SOME NOTES WITH NUMERICAL METHODS FOR STATIONARY PDES

ROB STEVENSON

1. Interpolation estimates in Sobolev spaces

Theorem 1.1 (Bramble-Hilbert "lemma"). *Let $\Omega \subset \mathbb{R}^n$ be a Lipschitz domain, and for some $m \in \mathbb{N}$, $q \in [1, \infty]$ and a normed space Y, let $L : W_q^m(\Omega) \to Y$ be a bounded linear mapping with $P_{m-1} \subset \mathrm{Ker} L$. Then $\exists C = C(\Omega)$ such that*

$$\|Lv\|_Y \leq C \|L\|_{W_q^m(\Omega) \to Y} |v|_{W_q^m(\Omega)} \quad (v \in W_q^m(\Omega)).$$

Lemma 1.2 (transformation lemma). *Let $G(\hat{x}) = B\hat{x} + c$ with $\det B \neq 0$, and $\hat{\Omega}$ and Ω be Lipschitz domains in $\mathbb{R}^n$ with $G(\hat{\Omega}) = \Omega$.*
For $m \geq 0$, $p \in [1, \infty]$ and $v \in W_p^m(\Omega)$, $\hat{v} := v \circ G \in W_p^m(\hat{\Omega})$.
$\exists C = C(n, m, p)$ with

$$|\hat{v}|_{W_p^m(\hat{\Omega})} \leq C \|B\|_2^m |\det B|^{-1/p} |v|_{W_p^m(\Omega)} \quad (v \in W_p^m(\Omega)),$$
$$|v|_{W_p^m(\Omega)} \leq C \|B^{-1}\|_2^m |\det B|^{1/p} |\hat{v}|_{W_p^m(\hat{\Omega})} \quad (\hat{v} \in W_p^m(\hat{\Omega})).$$

Theorem 1.3. *Let $\Omega, \hat{\Omega} \subset \mathbb{R}^n$ and G as in Lemma 1.2. Let*

$$h_\Omega := \inf\{\mathrm{diam}(S) : S \textit{ ball containing } \Omega\}$$
$$\rho_\Omega := \sup\{\mathrm{diam}(S) : S \textit{ ball in } \Omega\}$$

and let $h_{\hat{\Omega}}$ and $\rho_{\hat{\Omega}}$ be defined similarly. Then $\|B\|_2 \leq \frac{h_\Omega}{\rho_{\hat{\Omega}}}$, $\|B^{-1}\|_2 \leq \frac{h_{\hat{\Omega}}}{\rho_\Omega}$.

Theorem 1.4. *Let $\Omega, \hat{\Omega} \subset \mathbb{R}^n$ and G as in Lemma 1.2.*
Let $k, m \in \mathbb{N}_0$ and $p, q \in [1, \infty]$ be such that $W_p^{k+1}(\hat{\Omega}) \hookrightarrow W_q^m(\hat{\Omega})$, and let $\hat{\Pi} : W_p^{k+1}(\hat{\Omega}) \to W_q^m(\hat{\Omega})$ be a bounded linear mapping that preserves polynomials of degree k.
Define Π by $\Pi(v) \circ G = \hat{\Pi}(v \circ G)$.
Then $\exists C = C(\hat{\Pi}, \hat{\Omega})$, thus independent of Ω, *such that*

$$|v - \Pi(v)|_{W_q^m(\Omega)} \leq C (\mathrm{vol}(\Omega))^{\frac{1}{q} - \frac{1}{p}} \frac{h_\Omega^{k+1}}{\rho_\Omega^m} |v|_{W_p^{k+1}(\Omega)} \quad (v \in W_p^{k+1}(\Omega)).$$

Date: May 19, 2010.

2. Application to estimate local interpolation errors

Theorem 2.1. *Let $(\hat{K}, \hat{P}, \hat{N})$ be a finite element with s denoting the maximal order of partial derivatives occurring in the definition of $\hat{N}$. For some $m, k \in \mathbb{N}_0$, $p, q \in [1, \infty]$, let*

$$W_p^{k+1}(\hat{K}) \hookrightarrow C^s(\hat{K})$$
$$W_p^{k+1}(\hat{K}) \hookrightarrow W_q^m(\hat{K})$$
$$P_k(\hat{K}) \subset \hat{P} \subset W_q^m(\hat{K})$$

Then $\exists C = C(\hat{K}, \hat{P}, \hat{N})$ such that for all (K, P, N) that are affine interpolation equivalent to $(\hat{K}, \hat{P}, \hat{N})$,

$$|v - I_K v|_{W_q^m(K)} \leq C(\mathrm{vol}(K))^{\frac{1}{q}-\frac{1}{p}} \frac{h_K^{k+1}}{\rho_K^m} |v|_{W_p^{k+1}(K)} \quad (v \in W_p^{k+1}(K)).$$

Definition 2.2. A family of finite elements (K, P, N) is called *uniformly shape regular* when $\sup_K h_K/\rho_K < \infty$.

Corollary 2.3. *For a family of uniformly shape regular affine interpolation equivalent finite elements, result from Theorem 2.1 reads as*

$$|v - I_K v|_{W_q^m(K)} \leq C(\mathrm{vol}(K))^{\frac{1}{q}-\frac{1}{p}} h_K^{k+1-m} |v|_{W_p^{k+1}(K)} \quad (v \in W_p^{k+1}(K)).$$

3. Application to estimate global interpolation errors

Theorem 3.1. *Consider family $(\tau_h)_h$ of conforming subdivisions of a domain $\Omega \subset \mathbb{R}^n$ into n-simplices or n-rectangles that are uniformly shape regular, and such that all finite elements are affine interpolation equivalent to a reference element $(\hat{K}, \hat{P}, \hat{N})$.*
Let the global finite element spaces V_{τ_h} be continuous,

$$P_k(\hat{K}) \subset \hat{P} \subset H^1(\hat{K})$$
$$H^{k+1}(\hat{K}) \hookrightarrow C^s(\hat{K}) \quad (s \text{ as in Thm 2.1})$$

Then for $k + 1 \geq m \in \{0, 1\}$, with $h := \sup_{K \in \tau_h} \mathrm{diam}(K)$,

$$\|(I - I_{\tau_h})v\|_{H^m(\Omega)} \lesssim h^{k+1-m} |v|_{H^{k+1}(\Omega)} \quad (v \in H^{k+1}(\Omega)). \tag{1}$$

Remark 3.2. In these notes, by $C \lesssim D$ we will mean that C can be bounded on some absolute multiple of D, independently of parameters which C and D may depend on. Obviously, $C \gtrsim D$ is defined as $D \lesssim C$, and $C \eqsim D$ as $C \lesssim D$ and $C \gtrsim D$.

Remark 3.3. If the global finite element space is C^ℓ, then a statement similar to Theorem 3.1 is valid with the condition $m \in \{0, 1\}$ reading as $m \in \{0, \ldots, \ell + 1\}$.

Remark 3.4 (homogeneous Dirichlet boundary conditions). Let $V_{\tau_h,0}$ be a Lagrangian finite element space V_{τ_h} as in Thm 3.1 in which the degrees of freedom that are function evaluations in boundary nodes as well as the corresponding global basis function are removed. Then restricted to $H^1_0(\Omega)$, $I_{\mathcal{T}_h}$ maps into $V_{\mathcal{T}_h,0}$. Formula (1) now reads as

$$\|(I - I_{\mathcal{T}_h})v\|_{H^m(\Omega)} \lesssim h^{k+1-m}|v|_{H^{k+1}(\Omega)} \quad (v \in H^{k+1}(\Omega) \cap H^1_0(\Omega)).$$

Moreover, $V_{\mathcal{T}_h,0} \subset H^1_0(\Omega)$.

Using the Lax-Milgram lemma and Cea's lemma, we arrive at the following corollary.

Theorem 3.5. *Consider the situation of Th 3.1. Let* $a : H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$ *be bilinear, bounded, coercive,* $F : H^1(\Omega) \to \mathbb{R}$ *linear and bounded. Let* $u \in H^1(\Omega)$, $u_{\mathcal{T}_h} \in V_{\mathcal{T}_h}$ *be the solutions of*

$$\begin{aligned} a(u, v) &= F(v) \quad (v \in H^1(\Omega), \\ a(u_h, v_h) &= F(v_h) \quad (v_h \in V_{\mathcal{T}_h}), \end{aligned}$$

respectively. Then

$$\|u - u_h\|_{H^1(\Omega)} \lesssim h^k |u|_{H^{k+1}(\Omega)}$$

assuming $u \in H^{k+1}(\Omega)$.

Remark 3.6. Same conclusion when variational problem is formulated on $H^1_0(\Omega)$ and $V_{\mathcal{T}_h}$ reads as $V_{\mathcal{T}_h,0}$.

