Notes on Applied Mathematics

Interative Methods and Preconditioning

Guido Kanschat

March 27, 2015

Preface

This text is part of a set of notes I prepare for students and my own reference. They mainly serve the purpose of being short and concise introductions to mathematical topics. They are provided as is and in the hope that they are useful. Nevertheless, I am always thankful for possible corrections and suggestions for enhancements.

The material in these notes is not my original research. Most of it is adapted from textbooks and research publications. While I am striving to give credit to the original authors wherever possible, I will be delighted to include more citations, also in order to improve the value of these notes as a reference.

Finally, if you find these notes useful for your own research and decide to cite results from them, I would be most flattered if you decided to cite them as

> Guido Kanschat. Notes on Applied Mathematics. Universität Heidelberg. 0000.

Note: yellow boxes indicate text which is missing in the current version and will be added soon.

Contents

Chapter 1

Preliminaries

Notation 1.1. Very often in these notes, we will abbreviate the term "there is a positive constant C independent of certain parameters, such that $x \leq Cy$" by the expression

$$x \lesssim y. \tag{1.1}$$ eq:main:1

Furthermore, if additionally for $C \geq 1$, there holds $y \leq Cx$, then we write

$$x \simeq y.$$

Sometimes, it may be useful to be able refer to the implicit constant in such equations, which will be done in the form $C_{1.1}$.

Chapter 2

From finite to countable dimensions

Introduction 2.1. Linear algebra deals with abstract vector spaces, but most results on linear mappings are restricted to finite dimensional spaces, since they exploit the fact that we can choose a basis.

The choice of a basis becomes a more involved endeavor if we allow for spaces that do not have a finite basis. We can actually go by two very different routes. The route of Hamel bases, which are bases of inifintely many vectors, but in order to represent a vector in such a basis, we only allow for finite linear combinations.

The other route defines a Schauder basis as a set of vectors, such that every vector in the space is the linear combination of infinitely many basis vectors. In order to define such a linear combination, we have to define the meaning of such an infinite sum, namely the convergence of the sum. In the course of such a definition, we will learn about a natural extension of Euclidean[1] spaces, namely pre-Hilbert and Hilbert spaces.

2.1 Hilbert spaces and orthogonal bases

Definition 2.2. Let V be a vector space over $\mathbb{K}$ with $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$. An **inner product** on V is a mapping $\langle .,. \rangle : V \times V \to \mathbb{K}$ with the properties

$$\langle \alpha x + y, z \rangle = \alpha \langle x, z \rangle + \langle y, z \rangle \qquad \forall x, y, z \in V; \alpha \in \mathbb{K} \tag{2.1}$$

$$\langle x, y \rangle = \overline{\langle y, x \rangle} \qquad \forall x, y \in V \tag{2.2}$$

$$\langle x, x \rangle \geq 0 \quad \forall x \in V \qquad \text{and } \langle x, x \rangle = 0 \Leftrightarrow x = 0, \tag{2.3}$$

[1] And we will not have to distinguish between Euclidean real spaces and unitary complex spaces anymore.

usually referred to as (bi-)linearity, symmetry, and definiteness. We note that linearity in the second argument follows immediately by symmetry.

Definition 2.3. A vector space V equipped with an inner product $\langle ., . \rangle$ and a norm defined by

$$v = \sqrt{\langle v, v \rangle}$$

is called an **inner product space** or **pre-Hilbert space**. A **Hilbert space** is a pre-Hilbert space which is also complete, that is, every Cauchy sequence with elements in the space has a limit in the space.

Example 2.4. For any positive integer, the space $\mathbb{R}^n$ equipped with the Euclidean inner product

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

is a Hilbert space. The same holds for $\mathbb{C}^n$ and

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

Example 2.5. The spaces $\ell^2(\mathbb{R})$ and $\ell^2(\mathbb{C})$ of sequences $\{x_k\}_{k=1,\dots}$ of real and complex numbers, respectively, are Hilbert spaces, if equipped with the inner product

$$\langle x, y \rangle = \sum_{i=k}^{\infty} x_k \overline{y_k} = \lim_{n \to \infty} \sum_{k=1}^{n} x_k \overline{y_k}.$$

An example for a sequence in $\ell^2(\mathbb{R})$ is for instance the sequence $v = \{1/k\}$, since

$$||v||^2 = \sum_{i=k}^{\infty} \frac{1}{k^2} < \infty.$$

The sequence $w = \{1\}$ is not, since it does not converge quadratically.

