Multigrid Methods for Mixed Variational Problems and their Convergence Analysis

Joachim Schöberl[1] René Simon[2] Walter Zulehner[3]

[1]Center for Computational Engineering Science
RWTH Aachen, Germany

[2]SFB 013 Numerical and Symbolic Scientific Computing
JKU Linz, Austria

[3]Institute of Computational Mathematics
JKU Linz, Austria

Hamburg, March 27 - 29, 2008

Outline

Mixed variational problems

Find $x \in X$ and $p \in Q$ such that

$$\begin{aligned} a(x,w) + b(w,p) &= \langle f, w\rangle \quad \text{for all } w \in X, \\ b(x,q) - c(p,q) &= \langle g, q\rangle \quad \text{for all } q \in Q \end{aligned}$$

with

Hilbert spaces: X, Q,

bilinear forms: $a : X \times X \longrightarrow \mathbb{R}$, $b : X \times Q \longrightarrow \mathbb{R}$,
$c : Q \times Q \longrightarrow \mathbb{R}$,

linear functionals: $f : X \longrightarrow \mathbb{R}$, $g : Q \longrightarrow \mathbb{R}$.

Example: the Stokes problem

Find the velocity u and the pressure p such that

$$
\begin{aligned}
-\Delta u + \nabla p = f \quad & \text{in } \Omega, \\
\nabla \cdot u = 0 \quad & \text{in } \Omega, \\
u = 0 \quad & \text{on } \Gamma.
\end{aligned}
$$

Find $u \in H_0^1(\Omega)^d$ and $p \in L_0^2(\Omega)$ such that

$$
\begin{aligned}
(\nabla u, \nabla v)_{L^2(\Omega)^{d\times d}} - (p, \nabla \cdot v)_{L^2(\Omega)} &= (f, v)_{L^2(\Omega)^d} \quad && \text{for all } v \in H_0^1(\Omega)^d, \\
-(q, \nabla \cdot u)_{L^2(\Omega)} &= 0 && \text{for all } q \in L_0^2(\Omega).
\end{aligned}
$$

Example: the Stokes problem

Find the velocity u and the pressure p such that

$$\begin{aligned} -\Delta u + \nabla p &= f \quad \text{in } \Omega, \\ \nabla \cdot u &= 0 \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma. \end{aligned}$$

Find $u \in H_0^1(\Omega)^d$ and $p \in L_0^2(\Omega)$ such that

$$\begin{aligned} \underbrace{(\nabla u, \nabla v)_{L^2(\Omega)^{d\times d}}}_{a(u,v)} - (p, \nabla \cdot v)_{L^2(\Omega)} &= (f, v)_{L^2(\Omega)^d} \quad \text{for all } v \in H_0^1(\Omega)^d, \\ - (q, \nabla \cdot u)_{L^2(\Omega)} &= 0 \quad \text{for all } q \in L_0^2(\Omega). \end{aligned}$$

Example: the Stokes problem

Find the velocity u and the pressure p such that

$$\begin{aligned} -\Delta u + \nabla p = f \quad & \text{in } \Omega, \\ \nabla \cdot u = 0 \quad & \text{in } \Omega, \\ u = 0 \quad & \text{on } \Gamma. \end{aligned}$$

Find $u \in H_0^1(\Omega)^d$ and $p \in L_0^2(\Omega)$ such that

$$\begin{aligned} \underbrace{(\nabla u, \nabla v)_{L^2(\Omega)^{d\times d}}}_{a(u,v)} \underbrace{- (p, \nabla \cdot v)_{L^2(\Omega)}}_{b(v,p)} &= (f, v)_{L^2(\Omega)^d} \quad && \text{for all } v \in H_0^1(\Omega)^d, \\ \underbrace{- (q, \nabla \cdot u)_{L^2(\Omega)}}_{b(u,q)} &= 0 && \text{for all } q \in L_0^2(\Omega). \end{aligned}$$

Example: the Stokes problem

Minimize the cost functional

$$J(u) = \frac{1}{2} \|\nabla u\|^2_{L^2(\Omega)^{d\times d}} - (f, u)_{L^2(\Omega)^d}$$

for $u \in H^1_0(\Omega)^d$ subject to the constraint for u :

$$\nabla \cdot u = 0 \quad \text{in } \Omega.$$

Example: a model problem from optimal control

Minimize the cost functional

$$J(y,u) = \frac{1}{2}\,\|y - y_d\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\,\|u\|^2_{L^2(\Omega)}$$

subject to the state equation for y with distributed control u

$$\begin{aligned} -\Delta y + y &= u && \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= 0 && \text{on } \Gamma, \end{aligned}$$

where y_d is the desired state.

Example: a model problem from optimal control

Optimality system (KKT conditions):

- The adjoint state equation:

$$\begin{aligned} -\Delta p + p &= -(y - y_d) \quad \text{in } \Omega, \\ \frac{\partial p}{\partial n} &= 0 \quad \text{on } \Gamma. \end{aligned}$$

- The control equation:

$$\gamma\, u - p = 0 \quad \text{in } \Omega.$$

- The state equation:

$$\begin{aligned} -\Delta y + y &= u \quad \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= 0 \quad \text{on } \Gamma. \end{aligned}$$

Example: a model problem from optimal control

Find $y \in H^1(\Omega)$, $u \in L^2(\Omega)$ and $p \in H^1(\Omega)$ such that

$$\begin{aligned}
(y,z)_{L^2(\Omega)} \qquad\qquad\qquad + (z,p)_{H^1(\Omega)} &= (y_d,z)_{L^2(\Omega)} && \text{for all } z \in H^1(\Omega),\\
\gamma\,(u,v)_{L^2(\Omega)} - (v,p)_{L^2(\Omega)} &= 0 && \text{for all } v \in L^2(\Omega),\\
(y,q)_{H^1(\Omega)} - (u,q)_{L^2(\Omega)} \qquad\qquad &= 0 && \text{for all } q \in H^1(\Omega).
\end{aligned}$$

Find $(y,u) \in H^1(\Omega) \times L^2(\Omega)$ and $p \in H^1(\Omega)$ such that

$$\begin{aligned}
(y,z)_{L^2(\Omega)} + \gamma\,(u,v)_{L^2(\Omega)} + (z,p)_{H^1(\Omega)} - (v,p)_{L^2(\Omega)} &= (y_d,z)_{L^2(\Omega)},\\
(y,q)_{H^1(\Omega)} - (u,q)_{L^2(\Omega)} &= 0
\end{aligned}$$

for all $(z,v) \in H^1(\Omega) \times L^2(\Omega)$ and $q \in H^1(\Omega)$.