Under additional assumptions, higher order convergence can be demonstrated in the weaker $L_2(\Omega)$-norm:

Theorem 3.7 (Aubin-Nitsche duality 'trick')**.** *Let* $a(,)$ *be as in Thm 3.5. Suppose that for* $f \in L_2(\Omega)$*, the solution* $u_f \in H^1(\Omega)$ *(or in* $H^1_0(\Omega)$ *in case of hom. Dir.) of the adjoint problem* $a(v, u_f) = \int_\Omega fv dx$ $(v \in H^1(\Omega))$ $(H^1_0(\Omega))$ *is in* $H^2(\Omega)$ *with*

$$\|u_f\|_{H^2(\Omega)} \lesssim \|f\|_{L_2(\Omega)} \tag{2}$$

(this is known as a regularity condition*). Let* $(V_{\mathcal{T}_h})_h$ $((V_{\mathcal{T}_h,0})_h)$ *be such that*

$$\inf_{v_h \in V_{\mathcal{T}_h}} \|w - v_h\|_{H^1(\Omega)} \lesssim h\|w\|_{H^2(\Omega)} \text{ for all } w \in H^2(\Omega) \ (H^2(\Omega) \cap H^1_0(\Omega)). \tag{3}$$

Then for u *and* u_h *as in Thm 3.5, we have*

$$\|u - u_h\|_{L_2(\Omega)} \lesssim h\|u - u_h\|_{H^1(\Omega)}.$$

Proof. Let $w \in H^1(\Omega)$ $(H^1_0(\Omega))$ be the solution of the adjoint problem $a(v, w) = (u - u_h, v)_{L_2(\Omega)}$ $(v \in H^1(\Omega))$ $(H^1_0(\Omega))$. Then for any $w_h \in V_{\mathcal{T}_h}$ $(V_{\mathcal{T}_h,0})$,

$$\|u - u_h\|^2_{L_2(\Omega)} = a(u - u_h, w) = a(u - u_h, w - w_h) \lesssim \|u - u_h\|_{H^1(\Omega)}\|w - w_h\|_{H^1(\Omega)}$$

Using that $\inf_{w_h} \|w - w_h\|_{H^1(\Omega)} \lesssim h\|w\|_{H^2(\Omega)} \lesssim h\|u - u_h\|_{L_2(\Omega)}$, the proof is completed. □

Example 3.8. If $\Omega \subset \mathbb{R}^n$ has a C^2 boundary or is convex, then for $f \in L_2(\Omega)$, the solution $u \in H_0^1(\Omega)$ of $\int_\Omega \nabla u \cdot \nabla v dx = \int_\Omega f v dx$ ($v \in H_0^1(\Omega)$) is in $H^2(\Omega)$ and satisfies $\|u\|_{H^2(\Omega)} \lesssim \|f\|_{L_2(\Omega)}$. (Without such conditions on Ω, this regularity result is generally *not* true).

4. Inverse inequality

Theorem 4.1. *Let $(V_{\tau_h})_h$ be a family of affine equivalent f.e. spaces w.r.t. family $(\tau_h)_h$ of uniformly shape regular subdivisions of $\Omega \subset \mathbb{R}^n$. Let $h_{\min} := \min_{K\in\tau_h} \operatorname{diam}(K)$. Let $V_{\tau_h} \subset W_p^m(\Omega)$. Then on V_{τ_h},*

$$\|\cdot\|_{W_p^m(\Omega)} \lesssim h_{\min}^{-m} \|\cdot\|_{L_p(\Omega)}.$$

Proof. By the transformation lemma, equivalence of norms on finite dimensional spaces, and again the transformation lemma, for $v \in V_{\tau_h}$ we have

$$\begin{aligned}
|v|^p_{W_p^m(\Omega)} &= \sum_{K\in\tau_h} |v|_K|^p_{W_p^m(K)} \lesssim \sum_{K\in\tau_h} \|B^{-1}\|^{mp} |\det B| |\widehat{v|_K}|^p_{W_p^m(\hat{K})} \\
&\eqsim \sum_{K\in\tau_h} \|B^{-1}\|^{mp} |\det B| \|\widehat{v|_K}\|^p_{L_p(\hat{K})} \lesssim \sum_{K\in\tau_h} \|B^{-1}\|^{mp} \|v|_K\|^p_{L_p(K)} \\
&\lesssim \sum_{K\in\tau_h} \left(\frac{\hat{h}}{\rho_K}\right)^{mp} \|v|_K\|^p_{L_p(K)} \lesssim h_{\min}^{-pm} \|v\|^p_{L_p(\Omega)}.
\end{aligned}$$

□

Literature with Sections 1–4:

References

[Cia78] P.G. Ciarlet. *The finite element method for elliptic problems*. North-Holland, Amsterdam, 1978.

5. Matrix-vector formulation of finite element discretization

Let V be some finite dimension subspace of some real Hilbert space H, let $a : V \times V \to \mathbb{R}$ be bilinear, bounded and coercive, and let $f : V \to \mathbb{R}$ be linear and bounded (e.g., a and f are restrictions to V of (bi)linear forms on H having those properties). We consider the problem of finding $u \in V$ s.t.

$$a(u, v) = f(v) \quad (v \in V) \tag{4}$$

Defining $A : V \to V'$ by $(Au)(v) = a(u, v)$ an equivalent formulation is given by

$$Au = f. \tag{5}$$

Let $\Phi = \{\phi_1, \ldots, \phi_N\}$ be a basis for V. The corresponding dual basis $\Phi' = \{\phi_1', \ldots, \phi_N'\}$ for V' is defined by $\phi_i'(\phi_j) = \delta_{ij}$.

Exercise -1. Let $\mathbf{A} \in \mathbb{R}^{N\times N}$ be defined by $\mathbf{A}_{ij} = a(\phi_j, \phi_i)$, called the *stiffness matrix*.

- Show that $\mathbf{A}$ is the representation of A w.r.t. primal and dual bases of V and V', respectively, i.e., if $v = \sum_j \mathbf{v}_j \phi_j$, then $Av = \sum_i (\mathbf{A}\mathbf{v})_i \phi_i'$. Conclude that an equivalent formulation of (4) or (5) is given by $\mathbf{A}\mathbf{u} = \mathbf{f}$, where $u = \sum_i \mathbf{u}_i \phi_i$, $f = \sum_i \mathbf{f}_i \phi_i'$.
- With $u = \sum_i \mathbf{u}_i \phi_i$, $v = \sum_i \mathbf{v}_i \phi_i$, $f = \sum_i \mathbf{f}_i \phi_i'$, i.e., $\mathbf{f}_i = f(\phi_i)$, and $\langle \cdot, \cdot \rangle$ the standard scalar product on $\mathbb{R}^N$, show that $\langle \mathbf{A}\mathbf{u}, \mathbf{v}\rangle = a(u, v)$ and $f(v) = \langle \mathbf{f}, \mathbf{v}\rangle$.

Unless stated otherwise, with the norm $\|\cdot\|$ on $\mathbb{R}^N$ (or on $\mathbb{R}^{N\times N}$) we will always mean the standard norm (or the corresponding operator norm).

Exercise 0.
- Show that $a(\cdot,\cdot)$ is symmetric iff $\mathbf{A} = \mathbf{A}^T$.
- Show that $a(\cdot,\cdot)$ is coercive iff $\mathbf{A}$ is positive definite (denoted as $\mathbf{A} > 0$), i.e. $\langle \mathbf{A}\mathbf{v}, \mathbf{v}\rangle > 0$ for all $0 \neq \mathbf{v} \in \mathbb{R}^N$.

Remark 5.1. With the notations of Exercise -1, we have

$$\langle \mathbf{A}\mathbf{u}, \mathbf{v}\rangle = a(\sum_j \mathbf{u}_j \phi_j, \sum_i \mathbf{v}_i \phi_i) = \sum_{ij} \mathbf{u}_j \mathbf{v}_i \sum_K a(\phi_j|_K, \phi_i|_K).$$

The (set of non-zero entries of) the matrix $a(\phi_j|_K, \phi_i|_K)$ is known as the *element stiffness matrix*.

6. Conditioning of the stiffness matrix

Let $V \subset L_2(\Omega)$. $\mathbf{M} \in \mathbb{R}^{N\times N}$ defined by $\mathbf{M}_{ij} = (\phi_j, \phi_i)_{L_2(\Omega)}$ is called the *mass matrix*. Note that $\mathbf{M}$ is SPD.

Lemma 6.1. *If $\{\psi_1, \ldots, \psi_m\}$ is an independent set in a normed space $(V, \|\cdot\|)$, then $\|\sum_i \mathbf{c}_i \psi_i\|^2 \eqsim \sum_i |\mathbf{c}_i|^2$ (i.e. uniformly in $\mathbf{c} \in \mathbb{R}^m$).*

Proof. $\mathbf{c} \mapsto \|\sum_i \mathbf{c}_i \psi_i\|$ is continuous, so it attains a maximum and minimum on the unit ball in $\mathbb{R}^m$. By the independence of the set, the minimum is strictly positive. □

Theorem 6.2. *Let $(V_{\mathcal{T}_h})_h$ be a family of affine equivalent f.e. spaces w.r.t. a family of quasi-uniform, uniformly shape regular subdivisions of $\Omega \subset \mathbb{R}^n$. Then $\mathbf{M} = \mathbf{M}_h$ corresponding to the* nodal basis *is uniformly well-conditioned, i.e., $\sup_h \kappa(\mathbf{M}) < \infty$, where $\kappa(\mathbf{M}) = \|\mathbf{M}\|\|\mathbf{M}^{-1}\| = \frac{\lambda_{\max}(\mathbf{M})}{\lambda_{\min}(\mathbf{M})}$ is the* spectral condition number *of* $\mathbf{M}$.

