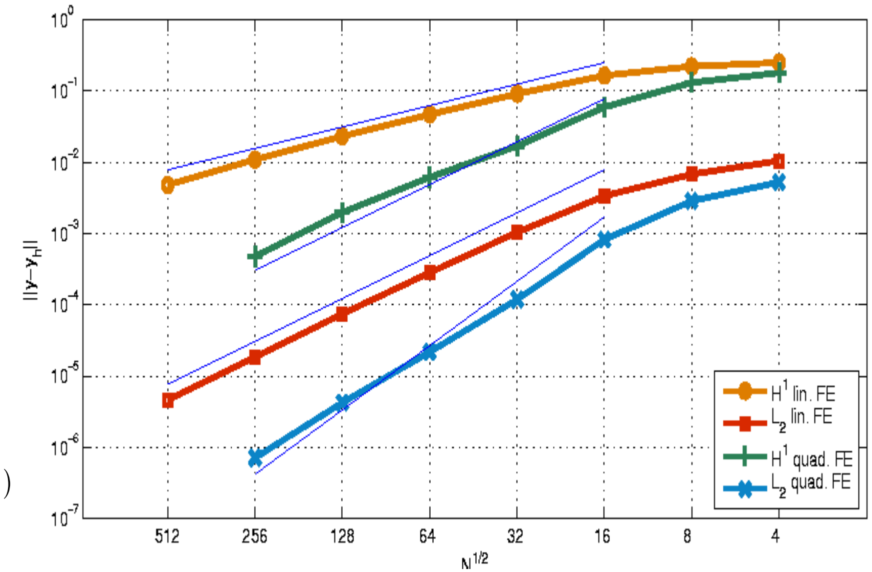
$$\min \frac{1}{2}\left\|y-y_d\right\|^2_{L_2(\Omega)}+\frac{\alpha}{2}\left\|u\right\|^2_{L_2(\Omega)}$$