Example 2.6. On the space of continuous functions on the interval $[-\pi/2, \pi/2]$ define the inner product

$$\langle f, g \rangle = \int_{-\pi/2}^{\pi/2} f(x) g(x)\, dx.$$

Let

$$V = \left\{ f \in C[-\pi/2, \pi/2] \middle| \langle f, f \rangle < \infty \right\}.$$

Then V is a vector space with an inner product and thus a pre-Hilbert space, but it is not a Hilbert space, since for any n the sum

$$f_n(x) = \frac{4}{\pi} \sum_{k=1}^{n} \frac{\sin\big((2k-1)x\big)}{2k-1}$$

is continuous, but

$$\lim_{n\to\infty} f_n = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0 \end{cases}$$

is not.

Definition 2.7. Let V be a vector space over a field $\mathbb{K}$. A basis of V is a set $\{x_i\}$ of linearly independent vectors with coefficients $i \in I$ from an index set I, such that each $v \in V$ has a representation of the form

$$v = \sum_{i \in I} \alpha_i x_i,$$

with coefficients $\alpha_i \in \mathbb{K}$. For a **Hamel basis**, it is required that only finitely many coefficients in this representation are nonzero. For a **Schauder basis**, we assume $I = \mathbb{N}$ and require that the sum in the linear combination exists as the limit of a series.

Notation 2.8. We will use the term sequence to denote an at most countable set. The elements of a sequence are numbered by indices and the index set is $\mathbb{N}$ or a subset thereof.

Definition 2.9. Let V be an inner product space over a field $\mathbb{K}$. Two vectors $x, y \in V$ are called orthogonal if $\langle x, y\rangle = 0$. We write $x \perp y$. Let W be a subspace of V. Then, a vector v is orthogonal to W, if it is orthogonal to every vector in W. By $w^\perp$ we denote the set of all vectors in V which are orthogonal to W.

A set of nonzero mutually orthogonal vectors $\{x_i\} \subset V$ is called **orthogonal set**. If additionally $x_i = 1$ for all vectors, it is called an **orthonormal set**. These notions transfer directly from finite to countable sets.

Theorem 2.10. *Let W be a closed subspace of a Hilbert space V. Then, $W^\perp$ is a closed subspace as well and every vector $v \in V$ has a unique decomposition*

$$v = w + u, \quad w \in W, u \in W^\perp. \tag{2.4}$$

eq:lafa:1

Proof. See for instance [Yos80, p. 82] □

Lemma 2.11 (Gram–Schmidt). *For every linearly independent sequence of vectors $\{v_i\}$ there is an up to scaling unique orthogonal set $\{x_i\}$ with the property that*

$$\forall n \in \mathbb{N}: \quad \operatorname*{span}_{i=1,\dots,n} \{x_i\} = \operatorname*{span}_{i=1,\dots,n} \{v_i\}.$$

Proof. □

Definition 2.12. A subset M of a Hilbert space V is called **dense**, if every vector in V is an accumulation point of M, that is, V is the closure of M. A Hilbert space is called **separable**, if it has a countable dense subset.

Note 2.13. From the point of view of numerical analysis and computation, spaces which are not separable are of limited interest. In fact, every result of a numerical calculation is in a finite set. When we look at convergence for $n \to \infty$ or $h \to 0$, we are usually studying sequences with countable index sets. Therefore, vectors in nonseparable spaces cannot be approximaetd reliably.

Theorem 2.14. *Every separable Hilbert space has an at most countable orthonormal basis.*

Proof. See e.g. [Yos80]. The proof is constructive and uses the Gram–Schmidt procedure. First, let M be a countable dense subset of V, which exists due to the separability assumption. Now choose any numbering of M and v_1 the first nonzero element in M. For $i = 2, \dots, \infty$ choose with $v_1, \dots, v_{i-1}$ given v_i as the next vector in M which is not in the subspace spanned by $v_1, \dots, v_{i-1}$. This procedure generates an at most countable sequence $\{v_i\}$ of linearly independent vectors. It will only stop, if V is finite dimensional, and we have that every element in M can be written as a finite linear combination of vectors v_i.