Proof. By the choice of the basis, in the relation $v = \sum_i \mathbf{v}_i \phi_i$ we have $\mathbf{v}_i = N_i(v)$ where N_i denotes the ith global degree of freedom.

$$\begin{aligned}\langle \mathbf{M}\mathbf{v}, \mathbf{v}\rangle = \|v\|^2_{L_2(\Omega)} &= \sum_K \|v|_K\|^2_{L_2(K)} = \sum_K |\det B| \|\widehat{v|_K}\|^2_{L_2(\hat{K})} \\ &\eqsim h^n \sum_K \|\widehat{v|_K}\|^2_{L_2(\hat{K})} \overset{\text{Lemma 6.1}}{\eqsim} h^n \sum_K \sum_j |\hat{N}_j^{\text{loc}}(\widehat{v|_K})|^2 \\ &\overset{\text{affine eq.}}{=} h^n \sum_K \sum_j |N_j^{\text{loc}}(v|_K)|^2 \eqsim h^n \sum_i |N_i(v)|^2 = h^n \|\mathbf{v}\|^2. \square\end{aligned}$$

Theorem 6.3. *For $\Omega \subset \mathbb{R}^n$, let $(V_{\tau_h})_h \subset H^m(\Omega)$ (or $\subset H_0^m(\Omega)$) be a family of f.e. spaces. Let $a(\cdot,\cdot) : H^m(\Omega) \times H^m(\Omega) \to \mathbb{R}$ be bil., bound. and coercive (or with $H^m(\Omega)$ reading as $H_0^m(\Omega)$). Then the stiffness matrix $\mathbf{A} = \mathbf{A}_h$ w.r.t. a basis of V_{τ_h} satisfies $\|\mathbf{A}\| \lesssim h^{-2m}\|\mathbf{M}\|$ and $\|\mathbf{A}^{-1}\| \lesssim \|\mathbf{M}^{-1}\|$, with $\mathbf{M}$ being the corresponding mass matrix.*

Proof. Using Theorem 4.1, we have

$$\begin{aligned}|\langle \mathbf{A}\mathbf{v}, \mathbf{w}\rangle| = |a(v,w)| &\lesssim \|v\|_{H^m(\Omega)} \|w\|_{H^m(\Omega)} \lesssim h^{-2m} \|v\|_{L_2(\Omega)} \|w\|_{L_2(\Omega)} \\ &\lesssim h^{-2m} \lambda_{\max}(\mathbf{M}) \|\mathbf{v}\| \|\mathbf{w}\|,\end{aligned}$$

or $\|\mathbf{A}\| \lesssim h^{-2m} \lambda_{\max}(\mathbf{M})$. On the other hand

$$\langle \mathbf{A}\mathbf{v}, \mathbf{v}\rangle \gtrsim \|v\|^2_{H^m(\Omega)} \geq \|v\|^2_{L_2(\Omega)} \gtrsim \lambda_{\min}(\mathbf{M}) \|\mathbf{v}\|^2,$$

and so

$$\|\mathbf{A}^{-1}\mathbf{v}\|^2 \lesssim \lambda_{\min}(\mathbf{M})^{-1} \langle \mathbf{v}, \mathbf{A}^{-1}\mathbf{v}\rangle \leq \lambda_{\min}(\mathbf{M})^{-1} \|\mathbf{v}\| \|\mathbf{A}^{-1}\mathbf{v}\|$$

or $\|\mathbf{A}^{-1}\mathbf{v}\| \lesssim \lambda_{\min}(\mathbf{M})^{-1}\|\mathbf{v}\|$ or $\|\mathbf{A}^{-1}\| \lesssim \lambda_{\min}(\mathbf{M})^{-1}$. $\square$

Remark 6.4. If the basis in Theorem 6.3 is the *nodal* basis, then under the conditions of Theorem 6.2 we have $\kappa(\mathbf{A}) \lesssim h^{-2m}$. Generally, this estimate is sharp.

7. Iterative methods

Let be given a linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b},$$

where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is SPD w.r.t. a scalar product $(\cdot,\cdot)$ on $\mathbb{R}^n$ (e.g., the standard scalar product $\langle \cdot,\cdot \rangle$). We define the energy scalar product $(\cdot,\cdot)_{\mathbf{A}} = (\mathbf{A}\cdot,\cdot)$ and (vector) norms $\|\cdot\| = (\cdot,\cdot)^{\frac{1}{2}}$ and $\|\cdot\|_{\mathbf{A}} = (\cdot,\cdot)^{\frac{1}{2}}_{\mathbf{A}}$. We will always use the same notation for a vector norm as for the associated matrix norm.

7.1. **The damped Richardson iteration.** This iteration is defined by

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \omega(\mathbf{b} - \mathbf{A}\mathbf{x}^k).$$

Exercise 1. Show that the error satisfies

$$\mathbf{x} - \mathbf{x}^k = (I - \omega\mathbf{A})^k(\mathbf{x} - \mathbf{x}^0),$$

and that the method converges if and only if $\omega \in (0, \frac{2}{\rho(\mathbf{A})})$.

Exercise 2. Proof that for both $|||\cdot||| = \|\cdot\|$ and for $|||\cdot||| = \|\cdot\|_{\mathbf{A}}$ it holds that $\sup_{\mathbf{x}^0 \neq \mathbf{x}} \frac{\|\mathbf{x}-\mathbf{x}^k\|}{\|\mathbf{x}-\mathbf{x}^0\|} = |||(I - \omega\mathbf{A})^k|||$, and that

$$\min_{\omega} |||(I - \omega\mathbf{A})^k|||^{\frac{1}{k}} = \rho(I - \omega_{\mathrm{opt}}\mathbf{A}) = \tfrac{\kappa(\mathbf{A})-1}{\kappa(\mathbf{A})+1},$$

where $\omega_{\mathrm{opt}} = \frac{2}{\lambda_{\min}(\mathbf{A})+\lambda_{\max}(\mathbf{A})}$. This means that the average convergence factor over an arbitrary number of k steps is minimal for $\omega = \omega_{\mathrm{opt}}$.

7.2. **Preconditioning.** We saw that the rate of convergence of the optimally damped Richardson iteration is determined by $\kappa(\mathbf{A})$. Note that $\kappa \to \frac{\kappa-1}{\kappa+1}$ is a monotonically increasing function on $[1, \infty)$ with limit 1. An obvious idea therefore is to apply damped Richardson on the equivalent system $\mathbf{B}^{-1}\mathbf{A}\mathbf{x} = \mathbf{B}^{-1}b$, where $\mathbf{B}$ is an approximation for $\mathbf{A}$ such that $\kappa(\mathbf{B}^{-1}\mathbf{A}) \ll \kappa(\mathbf{A})$, and such that system $\mathbf{B}y = z$ can be solved cheaply. Such an $\mathbf{B}$ is called a *preconditioner*. The resulting preconditioned damped Richardson iteration is given by

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \omega\mathbf{B}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}^k). \tag{6}$$

In the remainder we will consider $\mathbf{B}$ that are SPD w.r.t. $(\cdot,\cdot)$. We define

$$\tilde{\mathbf{A}} = \mathbf{B}^{-1}\mathbf{A}, \quad \tilde{\mathbf{b}} = \mathbf{B}^{-1}b$$

and $[\cdot,\cdot] = (\mathbf{B}\cdot,\cdot)$.

Exercise 3. Show that $\tilde{\mathbf{A}}$ is SPD w.r.t. $[\cdot,\cdot]$ and so also w.r.t $[\cdot,\cdot]_{\tilde{\mathbf{A}}} := [\tilde{\mathbf{A}}\cdot,\cdot]$.

Exercise 4. Show that for both $|||\cdot||| = [\cdot,\cdot]^{\frac{1}{2}}$ and for $|||\cdot||| = \|\cdot\|_{\tilde{\mathbf{A}}} := [\cdot,\cdot]_{\tilde{\mathbf{A}}}^{\frac{1}{2}}$,

$$\min_{\omega} |||(I - \omega\tilde{\mathbf{A}})^k|||^{\frac{1}{k}} = \rho(I - \omega_{\mathrm{opt}}\tilde{\mathbf{A}}) = \tfrac{\kappa(\tilde{\mathbf{A}})-1}{\kappa(\tilde{\mathbf{A}})+1}, \tag{7}$$

where $\omega_{\mathrm{opt}} = \frac{2}{\lambda_{\min}(\tilde{\mathbf{A}})+\lambda_{\max}(\tilde{\mathbf{A}})}$. Also show that $\|\cdot\|_{\mathbf{A}} = \|\cdot\|_{\tilde{\mathbf{A}}}$.

7.3. **(Preconditioned) Chebychev iteration.** As we have seen, the error in the iterands of the preconditioned Richardson iteration satisfy

$$\mathbf{x} - \mathbf{x}^k = p_k(\tilde{\mathbf{A}})(\mathbf{x} - \mathbf{x}^0), \tag{8}$$

where $p_k(\lambda) = (1 - \omega\lambda)^k$. The idea behind the Chebychev iteration is to vary in (6) the damping ω per iteration. Let ω_i the damping in the ith iteration, and let $p_k(\lambda) = \prod_{i=1}^{k}(1 - \omega_i\lambda)$.