Example: a model problem from optimal control

Find $y \in H^1(\Omega)$, $u \in L^2(\Omega)$ and $p \in H^1(\Omega)$ such that

$$\begin{aligned}
(y,z)_{L^2(\Omega)} \qquad\qquad\qquad + (z,p)_{H^1(\Omega)} &= (y_d,z)_{L^2(\Omega)} && \text{for all } z \in H^1(\Omega),\\
\gamma\,(u,v)_{L^2(\Omega)} - (v,p)_{L^2(\Omega)} &= 0 && \text{for all } v \in L^2(\Omega),\\
(y,q)_{H^1(\Omega)} - (u,q)_{L^2(\Omega)} \qquad\qquad &= 0 && \text{for all } q \in H^1(\Omega).
\end{aligned}$$

Find $(y,u) \in H^1(\Omega) \times L^2(\Omega)$ and $p \in H^1(\Omega)$ such that

$$\begin{aligned}
\underbrace{(y,z)_{L^2(\Omega)} + \gamma\,(u,v)_{L^2(\Omega)}}_{a((y,u),(z,v))} + (z,p)_{H^1(\Omega)} - (v,p)_{L^2(\Omega)} &= (y_d,z)_{L^2(\Omega)},\\
(y,q)_{H^1(\Omega)} - (u,q)_{L^2(\Omega)} &= 0
\end{aligned}$$

for all $(z,v) \in H^1(\Omega) \times L^2(\Omega)$ and $q \in H^1(\Omega)$.

Example: a model problem from optimal control

Find $y \in H^1(\Omega)$, $u \in L^2(\Omega)$ and $p \in H^1(\Omega)$ such that

$$\begin{aligned}
(y,z)_{L^2(\Omega)} \qquad\qquad\qquad\qquad + (z,p)_{H^1(\Omega)} &= (y_d,z)_{L^2(\Omega)} \quad \text{for all } z \in H^1(\Omega),\\
\gamma\,(u,v)_{L^2(\Omega)} - (v,p)_{L^2(\Omega)} &= 0 \qquad\qquad\quad \text{for all } v \in L^2(\Omega),\\
(y,q)_{H^1(\Omega)} - (u,q)_{L^2(\Omega)} \qquad\qquad\qquad &= 0 \qquad\qquad\quad \text{for all } q \in H^1(\Omega).
\end{aligned}$$

Find $(y,u) \in H^1(\Omega) \times L^2(\Omega)$ and $p \in H^1(\Omega)$ such that

$$\begin{aligned}
\underbrace{(y,z)_{L^2(\Omega)} + \gamma\,(u,v)_{L^2(\Omega)}}_{a((y,u),(z,v))} + \underbrace{(z,p)_{H^1(\Omega)} - (v,p)_{L^2(\Omega)}}_{b((z,v),p)} &= (y_d,z)_{L^2(\Omega)},\\
\underbrace{(y,q)_{H^1(\Omega)} - (u,q)_{L^2(\Omega)}}_{b((y,u),q)} &= 0
\end{aligned}$$

for all $(z,v) \in H^1(\Omega) \times L^2(\Omega)$ and $q \in H^1(\Omega)$.

Example: a model problem from optimal control

Second approach: use the control equation to eliminate the control:

$$u = \frac{1}{\gamma} p.$$

Find $y \in H^1(\Omega)$ and $p \in H^1(\Omega)$ such that

$$(y, z)_{L^2(\Omega)} + (z, p)_{H^1(\Omega)} = (y_d, z)_{L^2(\Omega)} \quad \text{for all } z \in H^1(\Omega),$$

$$(y, q)_{H^1(\Omega)} - \frac{1}{\gamma}(p, q)_{L^2(\Omega)} = 0 \quad \text{for all } q \in H^1(\Omega).$$

Example: a model problem from optimal control

Second approach: use the control equation to eliminate the control:

$$u = \frac{1}{\gamma} p.$$

Find $y \in H^1(\Omega)$ and $p \in H^1(\Omega)$ such that

$$\begin{aligned}
\underbrace{(y, z)_{L^2(\Omega)}}_{a(y,z)} + (z, p)_{H^1(\Omega)} &= (y_d, z)_{L^2(\Omega)} && \text{for all } z \in H^1(\Omega), \\
(y, q)_{H^1(\Omega)} - \frac{1}{\gamma}(p, q)_{L^2(\Omega)} &= 0 && \text{for all } q \in H^1(\Omega).
\end{aligned}$$

Example: a model problem from optimal control

Second approach: use the control equation to eliminate the control:

$$u = \frac{1}{\gamma} p.$$

Find $y \in H^1(\Omega)$ and $p \in H^1(\Omega)$ such that

$$\underbrace{(y, z)_{L^2(\Omega)}}_{a(y,z)} + \underbrace{(z, p)_{H^1(\Omega)}}_{b(z,p)} = (y_d, z)_{L^2(\Omega)} \quad \text{for all } z \in H^1(\Omega),$$

$$\underbrace{(y, q)_{H^1(\Omega)}}_{b(y,q)} - \frac{1}{\gamma}(p, q)_{L^2(\Omega)} = 0 \quad \text{for all } q \in H^1(\Omega).$$

Example: a model problem from optimal control

Second approach: use the control equation to eliminate the control:

$$u = \frac{1}{\gamma} p.$$

Find $y \in H^1(\Omega)$ and $p \in H^1(\Omega)$ such that

$$\underbrace{(y, z)_{L^2(\Omega)}}_{a(y,z)} + \underbrace{(z, p)_{H^1(\Omega)}}_{b(z,p)} = (y_d, z)_{L^2(\Omega)} \quad \text{for all } z \in H^1(\Omega),$$

$$\underbrace{(y, q)_{H^1(\Omega)}}_{b(y,q)} - \underbrace{\frac{1}{\gamma}(p, q)_{L^2(\Omega)}}_{c(p,q)} = 0 \quad \text{for all } q \in H^1(\Omega).$$