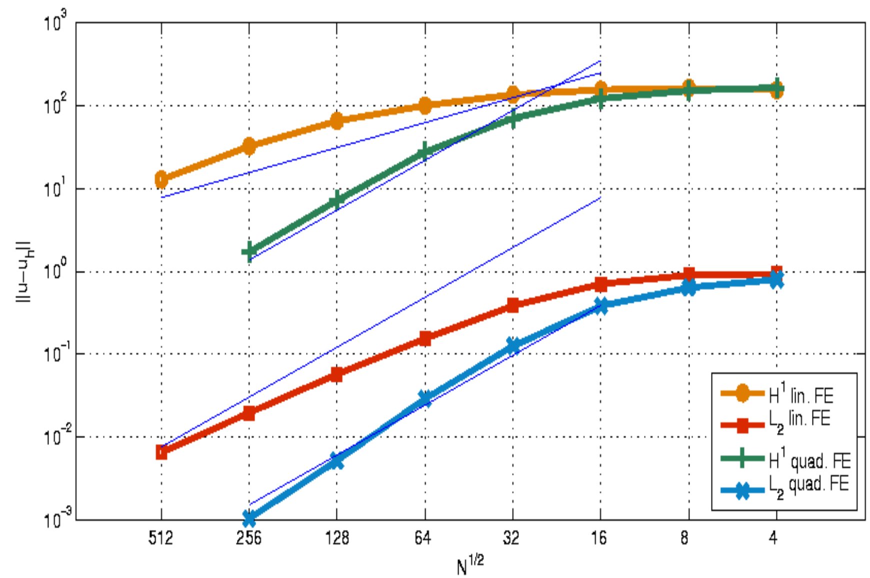
s.t. $\quad -0.5\,\Delta y+y=u$

$$-0.1\leqslant u\leqslant 1$$

$$\alpha=10^{-6}$$







Adaptive Multilevel Method

Adaptive stepsize selection
affine covariant strategy

Adaptive grid refinement
simple error indicator

stay inside region of Newton contraction

Matheon A1

Adaptive Multilevel Method

Adaptive stepsize selection
affine covariant strategy

Adaptive grid refinement
simple error indicator

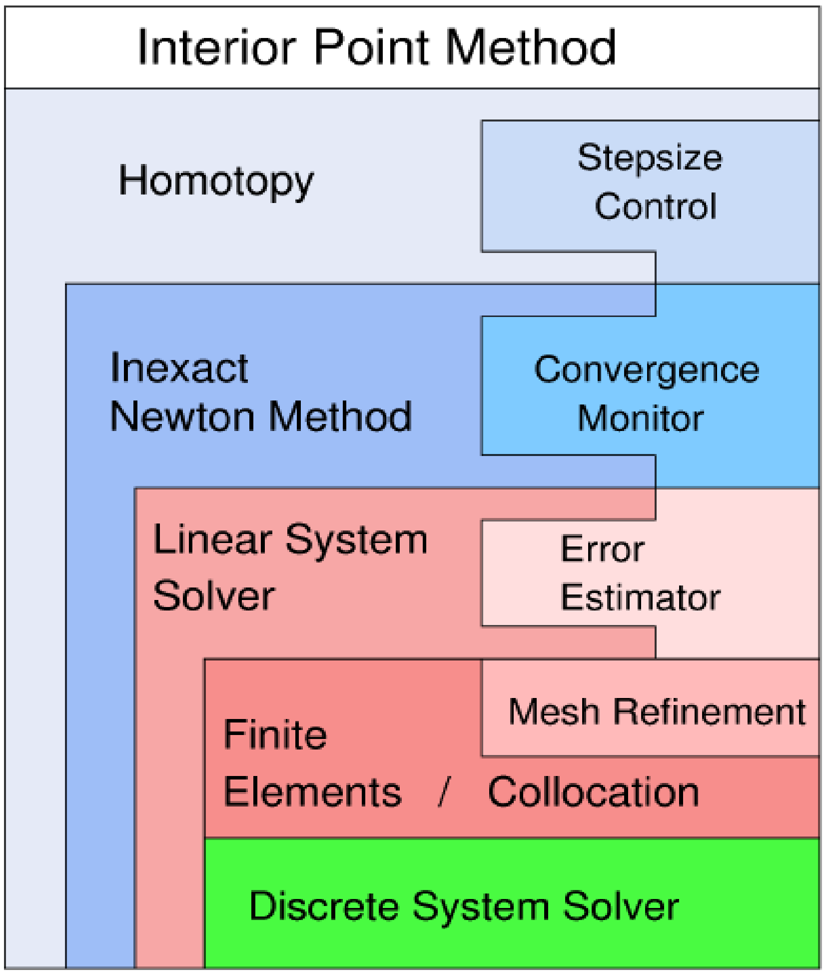
stay inside region of Newton contraction

Superlinear convergence & high order FEM

Final corrector step dominates computational cost

[Sch. PhD '06]