Exercise 5. Show that for both $\|\!|\cdot|\!\| = [\cdot,\cdot]^{\frac{1}{2}}$ and for $\|\!|\cdot|\!\| = \|\cdot\|_{\tilde{\mathbf{A}}}$,

$$\sup_{\mathbf{x}^0 \neq \mathbf{x}} \frac{\|\!|\mathbf{x}-\mathbf{x}^k|\!\|}{\|\!|\mathbf{x}-\mathbf{x}^0|\!\|} = \|\!|p_k(\tilde{\mathbf{A}})|\!\| = \max_{\lambda \in \sigma(\tilde{\mathbf{A}})} |p_k(\lambda)|.$$

Preferably, we would choose the ω_i's such that with $P_k^1 := \{p \in P_k : p(0) = 1\}$,

$$\max_{\lambda \in \sigma(\tilde{\mathbf{A}})} |p_k(\lambda)| = \min_{q_k \in P_k^1} \max_{\lambda \in \sigma(\tilde{\mathbf{A}})} |q_k(\lambda)|,$$

but this seems not realizable, since then we should know $\sigma(\tilde{\mathbf{A}})$ (cf., however, §7.5). Often estimates $0 < \mu$ and ν are available for $\lambda_{\min}(\tilde{\mathbf{A}})$ and $\lambda_{\max}(\tilde{\mathbf{A}})$, respectively. Now the idea is to choose p_k in such a way that

$$\max_{\lambda \in [\mu,\nu]} |p_k(\lambda)| = \min_{q_k \in P_k^1} \max_{\lambda \in [\mu,\nu]} |q_k(\lambda)|.$$

It can be show that this minimalization problem has the following unique solution:

$$p_k(\lambda) = T_k(\tfrac{\nu+\mu-2\lambda}{\nu-\mu})/T_k(\tfrac{\nu+\mu}{\nu-\mu}),$$

where $T_k(\xi) = \frac{1}{2}[(\xi + \sqrt{\xi^2-1})^k - (\xi - \sqrt{\xi^2-1})^k]$ is the Chebychev polynomial of degree k. It holds that

$$\max_{\lambda \in [\mu,\nu]} |p_k(\lambda)| = 1/T_k(\tfrac{\nu+\mu}{\nu-\mu}) = \tfrac{2c^k}{1+c^{2k}}$$

where $c = \frac{\sqrt{\frac{\mu}{\nu}}-1}{\sqrt{\frac{\mu}{\nu}}+1}$, and so $\lim_{k\to\infty}(\max_{\lambda \in [\mu,\nu]} |p_k(\lambda)|)^{\frac{1}{k}} = c$. For $\mu = \lambda_{\min}(\tilde{\mathbf{A}})$ and $\nu = \lambda_{\max}(\tilde{\mathbf{A}})$, compare this result with (7). The iteration with error recursion (8) and this p_k is called (preconditioned) Chebychev iteration.

7.4. **Practical implementation.** Given $\tilde{\mathbf{A}}$, $\tilde{\mathbf{b}}$ and $\mathbf{x}^0$, let $\mathbf{x}^k$ denote the kth iterand of the Chebychev iteration.

Exercise 6. Given k, what are the damping parameters $\omega_1^{(k)}, \ldots, \omega_k^{(k)}$ in terms of the zeros of T_k, whose are needed for the implementation (6) of this iteration?

The zeros of T_k interlace the zeros of T_{k+1}. As a consequence, one has to restart the whole computation if instead of $\mathbf{x}_k$ one would like to have the more accurate $\mathbf{x}_{k+1}$. A solution is provided by the following: The Chebychev polynomials satisfy the following recursion:

$$\left\{ \begin{array}{c} T_0(z) = 1,\ T_1(z) = z \\ T_{k+1}(z) - 2zT_k(z) + T_{k-1}(z) = 0 \quad k = 1, 2 \ldots . \end{array} \right. \tag{9}$$

Exercise 7. Show that with $d_k = T_k(\frac{\nu+\mu}{\nu-\mu})$,

$$\left\{ \begin{array}{rcl} d_1(\mathbf{x}-\mathbf{x}^1) &=& (\frac{\nu+\mu}{\nu-\mu}I - \frac{2}{\nu-\mu}\tilde{\mathbf{A}})(\mathbf{x}-\mathbf{x}^0) \\ d_{k+1}(\mathbf{x}-\mathbf{x}^{k+1}) &=& (\frac{2(\nu+\mu)}{\nu-\mu}I - \frac{4}{\nu-\mu}\tilde{\mathbf{A}})d_k(\mathbf{x}-\mathbf{x}^k) - d_{k-1}(\mathbf{x}-\mathbf{x}^{k-1}) \quad k = 1, 2 \ldots . \end{array} \right.$$

By using (9) again, show that

$$\begin{cases} d_1\mathbf{x}^1 &= \frac{\nu+\mu}{\nu-\mu}\mathbf{x}^0 + \frac{2}{\nu-\mu}(\tilde{\mathbf{b}} - \tilde{\mathbf{A}}\mathbf{x}^0) \\ d_{k+1}\mathbf{x}^{k+1} &= \frac{2(\nu+\mu)}{\nu-\mu} d_k\mathbf{x}^k - d_{k-1}\mathbf{x}^{k-1} + \frac{4}{\nu-\mu} d_k(\tilde{\mathbf{b}} - \tilde{\mathbf{A}}\mathbf{x}^k) \quad k = 1, 2 \ldots, \end{cases}$$

This leads to the following practical (preconditioned) Chebychev algorithm

$$\begin{cases} \mathbf{x}^1 &= \mathbf{x}^0 + \frac{1}{2}\beta_0\mathbf{B}^{-1}(-\mathbf{A}\mathbf{x}^0) \\ \mathbf{x}^{k+1} &= \alpha_k\mathbf{x}^k + (1-\alpha_k)\mathbf{x}^{k-1} + \beta_k\mathbf{B}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}^k) \quad k = 1, 2 \ldots, \end{cases},$$

where using (9) it holds that $\beta_0 = \frac{4}{\nu+\mu}$ and

$$\beta_k = \left(\tfrac{\nu+\mu}{2} - \left(\tfrac{\nu-\mu}{4}\right)^2\beta_{k-1}\right)^{-1}, \quad \alpha_k = \tfrac{\nu+\mu}{2}\beta_k, \qquad k = 1, 2, \ldots.$$

7.5. The (preconditioned) conjugate gradient ((P)CG) algorithm.

Exercise 8. Using (8), show that for both the (preconditioned) damped Richardson and Chebychev iteration,

$$\mathbf{x}^k = \mathbf{x}^0 + p_{k-1}(\tilde{\mathbf{A}})(\tilde{\mathbf{b}} - \tilde{\mathbf{A}}\mathbf{x}^0)$$

for certain $p_{k-1} \in P_{k-1}$. Note that $\mathbf{x}^k$ can also be computed as $\mathbf{x}^k := \hat{\mathbf{x}}^k + \mathbf{x}^0$, where $\hat{\mathbf{x}}^k$ is the kth iterand of the process applied to $\tilde{\mathbf{A}}\hat{\mathbf{x}} = \hat{\mathbf{b}}$ with $\hat{\mathbf{b}} = \tilde{\mathbf{b}} - \tilde{\mathbf{A}}\mathbf{x}^0$ en $\hat{\mathbf{x}}^0 = 0$.

So w.l.o.g. we may assume that $\mathbf{x}^0 = 0$. From this we have

$$\mathbf{x}^k \in \mathcal{K}_k, \tag{10}$$

where $\mathcal{K}_k = \operatorname{span}\{\tilde{\mathbf{b}}, \tilde{\mathbf{A}}\tilde{\mathbf{b}}, \ldots, \tilde{\mathbf{A}}^{k-1}\tilde{\mathbf{b}}\}$ is the so-called *Krylov* subspace of order k ($\mathcal{K}_0 = \{0\}$).

Given a norm, the *optimal* $\mathbf{x}^k \in \mathcal{K}_k$ would satisfy

$$\|\mathbf{x} - \mathbf{x}^k\| = \min_{y \in \mathcal{K}_k} \|\mathbf{x} - y\|. \tag{11}$$

We now choose the energy norm $\|\cdot\|_{\tilde{\mathbf{A}}}$. We denote $\tilde{\mathbf{r}}^k = \tilde{\mathbf{b}} - \tilde{\mathbf{A}}\mathbf{x}^k$, i.e. $\tilde{\mathbf{r}}^k$ is the residual of $\mathbf{x}^k$ w.r.t. the preconditioned system.

Exercise 9. Show that $\tilde{\mathbf{r}}^k \in \mathcal{K}_{k+1}$, and that (11) is equivalent to

$$[\tilde{\mathbf{r}}^k, \mathcal{K}_k] = [\mathbf{x} - \mathbf{x}^k, \mathcal{K}_k]_{\tilde{\mathbf{A}}} = 0.$$

Suppose that $\mathbf{x}^k \in \mathcal{K}_k$ is optimal, and let $\tilde{\mathbf{r}}^k \neq 0$ (otherwise ready). Let $0 \neq \mathbf{p}^k \in \mathcal{K}_{k+1}$ such that $[\mathbf{p}^k, \mathcal{K}_k]_{\tilde{\mathbf{A}}} = 0$.

Exercise 10. With $\alpha^k_{\text{opt}} := \frac{[\mathbf{p}^k, \tilde{\mathbf{r}}^k]}{[\mathbf{p}^k, \mathbf{p}^k]_{\tilde{\mathbf{A}}}}$, show that

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k_{\text{opt}}\mathbf{p}^k \in \mathcal{K}_{k+1}$$

is the next optimal iterand, and that

$$\tilde{\mathbf{r}}^{k+1} = \tilde{\mathbf{r}}^k - \alpha^k_{\text{opt}}\tilde{\mathbf{A}}\mathbf{p}^k.$$

The question is how $\mathbf{p}^k$ should be determined. We take $\mathbf{p}^0 = \tilde{\mathbf{b}}$ and, using induction, assume that $\{\mathbf{p}^0, \ldots, \mathbf{p}^{k-1}\}$ is an $[\cdot,\cdot]_{\tilde{\mathbf{A}}}$-orthogonal basis of $\mathcal{K}_k$.