Mixed variational problems

Find $x \in X$ and $p \in Q$ such that

$$\begin{aligned} a(x,w) + b(w,p) &= \langle f, w\rangle \quad \text{for all } w \in X,\\ b(x,q) - c(p,q) &= \langle g, q\rangle \quad \text{for all } q \in Q. \end{aligned}$$

$$\underbrace{a(x,w) + b(w,p) + b(x,q) - c(p,q)}_{\mathcal{B}((x,p),(w,q))} = \underbrace{\langle f, w\rangle + \langle g, q\rangle}_{\langle \mathcal{F}, (w,q)\rangle}$$

Find $(x,p) \in X \times Q$ such that

$$\mathcal{B}((x,p),(w,q)) = \langle \mathcal{F}, (w,q)\rangle \quad \text{for all } (w,q) \in X \times Q.$$

Mixed variational problems

Assumptions:

- $\mathcal{B}$ is bounded on $X \times Q$:

$$|\mathcal{B}((x,p),(w,q))| \leq C \,\|(x,p)\|_{X\times Q}\,\|(w,q)\|_{X\times Q}$$

  for all $(x,p),\ (w,q) \in X \times Q$.

- $\mathcal{B}$ is stable on $X \times Q$:

$$\sup_{0\neq(w,q)\in X\times Q} \frac{\mathcal{B}((x,p),(w,q))}{\|(w,q)\|_{X\times Q}} \geq c\,\|(x,p)\|_{X\times Q}$$

  for all $(x,p) \in X \times Q$.

Mixed variational problems

Brezzi's theorem (for the case $c \equiv 0$):

- a is bounded on X:

$$|a(x, w)| \leq \|a\| \, \|x\|_X \|w\|_X \quad \text{for all } x, w \in X.$$

- b is bounded on $X \times Q$:

$$|b(w, q)| \leq \|b\| \, \|w\|_X \|q\|_Q \quad \text{for all } w \in X, \; q \in Q.$$

- a is coercive on $\ker B = \{w \in X : b(w, q) = 0, \; q \in Q\}$:

$$a(w, w) \geq \alpha_0 \, \|w\|_X^2 \quad \text{for all } w \in \ker B.$$

- b satisfies the inf-sup condition on $X \times Q$:

$$\sup_{0 \neq w \in X} \frac{b(w, q)}{\|w\|_X} \geq k_0 \, \|q\|_Q \quad \text{for all } q \in Q.$$

Discretization

Find $(x, p) \in X \times Q$ such that

$$\mathcal{B}((x, p), (w, q)) = \langle \mathcal{F}, (w, q) \rangle \quad \text{for all } (w, q) \in X \times Q.$$

Galerkin's principle:

$$X_h \subset X, \quad Q_h \subset Q.$$

Find $(x_h, p_h) \in X_h \times Q_h$ such that

$$\mathcal{B}((x_h, p_h), (w_h, q_h)) = \langle \mathcal{F}, (w_h, q_h) \rangle \quad \text{for all } (w_h, q_h) \in X_h \times Q_h.$$

Find $x_h \in X_h$ and $p_h \in Q_h$ such that

$$\begin{aligned} a(x_h, w_h) + b(w_h, p_h) &= \langle f, w_h \rangle \quad \text{for all } w_h \in X_h, \\ b(x_h, q_h) - c(p_h, q_h) &= \langle g, q_h \rangle \quad \text{for all } q_h \in Q_h. \end{aligned}$$

Discretization

Matrix-vector representations:

$$\begin{aligned} a(x_h, w_h) &= (A_h \underline{x}_h, \underline{w}_h), & b(w_h, q_h) &= (B_h \underline{w}_h, \underline{q}_h), \\ c(p_h, q_h) &= (C_h \underline{p}_h, \underline{q}_h), & & \\ \langle f, w_h \rangle &= (\underline{f}_h, \underline{w}_h), & \langle g, q_h \rangle &= (\underline{g}_h, \underline{q}_h). \end{aligned}$$

Linear system:

$$\begin{pmatrix} A_h & B_h^T \\ B_h & -C_h \end{pmatrix} \begin{pmatrix} \underline{x}_h \\ \underline{p}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ \underline{g}_h \end{pmatrix}$$

Discretization

Brezzi's theorem (for the case $c \equiv 0$):

- a is bounded on X_h:

$$|a(x_h, w_h)| \leq \|a\| \, \|x_h\|_X \|w_h\|_X \quad \text{for all } x_h, w_h \in X_h.$$

- b is bounded on $X_h \times Q_h$:

$$|b(w_h, q_h)| \leq \|b\| \, \|w_h\|_X \|q_h\|_Q \quad \text{for all } w_h \in X_h,\ q_h \in Q_h.$$

- a is coercive on $\ker B_h = \{w_h \in X_h : b(w_h, q_h) = 0,\ q_h \in Q_h\}$:

$$a(w_h, w_h) \geq \alpha_0 \, \|w_h\|_X^2 \quad \text{for all } w_h \in \ker B_h.$$

- b satisfies a uniform inf-sup condition on $X_h \times Q_h$:

$$\sup_{0 \neq w_h \in X_h} \frac{b(w_h, q_h)}{\|w_h\|_X} \geq k_0 \, \|q_h\|_Q \quad \text{for all } q_h \in Q_h.$$

A class of iterative methods

Linear system

$$\mathcal{K}\begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \quad \text{with} \quad \mathcal{K} = \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix}$$

If A is non-singular, then

$$\begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} = \begin{pmatrix} A & 0 \\ B & -S \end{pmatrix} \begin{pmatrix} I & A^{-1}B^T \\ 0 & I \end{pmatrix} \text{ with } S = C + BA^{-1}B^T$$

Preconditioner

$$\hat{\mathcal{K}} = \begin{pmatrix} \hat{A} & 0 \\ B & -\hat{S} \end{pmatrix} \begin{pmatrix} I & \hat{A}^{-1}B^T \\ 0 & I \end{pmatrix} = \begin{pmatrix} \hat{A} & B^T \\ B & B\hat{A}^{-1}B^T - \hat{S} \end{pmatrix}$$

with $\hat{A}^T = \hat{A} > 0$ and $\hat{S}^T = \hat{S} > 0$.