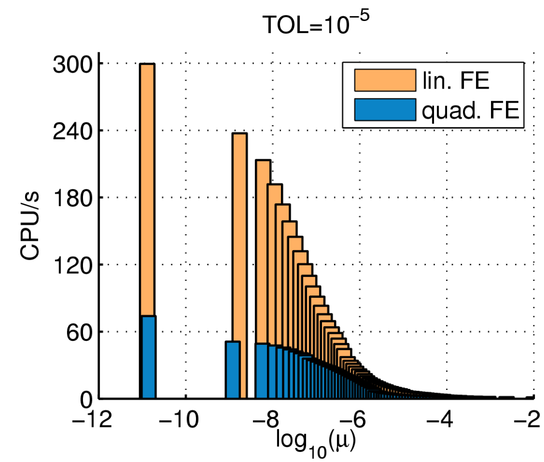
Adaptive Multilevel Method



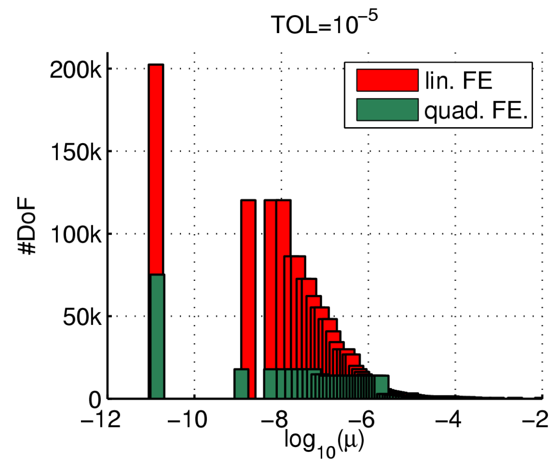
inequality constrained optimal control		sequence of nonlinear problems
nonlinear equation		sequence of linear problems
operator equation		sequence of finite dimensional equations

Adaptive Multilevel Method

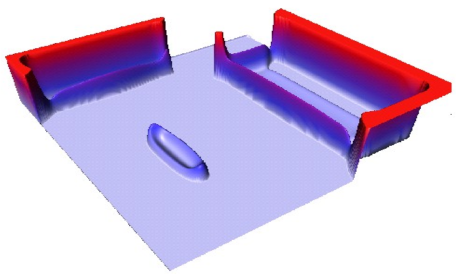
MATHEON A1

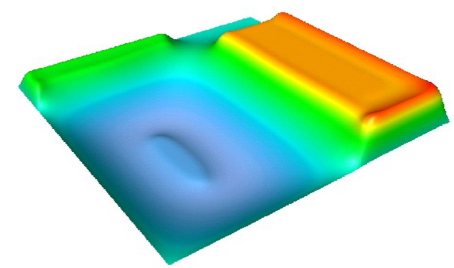


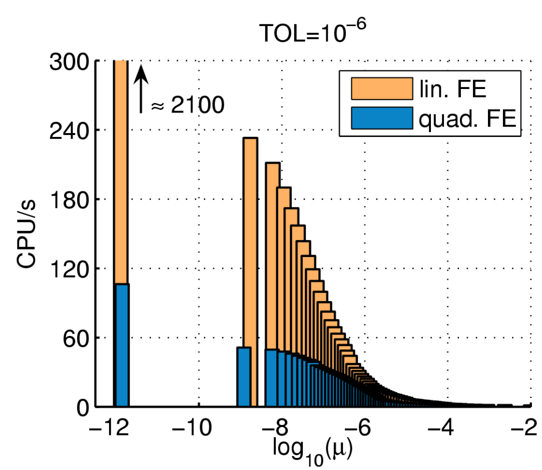
accumulated CPU-time

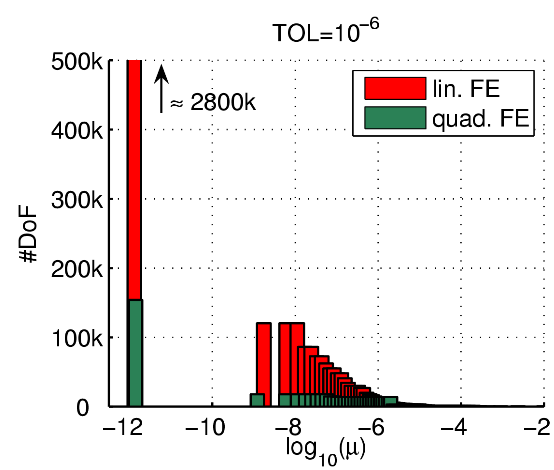


degrees of freedom









$$min \frac{1}{2}\left\|y-y_d\right\|^2_{L_2(\Omega)} + \frac{10^{-6}}{2}\left\|u\right\|^2_{L_2(\Omega)}$$

s.t. $\quad -0.5\,\Delta y + y = u$

$$-0.1 \leqslant u \leqslant 1$$

Summary

Control Reduced Methods for optimal control with PDEs

Control constraints

- Superlinear convergence
- Optimal discretization error estimates

State constraints

- Proof of convergence

Analysis & Algorithms

- Affine covariant strategy
- Fully adaptive algorithm

Thank you for your attention!