Exercise 11. From $\tilde{\mathbf{r}}^k \neq 0$, show that

$$\mathbf{p}^k := \tilde{\mathbf{r}}^k - \sum_{j=0}^{k-1} \frac{[\mathbf{p}^j, \tilde{\mathbf{r}}^k]_{\tilde{\mathbf{A}}}}{[\mathbf{p}^j, \mathbf{p}^j]_{\tilde{\mathbf{A}}}} \mathbf{p}^j.$$

Prove that $[\mathbf{p}^j, \tilde{\mathbf{r}}^k]_{\tilde{\mathbf{A}}} = 0$ when $j \leq k-2$, and thus that

$$\mathbf{p}^k := \tilde{\mathbf{r}}^k - \frac{[\mathbf{p}^{k-1}, \tilde{\mathbf{r}}^k]_{\tilde{\mathbf{A}}}}{[\mathbf{p}^{k-1}, \mathbf{p}^{k-1}]_{\tilde{\mathbf{A}}}} \mathbf{p}^{k-1}.$$

The resulting algorithm, known as the (preconditioned) conjugate gradient ((P)CG) algorithm[1], is therefore given by the following

$$\begin{cases} \mathbf{x}^0 := 0;\ \tilde{\mathbf{r}}^0 := \tilde{\mathbf{b}};\ \mathbf{p}^0 := \tilde{\mathbf{r}}^0; \\ \text{For } k = 0, 1, \ldots \text{ when } \tilde{\mathbf{r}}^k \neq 0, \text{ compute} \\ \alpha^k_{\text{opt}} := \frac{[\mathbf{p}^k, \tilde{\mathbf{r}}^k]}{[\mathbf{p}^k, \mathbf{p}^k]_{\tilde{\mathbf{A}}}}; \\ \mathbf{x}^{k+1} := \mathbf{x}^k + \alpha^k_{\text{opt}} \mathbf{p}^k; \\ \tilde{\mathbf{r}}^{k+1} := \tilde{\mathbf{r}}^k - \alpha^k_{\text{opt}} \tilde{\mathbf{A}} \mathbf{p}^k; \\ \mathbf{p}^{k+1} := \tilde{\mathbf{r}}^{k+1} - \frac{[\mathbf{p}^k, \tilde{\mathbf{r}}^{k+1}]_{\tilde{\mathbf{A}}}}{[\mathbf{p}^k, \mathbf{p}^k]_{\tilde{\mathbf{A}}}} \mathbf{p}^k; \end{cases}$$

Exercise 12. Show that, in exact arithmetic, the PCG algorithm yields the exact solution after at most n iterations. (In practice, one would like to compute much less iterations) . Prove that

$$\frac{\|\mathbf{x} - \mathbf{x}^k\|_{\tilde{\mathbf{A}}}}{\|\mathbf{x}\|_{\tilde{\mathbf{A}}}} \leq \frac{2c^k}{1 + c^{2k}}, \text{ with } c = \frac{\sqrt{\kappa(\tilde{\mathbf{A}})} - 1}{\sqrt{\kappa(\tilde{\mathbf{A}})} + 1}.$$

In contrast to damped Richardson and Chebychev iteration, note that the PCG iteration is non-linear; $\mathbf{x}^k$ is a linear combination of $\tilde{\mathbf{b}}, \tilde{\mathbf{A}}\tilde{\mathbf{b}}, \ldots, \tilde{\mathbf{A}}^{k-1}\tilde{\mathbf{b}}$ (cf. (10)) where the coefficients also depend on $\tilde{\mathbf{b}}$.

In view of obtaining an efficient implementation we now slightly rewrite the iteration.

Exercise 13. Show that $[\mathbf{p}^k, \tilde{\mathbf{r}}^k] = [\tilde{\mathbf{r}}^k, \tilde{\mathbf{r}}^k]$, and thus that

$$\alpha^k_{\text{opt}} = \frac{[\tilde{\mathbf{r}}^k, \tilde{\mathbf{r}}^k]}{[\mathbf{p}^k, \mathbf{p}^k]_{\tilde{\mathbf{A}}}}.$$

Show that $[\tilde{\mathbf{r}}^{k+1}, \tilde{\mathbf{r}}^{k+1}] = -\alpha^k_{\text{opt}} [\tilde{\mathbf{r}}^{k+1}, \mathbf{p}^k]_{\tilde{\mathbf{A}}}$ and so that

$$\mathbf{p}^{k+1} = \tilde{\mathbf{r}}^{k+1} + \frac{[\tilde{\mathbf{r}}^{k+1}, \tilde{\mathbf{r}}^{k+1}]}{[\tilde{\mathbf{r}}^k, \tilde{\mathbf{r}}^k]} \mathbf{p}^k.$$

[1]The name (P)CG algorithm can be explained by interpreting solving $\tilde{\mathbf{A}}\mathbf{x} = \tilde{\mathbf{b}}$ as minimizing $F(\mathbf{x}) := \frac{1}{2}[\tilde{\mathbf{A}}\mathbf{x}, \mathbf{x}] - [\tilde{\mathbf{b}}, \mathbf{x}]$. Note that the gradient of F in y equals minus the residual of y w.r.t. $\tilde{\mathbf{A}}\mathbf{x} = \tilde{\mathbf{b}}$, i.e. $F'(y) = -(\tilde{\mathbf{b}} - \tilde{\mathbf{A}}y)$. Vectors that are orthogonal w.r.t. to an energy scalar product are also called "conjugate".

We write $[\cdot,\cdot]$ and $[\cdot,\cdot]_{\tilde{\mathbf{A}}}$ in terms of $(\cdot,\cdot)$, $\mathbf{A}$ and $\mathbf{B}$, and introduce $\mathbf{r}^k = \mathbf{b} - \mathbf{A}\mathbf{x}^k$, i.e. $\mathbf{r}^k$ is the residual of $\mathbf{x}^k$ w.r.t. the unpreconditioned system. The final PCG algorithm now reads as

$$\begin{cases}
\mathbf{x}^0 := 0;\ \mathbf{r}^0 := \mathbf{b}; \\
\text{Compute}\,\tilde{\mathbf{r}}^0\,\text{from}\,\mathbf{B}\tilde{\mathbf{r}}^0 = \mathbf{r}^0;\ \mathbf{p}^0 = \tilde{\mathbf{r}}^0; \\
\text{For}\,k = 0,1,\ldots\ \text{when}\,\mathbf{r}^k \neq 0,\ \text{compute} \\
\alpha^k_{\text{opt}} := \frac{(\mathbf{r}^k,\tilde{\mathbf{r}}^k)}{(\mathbf{A}\mathbf{p}^k,\mathbf{p}^k)}; \\
\mathbf{x}^{k+1} := \mathbf{x}^k + \alpha^k_{\text{opt}}\mathbf{p}^k; \\
\mathbf{r}^{k+1} := \mathbf{r}^k - \alpha^k_{\text{opt}}\mathbf{A}\mathbf{p}^k; \\
\text{Compute}\,\tilde{\mathbf{r}}^{k+1}\,\text{from}\,\mathbf{B}\tilde{\mathbf{r}}^{k+1} = \mathbf{r}^{k+1}; \\
\mathbf{p}^{k+1} := \tilde{\mathbf{r}}^{k+1} + \frac{(\mathbf{r}^{k+1},\tilde{\mathbf{r}}^{k+1})}{(\mathbf{r}^k,\tilde{\mathbf{r}}^k)}\mathbf{p}^k;
\end{cases}$$

Literature with Section 7:

References

[GVL89] G.H. Golub and C.F. Van Loan. *Matrix computations*, volume 3 of *Johns Hopkins Series in the Mathematical Sciences*. Johns Hopkins University Press, Baltimore, MD, second edition, 1989.

[Hac91] W. Hackbusch. *Iterative Lösung großer schwachbesetzter Gleichungssysteme*. B.G. Teubner, Stuttgart, 1991.

8. Subspace correction methods

For a real linear space V of finite dimension N, $a : V \times V \to \mathbb{R}$ bilinear, bounded, symmetric and coercive (i.e., a is a scalar product on V), and $f \in V'$, consider the problem of finding $u \in V$ such that

$$a(u,v) = f(v) \quad (v \in V)$$

(e.g., a and f are the restrictions to V of such (bi)linear forms on an enclosing infinite dimensional Hilbert space). With Φ a basis for V, $\mathbf{u}$ ($\mathbf{v}$) denoting the representation of u (v) w.r.t. Φ, $\mathbf{f}$ the representation of f w.r.t. Φ' and $\mathbf{A}$ the stiffness matrix, an equivalent formulation is given by

$$\mathbf{A}\mathbf{u} = \mathbf{f},$$

where $\mathbf{A}$ is SPD w.r.t. $\langle\cdot,\cdot\rangle$ (cf. exercises -1 and 0).

Remark 8.1. If the problem is just given as a matrix vector problem $\mathbf{A}\mathbf{u} = \mathbf{f}$ with $\mathbf{A}$ being SPD, then V can be taken as $\mathbb{R}^N$ and $\Phi = \{e_1,\ldots,e_N\}$. Then $u = \sum_i \mathbf{u}_i e_i$, $v = \sum_i \mathbf{v}_i e_i$, $a(u,v) = \langle \mathbf{A}\mathbf{u},\mathbf{v}\rangle$, $f(v) = \langle \mathbf{f},\mathbf{v}\rangle$.