A class of iterative methods

Preconditioned Richardson method:

$$\begin{pmatrix} x^{(k+1)} \\ p^{(k+1)} \end{pmatrix} = \begin{pmatrix} x^{(k)} \\ p^{(k)} \end{pmatrix} + \hat{\mathcal{K}}^{-1} \left[\begin{pmatrix} f \\ g \end{pmatrix} - \mathcal{K} \begin{pmatrix} x^{(k)} \\ p^{(k)} \end{pmatrix} \right]$$

requires to solve

$$\hat{\mathcal{K}} \begin{pmatrix} w \\ q \end{pmatrix} = \begin{pmatrix} r \\ s \end{pmatrix},$$

which reduces to three linear systems:

$$\begin{aligned} \hat{A}\,\hat{w} &= r, \\ \hat{S}\,q &= B\,\hat{w} - s, \\ \hat{A}\,w &= r - B^T s. \end{aligned}$$

Symmetric indefinite preconditioners

Case: $A^T = A \geq 0$ and $C = 0$:

Theorem

If

$$A \leq \hat{A} \quad \text{and} \quad B\hat{A}^{-1}B^T \geq \hat{S},$$

then

$$\sigma(\hat{\mathcal{K}}^{-1}\mathcal{K}) \subset (0, \infty).$$

If

$$A < \hat{A} \quad \text{and} \quad B\hat{A}^{-1}B^T > \hat{S},$$

then $\hat{\mathcal{K}}^{-1}\mathcal{K}$ *is selfadjoint and positive definite w.r.t. the scalar product*

$$\left(\begin{pmatrix} x \\ p \end{pmatrix}, \begin{pmatrix} w \\ q \end{pmatrix} \right)_{\mathcal{D}} = ([\hat{A} - A]x, w)_{\ell^2} + ([B\hat{A}^{-1}B^T - \hat{S}]p, q)_{\ell^2}.$$

$\Longrightarrow$ CG method applied to the preconditioned problem

Symmetric indefinite preconditioners

Assumptions (for the case $A^T = A \geq 0$ and $C = 0$):

- $$A \leq \hat{A}$$
- $$B \hat{A}^{-1} B^T \leq \beta \hat{S}$$
- $$(Aw, w)_{\ell^2} \geq \alpha (\hat{A} w, w)_{\ell^2} \quad \text{for all } w \in \ker B$$
- $$B \hat{A}^{-1} B^T \geq \hat{S}$$

with

$$0 \leq \alpha \leq 1 \quad \text{and} \quad \beta \geq 1$$

Symmetric indefinite preconditioners

Theorem

$$\lambda_{min}(\hat{\mathcal{K}}^{-1}\mathcal{K}) \geq \alpha \left[\frac{2}{\sqrt{5-1/\beta}+\sqrt{1-1/\beta}} \right]^2$$

$$\lambda_{max}(\hat{\mathcal{K}}^{-1}\mathcal{K}) \leq \beta\,(1+\sqrt{1-1/\beta})$$

Therefore

$$\kappa(\hat{\mathcal{K}}^{-1}\mathcal{K}) \leq c(\beta)\,\frac{\beta}{\alpha}$$

with

$$1 \leq c(\beta) \leq 3+\sqrt{5}, \quad c(1)=1$$

Symmetric indefinite preconditioners

Brezzi's theorem (for the case a is symmetric and $c \equiv 0$):

- a is bounded on X_h:

$$|a(x_h, w_h)| \le \|a\| \, \|x_h\|_X \|w_h\|_X \quad \text{for all } x_h, w_h \in X_h.$$

With

$$(x_h, w_h)_X = (\underline{X}_h \underline{x}_h, \underline{w}_h)_{\ell^2}$$

we obtain

$$A_h \le \|a\| \, \underline{X}_h$$

Symmetric indefinite preconditioners

Brezzi's theorem (for the case a is symmetric and $c \equiv 0$):

- a is bounded on X_h:

$$|a(x_h, w_h)| \leq \|a\| \, \|x_h\|_X \|w_h\|_X \quad \text{for all } x_h, w_h \in X_h.$$

  With

$$(x_h, w_h)_X = (\underline{X}_h \underline{x}_h, \underline{w}_h)_{\ell^2}$$

  we obtain

$$\underline{A}_h \leq \|a\| \, \underline{X}_h$$

- Compare with

$$\underline{A}_h \leq \hat{\underline{A}}_h$$

Symmetric indefinite preconditioners

- b is bounded on $X_h \times Q_h$:

$$|b(w_h, q_h)| \le \|b\| \, \|w_h\|_X \|q_h\|_Q \quad \text{for all } x_h \in X_h, \; q_h \in Q_h.$$

With

$$(p_h, q_h)_Q = (\underline{Q}_h \underline{p}_h, \underline{q}_h)_{\ell^2}$$

we obtain

$$B_h \underline{X}_h^{-1} B_h^T \le \|b\| \, \underline{Q}_h$$

Symmetric indefinite preconditioners

- b is bounded on $X_h \times Q_h$:

$$|b(w_h, q_h)| \leq \|b\| \, \|w_h\|_X \|q_h\|_Q \quad \text{for all } x_h \in X_h, \; q_h \in Q_h.$$