Let

$$V = \sum_{i=0}^{J} V_i$$

be a space decomposition, where $N_i = \dim V_i$. Given $0 \leq i \leq J$, let $P^i : V \to V_i$ be defined by

$$a(P^i w, v_i) = a(w, v_i) \quad (v_i \in V_i).$$

Exercise 14. Show that P^i is the a-orthogonal projector onto V_i.

Let $\tilde{u}$ be some approximation of u. Define $\tilde{u}^{\text{new}}$ by

$$\tilde{u}^{\text{new}} = \tilde{u} + w_i,$$

where w_i is the solution of

$$a(w_i, v_i) = f(v_i) - a(\tilde{u}, v_i) \quad (v_i \in V_i) \tag{12}$$

Noting that $w_i = P^i(u - \tilde{u})$, we infer that

$$u - \tilde{u}^{\text{new}} = (I - P^i)(u - \tilde{u}).$$

Unless $V_i = V$, this *subspace correction iteration* cannot be convergent. Indeed, components in the error $u - \tilde{u}$ that are a-orthogonal to V_i are not affected.

Before we consider iterations involving corrections on all subspaces, we consider the implementation of this basic step. Let $\mathbf{I}_i : \mathbb{R}^{N_i} \to \mathbb{R}^N$ denote the representation of the inclusion of V_i into V w.r.t. some basis Φ^i of V_i and the basis Φ of V, respectively.

Exercise 15. Let V_i be the space of continuous piecewise linears that vanish at the boundary w.r.t. a subdivision of $[0, 1]$ into 3 equal subintervals, and let V be such a space with 6 equal subintervals. Give the matrix $\mathbf{I}_i$ when the finite element spaces are equipped with nodal bases.

Recall that $\mathbf{A}$ is the stiffness matrix of a w.r.t. Φ, and that $\mathbf{f}$ denotes the representation of f w.r.t. Φ'. Let $\mathbf{w}_i$ ($\mathbf{v}_i$) denote representations of w_i (v_i) w.r.t. Φ^i. Let $\mathbf{A}_i$ denote the stiffness matrix of a (restricted to V_i) w.r.t. Φ^i. Let $\tilde{\mathbf{u}}$ ($\tilde{\mathbf{u}}^{\text{new}}$) denote the representation of $\tilde{u}$ ($\tilde{u}^{\text{new}}$) w.r.t. Φ. Then $a(w_i, v_i) = \langle \mathbf{A}_i \mathbf{w}_i, \mathbf{v}_i \rangle$, $f(v_i) = \langle \mathbf{f}, \mathbf{I}_i \mathbf{v}_i \rangle = \langle \mathbf{I}_i^T \mathbf{f}, \mathbf{v}_i \rangle$, $a(\tilde{u}, v_i) = \langle \mathbf{A}\tilde{\mathbf{u}}, \mathbf{I}_i \mathbf{v}_i \rangle = \langle \mathbf{I}_i^T \mathbf{A}\tilde{\mathbf{u}}, \mathbf{v}_i \rangle$, i.e., (12) reads as

$$\langle \mathbf{A}_i \mathbf{w}_i, \mathbf{v}_i \rangle = \langle \mathbf{I}_i^T \mathbf{f}, \mathbf{v}_i \rangle - \langle \mathbf{I}_i^T \mathbf{A}\tilde{\mathbf{u}}, \mathbf{v}_i \rangle \quad (\mathbf{v}_i \in \mathbb{R}^{N_i})$$

or

$$\mathbf{A}_i \mathbf{w}_i = \mathbf{I}_i^T (\mathbf{f} - \mathbf{A}\tilde{\mathbf{u}})$$

The representation of w_i w.r.t. Φ is given by $\mathbf{I}_i \mathbf{w}_i$, so that

$$\tilde{\mathbf{u}}^{\text{new}} = \tilde{\mathbf{u}} + \mathbf{I}_i \mathbf{A}_i^{-1} \mathbf{I}_i^T (\mathbf{f} - \mathbf{A}\tilde{\mathbf{u}}). \tag{13}$$

Remark 8.2. From $a(w_i, v_i) = \langle \mathbf{A}\mathbf{I}_i \mathbf{w}_i, \mathbf{I}_i \mathbf{v}_i \rangle = \langle \mathbf{I}_i^T \mathbf{A}\mathbf{I}_i \mathbf{w}_i, \mathbf{v}_i \rangle$, note that

$$\mathbf{A}_i = \mathbf{I}_i^T \mathbf{A}\mathbf{I}_i,$$

which gives an alternative way to compute $\mathbf{A}_i$ once $\mathbf{A}$ is available.

Often, one replaces the application of $\mathbf{A}_i^{-1}$ by that of some approximation $\mathbf{R}_i$, i.e.,

$$\tilde{\mathbf{u}}^{\text{new}} = \tilde{\mathbf{u}} + \mathbf{I}_i \mathbf{R}_i \mathbf{I}_i^T (\mathbf{f} - \mathbf{A}\tilde{\mathbf{u}}). \tag{14}$$

Having more than one subspace, one can choose to perform the corresponding (approximate) subspace corrections in parallel ((damped) *Parallel Subspace Correction* method)

$$\tilde{\mathbf{u}}^{\text{new}} = \tilde{\mathbf{u}} + \omega \sum_{i=0}^{J} \mathbf{I}_i \mathbf{R}_i \mathbf{I}_i^T (\mathbf{f} - \mathbf{A}\tilde{\mathbf{u}}),$$

or successively (*Successive Subspace Correction* method)

$$\begin{aligned} &\mathbf{v} \leftarrow \tilde{\mathbf{u}}, \\ &\texttt{for } i = 0 : J \texttt{ do } \mathbf{v} \leftarrow \mathbf{v} + \mathbf{I}_i \mathbf{R}_i \mathbf{I}_i^T (\mathbf{f} - \mathbf{A}\mathbf{v}) \texttt{ endfor}, \\ &\tilde{\mathbf{u}}^{\text{new}} \leftarrow \mathbf{v}. \end{aligned}$$

With the latter method, we have

$$\begin{aligned} &\mathbf{u} - \tilde{\mathbf{u}}^{\text{new}} = \\ &(\mathbf{I} - \mathbf{I}_J \mathbf{R}_J \mathbf{I}_J^T \mathbf{A})(\mathbf{I} - \mathbf{I}_{J-1} \mathbf{R}_{J-1} \mathbf{I}_{J-1}^T \mathbf{A}) \cdots (\mathbf{I} - \mathbf{I}_0 \mathbf{R}_0 \mathbf{I}_0^T \mathbf{A})(\mathbf{u} - \tilde{\mathbf{u}}). \end{aligned} \tag{15}$$

Remark 8.3. Without damping the PSC does not necessarily converge. Instead of applying a simple damping, assuming that the $\mathbf{R}_i$ are SPD, usually $\sum_{i=0}^J \mathbf{I}_i \mathbf{R}_i \mathbf{I}_i^T$ is used as a preconditioner in CG.

Exercise 16. Consider the situation of Remark 8.1. Let $J = N - 1$, $V_i = \text{span}\{e_{i+1}\}$ $(0 \leq i \leq J)$. Write $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$, where $\mathbf{L}_{k,l} = \mathbf{A}_{k,\ell}$ when $k > \ell$, $\mathbf{D}_{k,\ell} = \mathbf{A}_{k,\ell}$ when $k = \ell$, $\mathbf{U}_{k,\ell} = \mathbf{A}_{k,\ell}$ when $k < \ell$, and zeros otherwise. With exact subspace solvers, i.e., "$\mathbf{R}_i = \mathbf{A}_i^{-1}$", show that the (damped) PSC or SSC is given by $\tilde{\mathbf{u}}^{\text{new}} = \tilde{\mathbf{u}} + \omega \mathbf{D}^{-1}(\mathbf{f} - \mathbf{A}\tilde{\mathbf{u}})$, known as the (damped) Jacobi iteration, or $\tilde{\mathbf{u}}^{\text{new}} = \tilde{\mathbf{u}} + (\mathbf{L} + \mathbf{D})^{-1}(\mathbf{f} - \mathbf{A}\tilde{\mathbf{u}})$, known as Gauss-Seidel iteration, respectively.

Besides the Jacobi and Gauss-Seidel iterations, well-known classes of methods that fit into the framework of subspace correction methods are *domain decomposition methods*, and *multigrid method*, the latter being discussed in the next section.

9. Multigrid

Let $V_0 \subset V_1 \subset V_2 \subset \cdots \subset H$ be a nested sequence of finite element spaces on some domain $\Omega \subset \mathbb{R}^n$ contained in some Hilbert space H, with $N_i := \dim V_i$.

Let a be a scalar product on H, Φ^i be a basis for V_i, for $j \geq i$ let $\mathbf{I}_i^j$ be the representation of the inclusion of V_i in V_j, $\mathbf{I}_j^i := (\mathbf{I}_i^j)^T$, $\mathbf{A}_i$ be the stiffness matrix of a restricted to $V_i \times V_i$ w.r.t. Φ^i, and, for $i \geq 1$, let $\mathbf{R}_i$, $\tilde{\mathbf{R}}_i$ be some approximate inverses of $\mathbf{A}_i$.

We define multigrid in the form of a preconditioner $\mathbf{B}_J$ for $\mathbf{A}_J$, i.e., as an approximate inverse.

Algorithm 9.1. (multigrid) For $J = 0$, define $\mathbf{B}_0 = \mathbf{R}_0 = \mathbf{A}_0^{-1}$. Assume that $\mathbf{B}_{J-1}$ has been defined. For some fixed $m_1, m_2 \in \mathbb{N}$, the application of $\mathbf{B}_J$ to a vector $\mathbf{g} \in \mathbb{R}^{N_j}$ is defined as follows:

(1) Pre-smoothing: For $\mathbf{v}^0 = 0$ and $\ell = 1, 2, \ldots, m_1$,
$$\mathbf{v}^\ell = \mathbf{v}^{\ell-1} + \mathbf{R}_J(\mathbf{g} - \mathbf{A}_J\mathbf{v}^{\ell-1})$$
(2) Coarse grid correction:
$$\mathbf{w}^0 = \mathbf{v}^{m_1} + \mathbf{I}_{J-1}^J\mathbf{B}_{J-1}\mathbf{I}_J^{J-1}(\mathbf{g} - \mathbf{A}_J\mathbf{v}^{m_1})$$
(3) Post-smoothing: For $\ell = 1, 2, \ldots, m_2$,
$$\mathbf{w}^\ell = \mathbf{w}^{\ell-1} + \tilde{\mathbf{R}}_J(\mathbf{g} - \mathbf{A}_J\mathbf{w}^{\ell-1})$$

$\mathbf{B}_J\mathbf{g}$ is now defined as $\mathbf{w}^{m_2}$.

The matrices $\mathbf{I}_{J-1}^J$ and $\mathbf{I}_J^{J-1}$ are known as *prolongations* and *restrictions*, respectively.

Exercise 17. (i) For some $\rho > 1$, let $N_i \eqsim \rho^i$, and let the application of any of $\mathbf{A}_i$, $\mathbf{I}_i^{i-1}$, $\mathbf{I}_{i-1}^i$, $\mathbf{R}_i$, $\tilde{\mathbf{R}}_i$ to a vector takes $\mathcal{O}(N_i)$ operations. Show that the application of $\mathbf{B}_J$ to a vector takes $\mathcal{O}(N_J)$ operations (dependent on N_0).

(ii) With $0 < \omega < 1$ being a constant, show that the forthcoming condition (21) is satisfied for $\mathbf{R}_i = \frac{\omega}{\|\mathbf{A}_i\|}\mathbf{I}_i$.

(iii) Consider the Poisson problem on a polygon in $\mathbb{R}^2$. Let τ_0 be some initial conforming triangulation of Ω, and for $i \in \mathbb{N}$, let τ_{i+1} be defined from τ_i by connecting the midpoints of the edges of each triangle from τ_i. Let V_i be the continuous piecewise linear finite element space w.r.t. τ_i, and Φ^i be its nodal basis. Let $\mathbf{R}_i = \tilde{\mathbf{R}}_i$ be as in (ii). Show that the assumptions made in (i) are satisfied.

Since $m > 1$ steps of a stationary iterative method can be rewritten as a one step modified iterative method, it is sufficient to analyse the multigid method for $m_1, m_2 \in \{0, 1\}$ with $m_1 + m_2 > 0$. We start with considering the case that $m_1 = 1$ and $m_2 = 0$. The general case will be considered in Remark 9.5.

We will show that $\mathbf{B}_J$ can be used without CG acceleration or damping as a convergent "stand alone" iteration
$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{B}_J(\mathbf{b} - \mathbf{A}_J\mathbf{x}^k).$$

With $\mathbf{I}_J$ denoting the identity matrix on $\mathbb{R}^{N_j}$, one infers that
$$\mathbf{B}_J = \mathbf{R}_J + \mathbf{I}_{J-1}^J\mathbf{B}_{J-1}\mathbf{I}_J^{J-1}(\mathbf{I}_J - \mathbf{A}_J\mathbf{R}_J), \tag{16}$$
and so
$$\mathbf{I}_J - \mathbf{B}_J\mathbf{A}_J = (\mathbf{I}_J - \mathbf{I}_{J-1}^J\mathbf{B}_{J-1}\mathbf{I}_J^{J-1}\mathbf{A}_J)(\mathbf{I}_J - \mathbf{R}_J\mathbf{A}_J). \tag{17}$$

Exercise 18. For $0 \leq i \leq J$, set

$$\mathbf{P}_J^i := \mathbf{A}_i^{-1}\mathbf{I}_J^i\mathbf{A}_J.$$

(i) Show that $\mathbf{P}_J^i\mathbf{I}_i^J = \mathbf{I}_i$, and that $\mathbf{I}_i^J\mathbf{P}_J^i$ is the $\langle\cdot,\cdot\rangle_{\mathbf{A}_J}$-orthogonal projector onto $\mathrm{Im}\mathbf{I}_i^J$.

(ii) Show that $\mathbf{I}_i^J\mathbf{I}_{i-1}^i = \mathbf{I}_{i-1}^J$, and so that $\mathbf{I}_i^{i-1}\mathbf{I}_J^i = \mathbf{I}_J^{i-1}$.

(iii) From (16), with J reading as i, show that

$$\mathbf{I}_J - \mathbf{I}_i^J\mathbf{B}_i\mathbf{A}_i\mathbf{P}_J^i = (\mathbf{I}_J - \mathbf{I}_{i-1}^J\mathbf{B}_{i-1}\mathbf{A}_{i-1}\mathbf{P}_J^{i-1})(\mathbf{I}_J - \mathbf{I}_i^J\mathbf{R}_i\mathbf{A}_i\mathbf{P}_J^i). \tag{18}$$

Conclude that

$$\mathbf{I}_J - \mathbf{B}_J\mathbf{A}_J = (\mathbf{I}_J - \mathbf{I}_0^J\mathbf{R}_0\mathbf{I}_J^0\mathbf{A}_J)\cdots(\mathbf{I}_J - \mathbf{I}_J^J\mathbf{R}_J\mathbf{I}_J^J\mathbf{A}_J) \tag{19}$$

and thus that multigrid is an efficient implementation of SSC with the space decomposition $V_J = V_J + V_{J-1} + \ldots + V_0$.

Remark 9.2. A similar recursive implementation of PSC with the space decomposition $V_J = V_J + V_{J-1} + \ldots + V_0$ is known as the BPX (Bramble-Pasciak-Xu) preconditioner.

Below, we present the classical multigrid convergence theory showing that the method has the unique property to converge with a rate that does not deteriorate with an increasing problem size. The theory will require a regularity assumption (cf. (26)) that can be avoided by an analysis based on the concept of subspace corrections. This, however, goes beyond the scope of these notes.

Theorem 9.3. *For some constants c_0 and c_1 (i.e., independent of i), let*

$$\|(\mathbf{I}_i - \mathbf{I}_{i-1}^i\mathbf{P}_i^{i-1})\mathbf{v}\|_{\mathbf{A}_i}^2 \leq c_1\|\mathbf{A}_i\|^{-1}\|\mathbf{A}_i\mathbf{v}\|^2 \quad (\mathbf{v} \in \mathbb{R}^{N_i}), \tag{20}$$

and let the $\mathbf{R}_i$ be SPD, and satisfy

$$\lambda_{\min}(\mathbf{R}_i) \geq \frac{c_0}{\|\mathbf{A}_i\|} \text{ and } \rho(\mathbf{R}_i\mathbf{A}_i) \leq 1. \tag{21}$$

Then

$$\|\mathbf{I}_J - \mathbf{B}_J\mathbf{A}_J\|_{\mathbf{A}_J\to\mathbf{A}_J} \leq \frac{c_1}{2c_0 + c_1}. \tag{22}$$

Proof. With $\mathbf{E}_i := \mathbf{I}_i - \mathbf{B}_i\mathbf{A}_i$, $\mathbf{S}_i := \mathbf{I}_i - \mathbf{R}_i\mathbf{A}_i$, from (17) we infer that

$$\mathbf{E}_J = (\mathbf{I}_J - \mathbf{I}_{J-1}^J\mathbf{P}_J^{J-1} + \mathbf{I}_{J-1}^J\mathbf{E}_{J-1}\mathbf{P}_J^{J-1})\mathbf{S}_J.$$

Using Exercise 18(i), we find that for $\mathbf{v} \in \mathbb{R}^{N_J}$,

$$\|\mathbf{E}_J\mathbf{v}\|_{\mathbf{A}_J}^2 = \|(\mathbf{I}_J - \mathbf{I}_{J-1}^J\mathbf{P}_J^{J-1})\mathbf{S}_J\mathbf{v}\|_{\mathbf{A}_J}^2 + \|\mathbf{E}_{J-1}\mathbf{P}_J^{J-1}\mathbf{S}_J\mathbf{v}\|_{\mathbf{A}_{J-1}}^2, \tag{23}$$

where we used that

$$\|\mathbf{I}_{J-1}^J\mathbf{w}\|_{\mathbf{A}_J} = \|\mathbf{w}\|_{\mathbf{A}_{J-1}} \quad (\mathbf{w} \in \mathbb{R}^{N_{J-1}}). \tag{24}$$

From the conditions of the theorem, it follows that

$$\begin{aligned}
&\|(\mathbf{I}_J - \mathbf{I}_{J-1}^J \mathbf{P}_J^{J-1})\mathbf{S}_J \mathbf{v}\|_{\mathbf{A}_J}^2 \le c_1 \|\mathbf{A}_J\|^{-1} \|\mathbf{A}_J \mathbf{S}_J \mathbf{v}\|^2 \\
&\le \tfrac{c_1}{c_0} \langle \mathbf{R}_J \mathbf{A}_J \mathbf{S}_J \mathbf{v}, \mathbf{A}_J \mathbf{S}_J \mathbf{v}\rangle = \tfrac{c_1}{c_0} \langle (\mathbf{I}_J - \mathbf{S}_J)\mathbf{S}_J^2 \mathbf{v}, \mathbf{v}\rangle_{\mathbf{A}_J} \\
&\le \tfrac{c_1}{2c_0} (\|\mathbf{v}\|_{\mathbf{A}_J}^2 - \|\mathbf{S}_J \mathbf{v}\|_{\mathbf{A}_J}^2),
\end{aligned} \tag{25}$$

where in the last step we have used that $\sigma(\mathbf{R}_J \mathbf{A}_J) \subset [0,1]$, $\mathbf{S}_J$ is symmetric w.r.t. $\langle \cdot, \cdot \rangle_{\mathbf{A}_J}$, and the inequality $(1-t)t^2 \le \frac{1}{2}(1-t^2)$ for $t \in [0,1]$.