With

$$(p_h, q_h)_Q = (\underline{Q}_h \underline{p}_h, \underline{q}_h)_{\ell^2}$$

we obtain

$$B_h \underline{X}_h^{-1} B_h^T \leq \|b\| \, \underline{Q}_h$$

- Compare with

$$B_h \hat{A}_h^{-1} B_h^T \leq \beta \, \underline{S}_h$$

Symmetric indefinite preconditioners

- a is coercive on $\ker B_h$:

$$a(w_h, w_h) \geq \alpha_0 \, \|w_h\|_X^2 \quad \text{for all } w_h \in \ker B_h$$

becomes

$$(A_h \underline{w}_h, \underline{w}_h)_{\ell^2} \geq \alpha_0 \, (\underline{X}_h \, \underline{w}_h, \underline{w}_h)_{\ell^2} \quad \text{for all } \underline{w}_h \in \ker B_h$$

Symmetric indefinite preconditioners

- a is coercive on $\ker B_h$:

$$a(w_h, w_h) \geq \alpha_0 \, \|w_h\|_X^2 \quad \text{for all } w_h \in \ker B_h$$

becomes

$$(A_h \underline{w}_h, \underline{w}_h)_{\ell^2} \geq \alpha_0 \, (\underline{X}_h \, \underline{w}_h, \underline{w}_h)_{\ell^2} \quad \text{for all } \underline{w}_h \in \ker B_h$$

- Compare with

$$(A_h \underline{w}_h, \underline{w}_h)_{\ell^2} \geq \alpha \, (\hat{A}_h \, \underline{w}_h, \underline{w}_h)_{\ell^2} \quad \text{for all } \underline{w}_h \in \ker B_h$$

Symmetric indefinite preconditioners

- b satisfies a uniform inf-sup condition on $X_h \times Q_h$:

$$\sup_{0 \neq w_h \in X_h} \frac{b(w_h, q_h)}{\|w_h\|_X} \geq k_0 \, \|q_h\|_Q \quad \text{for all } q_h \in Q_h.$$

becomes

$$B_h \, \underline{X}_h^{-1} B_h^T \geq k_0 \, \underline{Q}_h$$

- Compare with

$$B \, \hat{A}^{-1} B^T \geq \hat{S}$$

Symmetric indefinite preconditioners

Choice:

$$\hat{A}_h = \frac{1}{\sigma} \underline{X}_h \quad \text{and} \quad \hat{S}_h = \frac{\sigma}{\tau} \underline{Q}_h$$

for $\sigma > 0$ and $\tau > 0$. If

$$\sigma < \frac{1}{\|a\|} \quad \text{and} \quad \tau > \frac{1}{k_0^2},$$

then the conditions are satisfied with

$$\alpha = \sigma \, \alpha_0 < \frac{\alpha_0}{\|a\|} \quad \text{and} \quad \beta = \tau \, \|b\|^2 > \frac{\|b\|^2}{k_0^2}.$$

Example: a model problem from optimal control

Find $(y,u) \in H^1(\Omega) \times L^2(\Omega)$ and $p \in H^1(\Omega)$ such that

$$
\begin{aligned}
\underbrace{(y,z)_{L^2(\Omega)} + \gamma\,(u,v)_{L^2(\Omega)}}_{a((y,u),(z,v))} + \underbrace{(z,p)_{H^1(\Omega)} - (v,p)_{L^2(\Omega)}}_{b((z,v),p)} &= (y_d, z)_{L^2(\Omega)} \\
\underbrace{(y,q)_{H^1(\Omega)} - (u,q)_{L^2(\Omega)}}_{b((y,u),q)} &= 0
\end{aligned}
$$

for all $(z,v) \in H^1(\Omega) \times L^2(\Omega)$ and $q \in H^1(\Omega)$.

Discretization: $\mathcal{T}_h$ triangulation of Ω, P_1-P_1-P_1 element

$$
\begin{pmatrix} A_h & B_h^T \\ B_h & 0 \end{pmatrix} \begin{pmatrix} \underline{x}_h \\ \underline{p}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ 0 \end{pmatrix}, \quad A_h = \begin{pmatrix} M_h & 0 \\ 0 & \gamma\, M_h \end{pmatrix}, \quad B_h = (K_h, -M_h)
$$

with $\underline{x}_h = (\underline{y}_h, \underline{u}_h)^T$, mass matrix M_h and stiffness matrix K_h.

Example: a model problem from optimal control

$X = H^1(\Omega) \times L^2(\Omega)$:

$$(x, w)_X \;=\; (y, z)_{H^1(\Omega)} + (u, v)_{L^2(\Omega)}$$

$Q = H^1(\Omega)$:

$$(p, q)_Q \;=\; (p, q)_{H^1(\Omega)}$$

On $X_h \times Q_h$:

$$\underline{X}_h = \begin{pmatrix} K_h & 0 \\ 0 & M_h \end{pmatrix}, \quad \underline{Q}_h = K_h.$$

$$\|a\| = \max(1, \gamma),\ \|b\| = \sqrt{2},\ \alpha_0 = \frac{\gamma}{2},\ k_0 = 1 \Longrightarrow \kappa(\hat{\mathcal{K}}_h^{-1}\mathcal{K}_h) = O(\frac{1}{\gamma})$$

Example: a model problem from optimal control

$X = H^1(\Omega) \times L^2(\Omega)$:

$$(x, w)_X \;=\; (y, z)_{L^2(\Omega)} + \sqrt{\gamma}\,(y, z)_{H^1(\Omega)} + \gamma\,(u, v)_{L^2(\Omega)}$$

$Q = H^1(\Omega)$:

$$(p, q)_Q \;=\; \frac{1}{\gamma}\left[(p, q)_{L^2(\Omega)} + \sqrt{\gamma}\,(p, q)_{H^1(\Omega)}\right]$$

On $X_h \times Q_h$:

$$\underline{X}_h = \begin{pmatrix} \underline{Y}_h & 0 \\ 0 & \gamma\, M_h \end{pmatrix}, \quad \underline{Q}_h = \frac{1}{\gamma}\,\underline{Y}_h \quad \text{with} \quad \underline{Y}_h = M_h + \sqrt{\gamma}\, K_h$$

$$\|a\| = 1, \quad \|b\| = 1, \quad \alpha_0 = \frac{2}{3}, \quad k_0 = \sqrt{\frac{3}{4}} \Longrightarrow \kappa(\hat{\mathcal{K}}_h^{-1}\mathcal{K}_h) \approx 4$$

Example: a model problem from optimal control

Summary:

$$\mathcal{K}_h \begin{pmatrix} \underline{x}_h \\ \underline{p}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ 0 \end{pmatrix} \quad \text{with} \quad \mathcal{K}_h = \begin{pmatrix} A_h & B_h^T \\ B_h & 0 \end{pmatrix}$$

with

$$A_h = \begin{pmatrix} M_h & 0 \\ 0 & \gamma\, M_h \end{pmatrix}, \quad B_h = \begin{pmatrix} K_h & -M_h \end{pmatrix}$$

Preconditioner

$$\hat{\mathcal{K}}_h = \begin{pmatrix} \hat{A}_h & B_h^T \\ B_h & B_h \hat{A}_h^{-1} B_h^T - \hat{S}_h \end{pmatrix}$$

where

$$\hat{A}_h = \frac{1}{\sigma} \begin{pmatrix} \underline{Y}_h & 0 \\ 0 & \gamma\, M_h \end{pmatrix} \quad \text{and} \quad \hat{S}_h = \frac{\sigma}{\tau} \frac{1}{\gamma} \underline{Y}_h$$

with

$$\underline{Y}_h = M_h + \sqrt{\gamma}\, K_h$$

Example: a model problem from optimal control

Summary:

$$\mathcal{K}_h \begin{pmatrix} \underline{x}_h \\ \underline{p}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ 0 \end{pmatrix} \quad \text{with} \quad \mathcal{K}_h = \begin{pmatrix} A_h & B_h^T \\ B_h & 0 \end{pmatrix}$$

with

$$A_h = \begin{pmatrix} M_h & 0 \\ 0 & \gamma\, M_h \end{pmatrix}, \quad B_h = \begin{pmatrix} K_h & -M_h \end{pmatrix}$$

Preconditioner

$$\hat{\mathcal{K}}_h = \begin{pmatrix} \hat{A}_h & B_h^T \\ B_h & B_h \hat{A}_h^{-1} B_h^T - \hat{S}_h \end{pmatrix}$$

where

$$\hat{A}_h = \frac{1}{\sigma} \begin{pmatrix} \hat{Y}_h & 0 \\ 0 & \gamma\, \hat{M}_h \end{pmatrix} \quad \text{and} \quad \hat{S}_h = \frac{\sigma}{\tau} \frac{1}{\gamma}\, \hat{Y}_h$$

with

$$\hat{Y}_h \approx \underline{Y}_h = M_h + \sqrt{\gamma}\, K_h \quad \text{and} \quad \hat{M}_h \approx M_h.$$

Numerical experiments

Geometry and mesh:

- $\Omega = (0,1)^3$.
- initial mesh $l = 1$: 24 tetrahedra
- uniform refinement, final mesh $l = L$.

Preconditioner:

- $\hat{Y}_h$ one V(3,3)-cycle of the multigrid method
- $\hat{M}_h$ 3 steps of the symmetric Gauß-Seidel method

Iterative method

- CG-method for the preconditioned problem
- Starting values $\underline{x}_h^{(0)}$ and $\underline{p}_h^{(0)}$ randomly generated.
- stopping rule $r^{(k)} \le \varepsilon\, r^{(0)}$ with $\varepsilon = 10^{-8}$.

Numerical experiments

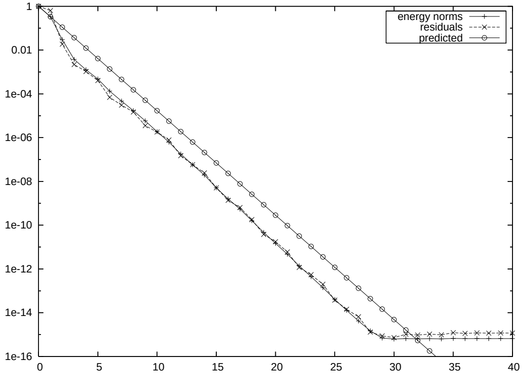

Figure: Number of iterations versus relative accuracy.

Numerical experiments

Table: Fixed regularization parameter $\gamma = 1$.

level L	number of unknowns $n + m$	iterations k	cpu time t (in seconds)
3	1 107	14	0.06
4	7 395	15	0.61
5	53 955	15	6.96
6	412 035	16	62.04
7	3 200 227	15	559.16

Table: Fixed refinement level $L = 5$.

γ	iterations k
10^{-4}	15
10^{-2}	14
1	15
10^{2}	14
10^{4}	15

Multigrid methods

Problem:

$$\mathcal{B}((x_h, p_h), (w_h, q_h)) = \langle \mathcal{F}, (w_h, q_h) \rangle \quad (w_h, q_h) \in X_h \times Q_h.$$

Two-grid method: $X_H \subset X_h, \quad Q_H \subset Q_h$

$(x_h^{(0)}, p_h^{(0)})$ starting value

1. Smoothing step:

$$(x_h^{(k+1)}, p_h^{(k+1)}) = \mathcal{S}_h\,(x_h^{(k)}, p_h^{(k)}) \quad k = 0, \ldots, m-1.$$

$(x_h^{(m)}, p_h^{(m)})$:

$$\langle \mathcal{R}, (w_h, q_h) \rangle = \langle \mathcal{F}, (w_h, q_h) \rangle - \mathcal{B}((x_h^{(m)}, p_h^{(m)}), (w_h, q_h))$$

2. Coarse grid correction:

$$\mathcal{B}((s_H, r_H), (w_H, q_H)) = \langle \mathcal{R}, (w_H, q_H) \rangle \quad (w_H, q_H) \in X_H \times Q_H$$

$$(x_h^{(m+1)}, p_h^{(m+1)}) = (x_h^{(m)}, p_h^{(m)}) + (s_H, r_H)$$

Multigrid methods

1. Smoothing step:

$$(x_h^{(k+1)}, p_h^{(k+1)}) = \mathcal{S}_h(x_h^{(k)}, p_h^{(k)}) \quad k = 0, \ldots, m-1.$$

iterative method as smoother:

$$\begin{pmatrix} \underline{x}_h^{(k+1)} \\ \underline{p}_h^{(k+1)} \end{pmatrix} = \begin{pmatrix} \underline{x}_h^{(k)} \\ \underline{p}_h^{(k)} \end{pmatrix} + \hat{\mathcal{K}}_h^{-1} \left[\begin{pmatrix} \underline{f}_h \\ \underline{g}_h \end{pmatrix} - \mathcal{K}_h \begin{pmatrix} \underline{x}_h^{(k)} \\ \underline{p}_h^{(k)} \end{pmatrix} \right]$$

with

$$\hat{\mathcal{K}}_h = \begin{pmatrix} \hat{A}_h & B_h^T \\ B_h & B_h \hat{A}_h^{-1} B_h^T - \hat{S}_h \end{pmatrix}$$