Let $\delta = \frac{c_1}{2c_0 + c_1}$. Obviously (22) is valid for $J = 0$. Assume that is holds for $J - 1$. Then from (23), (24), Exercise 18(i) and (25), we have

$$\begin{aligned}
\|\mathbf{E}_J\|_{\mathbf{A}_J}^2 &\le \|(\mathbf{I}_J - \mathbf{I}_{J-1}^J \mathbf{P}_J^{J-1})\mathbf{S}_J \mathbf{v}\|_{\mathbf{A}_J}^2 + \delta \|\mathbf{I}_{J-1}^J \mathbf{P}_J^{J-1} \mathbf{S}_J \mathbf{v}\|_{\mathbf{A}_J}^2 \\
&\le (1-\delta)\|(\mathbf{I}_J - \mathbf{I}_{J-1}^J \mathbf{P}_J^{J-1})\mathbf{S}_J \mathbf{v}\|_{\mathbf{A}_J}^2 + \delta \|\mathbf{S}_J \mathbf{v}\|_{\mathbf{A}_J}^2 \\
&\le (1-\delta)\tfrac{c_1}{2c_0}(\|\mathbf{v}\|_{\mathbf{A}_J}^2 - \|\mathbf{S}_J \mathbf{v}\|_{\mathbf{A}_J}^2) + \delta \|\mathbf{S}_J \mathbf{v}\|_{\mathbf{A}_J}^2 = \delta \|\mathbf{v}\|_{\mathbf{A}_J}^2,
\end{aligned}$$

which completes the proof. □

Theorem 9.4. *Let $H = H_0^1(\Omega)$. Let the solution $u_f \in H_0^1(\Omega)$ of $a(u_f, v) = \int_\Omega f v dx$ ($v \in H_0^1(\Omega)$) be in $H^2(\Omega)$ with*

$$\|u_f\|_{H^2(\Omega)} \lesssim \|f\|_{L_2(\Omega)} \tag{26}$$

(*cf.* (2))*. Let $h_{i-1} \lesssim h_i$, and*

$$\inf_{v^i \in V_i} \|w - v^i\|_{H^1(\Omega)} \lesssim h_i \|w\|_{H^2(\Omega)} \quad (w \in H^2(\Omega) \cap H_0^1(\Omega)) \tag{27}$$

(*cf.* (3))*. Let $\sup_i \kappa(\mathbf{M}_i) < \infty$ (cf. Theorem 6.2), and $\|\mathbf{A}_i\| \lesssim h_i^{-2}\|\mathbf{M}_i\|$ and $\|\mathbf{A}_i^{-1}\| \lesssim \|\mathbf{M}_i^{-1}\|$ (cf. Theorem 6.3). Then* (20) *is valid.*

Proof. Given $\mathbf{v}^i \in \mathbb{R}^{N_i}$, let $\mathbf{v}^{i-1} = \mathbf{P}_i^{i-1}\mathbf{v}^i$. Write $v^i = \sum_j \mathbf{v}_j^i \phi_j^i$ and $v^{i-1} = \sum_j \mathbf{v}_j^{i-1}\phi_j^{i-1}$. From Exercise 18(i), we have

$$\langle \mathbf{A}^i \mathbf{I}_{i-1}^i \mathbf{v}^{i-1}, \mathbf{I}_{i-1}^i \mathbf{z}^{i-1}\rangle = \langle \mathbf{A}^i \mathbf{v}^i, \mathbf{I}_{i-1}^i \mathbf{z}^{i-1}\rangle \quad (\mathbf{z}^{i-1} \in \mathbb{R}^{N_{i-1}}).$$

Define $\mathbf{f}^i = \mathbf{A}_i \mathbf{v}^i$, $\mathbf{g}^i = \mathbf{M}_i^{-1}\mathbf{f}^i$. Write $g^i = \sum_j \mathbf{g}_j^i \phi_j^i$, and for any $\mathbf{z}^i \in \mathbb{R}^{N_i}$, $z^i = \sum_j \mathbf{z}_j^i \phi_j^i$. Then

$$\langle \mathbf{f}^i, \mathbf{z}^i\rangle = \langle \mathbf{M}_i \mathbf{g}^i, \mathbf{z}^i\rangle = (g^i, z^i)_{L_2(\Omega)} \quad (z^i \in V_i),$$

and

$$\|g^i\|_{L_2(\Omega)}^2 = \langle \mathbf{M}_i \mathbf{g}^i, \mathbf{g}^i\rangle = \langle \mathbf{f}^i, \mathbf{M}_i^{-1}\mathbf{f}^i\rangle \le \|\mathbf{M}_i^{-1}\|\|\mathbf{f}^i\|^2. \tag{28}$$

Let $v \in H_0^1(\Omega)$ denote the solution of $a(v, z) = (g^i, z)_{L_2(\Omega)}$ ($z \in H_0^1(\Omega)$). We have

$$a(v^i, z^i) = \langle \mathbf{A}_i \mathbf{v}^i, \mathbf{z}^i\rangle = \langle \mathbf{f}^i, \mathbf{z}^i\rangle = (g^i, z^i)_{L_2(\Omega)} \quad (z^i \in V_i),$$

and, writing $z^{i-1} = \sum_j \mathbf{z}_j^{i-1}\phi_j^{i-1}$,

$$\begin{aligned}
a(v^{i-1}, z^{i-1}) &= \langle \mathbf{A}_{i-1}\mathbf{v}^{i-1}, \mathbf{z}^{i-1}\rangle = \langle \mathbf{A}_i \mathbf{I}_{i-1}^i \mathbf{v}^{i-1}, \mathbf{I}_{i-1}^i \mathbf{z}^{i-1}\rangle \\
&= \langle \mathbf{A}_i \mathbf{v}^i, \mathbf{I}_{i-1}^i \mathbf{z}^{i-1}\rangle = (g^i, z^{i-1})_{L_2(\Omega)} \quad (z^{i-1} \in V_{i-1}).
\end{aligned}$$

From (26) and (27), we infer that

$$\|v^i - v^{i-1}\|_{H^1(\Omega)} \leq \|v^i - v\|_{H^1(\Omega)} + \|v - v^{i-1}\|_{H^1(\Omega)} \lesssim h_i \|g^i\|_{L_2(\Omega)}. \tag{29}$$

From

$$\|\mathbf{v}^i - \mathbf{I}_{i-1}^i \mathbf{v}^{i-1}\|_{\mathbf{A}_i} = a(v^i - v^{i-1}, v^i - v^{i-1})^{\frac{1}{2}} \eqsim \|v^i - v^{i-1}\|_{H^1(\Omega)},$$

(29), (28), $\mathbf{f}^i = \mathbf{A}_i \mathbf{v}^i$ and the remaining assumptions of the theorem, the proof is completed. □

Remark 9.5. Let us denote $\mathbf{B}_J$ for general $m_1, m_2 \in \{0, 1\}$ as $\mathbf{B}_J^{(m_1,m_2)}$. Along the lines we derived (19), it is not hard to show that

$$\mathbf{I}_J - \mathbf{B}_J^{(1,1)} \mathbf{A}_J = (\mathbf{I}_J - \mathbf{B}_J^{(0,1)} \mathbf{A}_J)(\mathbf{I}_J - \mathbf{B}_J^{(1,0)} \mathbf{A}_J)$$

and that $\mathbf{I}_J - \mathbf{B}_J^{(0,1)} \mathbf{A}_J = (\mathbf{A}_J(\mathbf{I}_J - \tilde{\mathbf{B}}_J^{(1,0)} \mathbf{A}_J)\mathbf{A}_J^{-1})^T$, where $\tilde{\mathbf{B}}_J^{(1,0)}$ is $\mathbf{B}_J^{(1,0)}$ with any $\mathbf{R}_i$ reading as $\tilde{\mathbf{R}}_i^T$. A direct consequence is that $\|\mathbf{I}_J - \mathbf{B}_J^{(0,1)} \mathbf{A}_J\|_{\mathbf{A}_J \to \mathbf{A}_J} = \|\mathbf{I}_J - \tilde{\mathbf{B}}_J^{(1,0)} \mathbf{A}_J\|_{\mathbf{A}_J \to \mathbf{A}_J}$.

Literature with Sections 8–9:

References

[BH83] D. Braess and W. Hackbusch. A new convergence proof for the multigrid method including the V-cycle. *SIAM J. Numer. Anal.*, 20(5):967–975, 1983.

[Xu01] J. Xu. The method of subspace corrections. *J. Comput. Appl. Math.*, 128(1-2), 2001. Numerical analysis 2000, Vol. VII, Partial differential equations.