Multigrid methods

Theorem

If

$$A \le \hat{A}, \quad C + B\hat{A}^{-1}B^T \le \hat{S},$$

then, for $\mathcal{M} = I - \hat{\mathcal{K}}^{-1}\mathcal{K}$,

$$\|\mathcal{K}\mathcal{M}^m\|_{\mathcal{L}} \le \eta_0(m)\, \|\mathcal{Q}\|_{\mathcal{L}},$$

where

$$\mathcal{Q} = \begin{pmatrix} \hat{A} - A & 0 \\ 0 & C + B\hat{A}^{-1}B^T - \hat{S} \end{pmatrix},$$

$\mathcal{L}$ *is an arbitrary symmetric and positive definite matrix, and*

$$\eta_0(m) = \frac{1}{2^{m-1}} \binom{m-1}{[m/2]} = O\left(\frac{1}{\sqrt{m}}\right).$$

Multigrid convergence analysis

For $w \in X_h$, $q \in Q_h$:

$$|||(w,q)|||_{0,h} \quad \text{and} \quad |||(s,r)|||_{2,h} = \sup_{0 \neq (w,q) \in X_h \times Q_h} \frac{|\mathcal{B}((s,r),(w,q))|}{|||(w,q)|||_{0,h}}.$$

Approximation property:

$$|||\mathcal{B}|||_{0,h}\, |||(x_h^{(m+1)} - x_h, p_h^{(m+1)} - p_h)|||_{0,h} \leq c_A\, |||(x_h^{(m)} - x_h, p_h^{(m)} - p_h)|||_{2,h}$$

Smoothing property:

$$|||(x_h^{(m)} - x_h, p_h^{(m)} - p_h)|||_{2,h} \leq \eta(m)\, |||\mathcal{B}|||_{0,h}\, |||(x_h^{(0)} - x_h, p_h^{(0)} - p_h)|||_{0,h}$$

with

$$\eta(m) \to 0 \quad \text{for } m \to \infty.$$

Multigrid convergence analysis

For $w \in X_h$, $q \in Q_h$ with vector representations $\underline{w} \in \mathbb{R}^{n_h}$, $\underline{q} \in \mathbb{R}^{m_h}$:

$$|||(w,q)|||_{0,h} = \left(\left(\mathcal{L}_h \begin{pmatrix} \underline{w} \\ \underline{q} \end{pmatrix}, \begin{pmatrix} \underline{w} \\ \underline{q} \end{pmatrix} \right)_{\ell^2} \right)^{1/2} = \left\| \begin{pmatrix} \underline{w} \\ \underline{q} \end{pmatrix} \right\|_{\mathcal{L}_h}$$

Smoothing property

$$\|\mathcal{K}_h \mathcal{M}_h^m\|_{\mathcal{L}_h} \le \eta(m)\, \|\mathcal{K}_h\|_{\mathcal{L}_h},$$

which follows with $\eta(m) = c_R\, \eta_0(m) = O(1/\sqrt{m})$ if

$$\hat{A}_h \ge A_h \quad \text{and} \quad \hat{S}_h \ge C_h + B_h \hat{A}_h^{-1} B_h^T$$

and

$$\|\mathcal{Q}_h\|_{\mathcal{L}_h} \le c_R\, \|\mathcal{K}_h\|_{\mathcal{L}_h} \quad \text{with} \quad \mathcal{Q}_h = \begin{pmatrix} \hat{A}_h - A_h & 0 \\ 0 & C_h + B_h \hat{A}_h^{-1} B_h^T - \hat{S}_h \end{pmatrix}$$

Example: a model problem from optimal control

Find $(y,u) \in H^1(\Omega) \times L^2(\Omega)$ and $p \in H^1(\Omega)$ such that

$$\begin{aligned}
\underbrace{(y,z)_{L^2(\Omega)} + \gamma\,(u,v)_{L^2(\Omega)}}_{a((y,u),(z,v))} + \underbrace{(z,p)_{H^1(\Omega)} - (v,p)_{L^2(\Omega)}}_{b((z,v),p)} &= (y_d,z)_{L^2(\Omega)} \\
\underbrace{(y,q)_{H^1(\Omega)} - (u,q)_{L^2(\Omega)}}_{b((y,u),q)} &= 0
\end{aligned}$$

for all $(z,v) \in H^1(\Omega) \times L^2(\Omega)$ and $q \in H^1(\Omega)$.

Discretization:

$$\begin{pmatrix} A_h & B_h^T \\ B_h & 0 \end{pmatrix} \begin{pmatrix} \underline{x}_h \\ \underline{p}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ 0 \end{pmatrix}, \quad A_h = \begin{pmatrix} M_h & 0 \\ 0 & \gamma\,M_h \end{pmatrix}, \quad B_h = (K_h, -M_h)$$

with $\underline{x}_h = (\underline{y}_h, \underline{u}_h)^T$, mass matrix M_h and stiffness matrix K_h.

Example: a model problem from optimal control

Mesh dependent norm on $X_h \times Q_h$:

$$\begin{aligned} |||(w,q)|||_{0,h} &= h \left(\|\underline{z}\|_{\ell^2} + h^2 \|\underline{v}\|_{\ell^2} + \|\underline{q}\|_{\ell^2} \right)^{1/2} \\ &= \left(\left(\mathcal{L}_h \begin{pmatrix} \underline{w} \\ \underline{q} \end{pmatrix}, \begin{pmatrix} \underline{w} \\ \underline{q} \end{pmatrix} \right)_{\ell^2} \right)^{1/2} \end{aligned}$$

with $w = (z, v) \in X_h$, $q \in Q_h$ and

$$\mathcal{L}_h = h^2 \begin{pmatrix} I & 0 & 0 \\ 0 & h^2 I & 0 \\ 0 & 0 & I \end{pmatrix}.$$

Theorem

The approximation property is satisfied.

Example: a model problem from optimal control

Choice of $\hat{A}_h$ and $\hat{S}_h$:

$$\hat{A}_h = \frac{1}{\sigma} \begin{pmatrix} \operatorname{diag} K_h & 0 \\ 0 & \gamma \operatorname{diag} M_h \end{pmatrix}, \quad \hat{S}_h = \frac{1}{\tau} \operatorname{diag}\left(B_h^T \hat{A}_h^{-1} B_h^T \right).$$

with $\sigma > 0$, $\tau > 0$ such that

$$\hat{A}_h \geq A_h \quad \text{and} \quad \hat{S}_h \geq B_h \hat{A}^{-1} B_h^T.$$

Observe:

$$\hat{A}_h \neq \operatorname{diag} A_h = \frac{1}{\sigma} \begin{pmatrix} \operatorname{diag} M_h & 0 \\ 0 & \gamma \operatorname{diag} M_h \end{pmatrix}$$

Theorem

The smoothing property is satisfied with $\eta(m) = O(1/\sqrt{m})$.

Example: a model problem from optimal control

Theorem

There exists a constant $c > 0$ such that

$$\|\!|(x_h^{(m+1)} - x_h, p_h^{(m+1)} - p_h)|\!\|_{0,h} \leq \frac{c}{\sqrt{m}} \|\!|(x_h^{(0)} - x_h, p_h^{(0)} - p_h)|\!\|_{0,h}$$

Numerical experiments

Geometry and mesh:

- $\Omega = [0,1]^2$
- initial mesh $l = 1$: 2 triangles
- uniform refinement, final mesh $l = L$.

Iterative method

- Starting values $\underline{x}_h^{(0)}$ and $\underline{p}_h^{(0)}$ randomly generated.
- stopping rule $r^{(k)} \leq \varepsilon\, r^{(0)}$ with $\varepsilon = 10^{-8}$.

		smoothing steps					
level L	$n+m$	5+5		7+7		10+10	
5	3 267	46	0.668	30	0.538	21	0.411
6	12 675	48	0.679	34	0.578	24	0.455
7	49 923	49	0.685	35	0.587	25	0.467
8	198 147	49	0.685	35	0.588	25	0.469
9	789 507	49	0.685	35	0.589	25	0.469

Example: a model problem from optimal control

Second approach:
Find $y \in H^1(\Omega)$ and $p \in H^1(\Omega)$ such that

$$\underbrace{(y,z)_{L^2(\Omega)}}_{a(y,z)} + \underbrace{(z,p)_{H^1(\Omega)}}_{b(z,p)} = (y_d, z)_{L^2(\Omega)} \quad \text{for all } z \in H^1(\Omega)$$

$$\underbrace{(y,q)_{H^1(\Omega)}}_{b(y,q)} - \underbrace{\frac{1}{\gamma}(p,q)_{L^2(\Omega)}}_{c(p,q)} = 0 \quad \text{for all } q \in H^1(\Omega)$$

Discretization:

$$\mathcal{K}_h \begin{pmatrix} \underline{y}_h \\ \underline{p}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ 0 \end{pmatrix} \quad \text{with} \quad \mathcal{K}_h = \begin{pmatrix} A_h & B_h^T \\ B_h & -C_h \end{pmatrix} = \begin{pmatrix} M_h & K_h \\ K_h & -\frac{1}{\gamma} M_h \end{pmatrix}$$

with the mass matrix M_h and the stiffness matrix K_h.

Example: a model problem from optimal control

$$\mathcal{K}_h \begin{pmatrix} \underline{p}_h \\ \underline{y}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ 0 \end{pmatrix} \quad \text{with} \quad \mathcal{K}_h = \begin{pmatrix} A_h & B_h^T \\ B_h & -C_h \end{pmatrix} = \begin{pmatrix} K_h & M_h \\ M_h & -\gamma\, K_h \end{pmatrix}$$

Smoother:

$$\begin{pmatrix} p_h^{(k+1)} \\ y_h^{(k+1)} \end{pmatrix} = \begin{pmatrix} p_h^{(k)} \\ y_h^{(k)} \end{pmatrix} + \rho\, \hat{\mathcal{K}}_h^{-1} \left[\begin{pmatrix} f_h \\ 0 \end{pmatrix} - \mathcal{K}_h \begin{pmatrix} p_h^{(k)} \\ y_h^{(k)} \end{pmatrix} \right]$$

with $1 < \rho < 2$,

$$\hat{\mathcal{K}}_h = \begin{pmatrix} \hat{A}_h & B_h^T \\ B_h & B_h \hat{A}_h^{-1} B_h^T - \hat{S}_h \end{pmatrix}$$

and

$$\hat{A}_h = \frac{1}{\sigma} \operatorname{diag} K_h, \quad \hat{S}_h = \frac{1}{\tau} \operatorname{diag} \left(\gamma\, K_h + M_h \hat{A}_h^{-1} M_h \right).$$

Example: a model problem from optimal control

Numerical experiments:

		smoothing steps					
level L	$n+m$	1+1		2+2		3+3	
5	2 178	16	0.301	9	0.127	7	0.067
6	8 450	16	0.302	9	0.128	7	0.066
7	33 282	16	0.302	10	0.135	7	0.067
8	132 098	16	0.302	10	0.135	7	0.067
9	526 338	16	0.302	10	0.135	7	0.068

Dependence on γ:

γ	Iter.	conv. rate
1	16	0.302
10^{-2}	16	0.302
10^{-4}	16	0.302
10^{-6}	16	0.302

Concluding remarks

- General approach: one-shot multigrid
- Exploit special properties
  (much better results for the reduced form)
- Block preconditioners: require more information
- Rigorous analysis
- Other approaches: Braess-Sarazin, ...