The sequence $\{v_i\}$ is a Schauder basis for V. In fact, given a vector $v \in V$ we have to show that for every ε, there is a finite linear combination $s_n = \sum_{i=1}^{n} \alpha_i v_i$ such that $v - s_n < \varepsilon$. Let by separability w_ε in M be such that $v - w_\varepsilon < \varepsilon$ and choose $s_n = w_\varepsilon$.

Finally, by the Gram–Schmidt procedure, we can construct an orthonormal set $\{x_i\}$ from the linearly independent set $\{v_i\}$. □

Example 2.15. In the Hilbert spaces $\mathbb{R}^n$, $\mathbb{C}^n$, $\ell^2(\mathbb{R})$, and $\ell^2(\mathbb{C})$, an orthonormal basis is obtained by choosing basis vectors x_i with entries $x_{i,j} = \delta_{ij}$.

2.2 Linear operators

Introduction 2.16. Linear mappings are the next central topic of linear algebra, which we want to extend to infinite dimensional spaces. Here, the basic definition remains the same, that is, a **linear operator** is a mapping of a Hilbert space V to a Hilbert space W which is compatible with vector operations. But Hilbert spaces have additional structure by their norms and their completeness.

Operators are well defined on a basis

Example 2.17. Let $\varphi : \ell^2(\mathbb{R}) \to \ell^2(\mathbb{R})$ be such that

$$v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ \vdots \end{pmatrix} \mapsto \varphi(v) = \begin{pmatrix} x_1 \\ 2x_2 \\ \vdots \\ kx_k \\ \vdots \end{pmatrix}.$$

Clearly, φ is linear. But if we consider the sequence of vectors $v_n = \{\delta_{nk}\}$, we see that $v_n \mapsto nv_n$ and thus, while the sequence is bounded in $\ell^2(\mathbb{R})$, its image is not.

Moreover, take now the sequence of vectors

$$v_n = \sum_{k=1}^{n} \frac{1}{k} \quad \mapsto \quad \varphi(v_n) = \sum_{k=1}^{n} 1.$$

The sequence v_n converges to a limit $v \in \ell^2(\mathbb{R})$, while the sequence $\varphi(v_n)$ diverges. While $\varphi(v_n)$ is defined for all v_n, it is not for the limit v.

Definition 2.18. A linear operator $\varphi : V \to W$ is **bounded**, if there is a constant $C > 0$ such that

$$\forall v \in V : \left\|\varphi(v)\right\|_W \leq C \left\|v\right\|_V.$$

Remark 2.19. By virtue of completeness of the space, whenever a linear operator is not bounded, it must be undefined for some vectors. We could exclude such operators from our considerations, but we would severely limit the theory we want to develop. Instead, we will accept the fact, that we have to extend the notion of a linear mapping $\varphi : V \to W$ to a linear operator $\varphi : V \to W$, which may not be defined on all of V. The following definition fixes this problem somewhat.

Definition 2.20. Let $\varphi : V \to W$ be a linear operator. Then, the **domain** of φ is

$$\mathcal{D}(\varphi) = \left\{v \in V \middle| \varphi(v) \in W\right\}.$$

Here, $\varphi(v) \in W$ implies that $\varphi(v)$ is also well defined.

Chapter 3

Iterative methods in finite and infinite dimensional spaces

Remark 3.1. This part of the notes deals with preconditioning of symmetric operators, or those, which have a dominating symmetric part. The theory of preconditioning methods for nonsymmetric and in particular non-normal operators is currently not well developed and thus will not be covered by these notes.

Notation 3.2. Iterative methods will be considered in a Hilbert space X with inner product $\langle ., . \rangle$.

Remark 3.3. For purposes of analysis we typically choose the space $X = L^2(\Omega)$. We admit a small inaccuracy here: when we run the algorithms on a computer, we usually employ the Eiclidean inner product, thus X should be the space of degrees of freedom. But this is a discrete space, where we cannot use theory of function spaces easily. Instead, we note that the L^2-inner product of standard finite element bases yield inner products equivalent to the Euclidean up to the local mesh size (see Lemma 3.13).

Example 3.4. While the methods developed in this chapter are fairly general, we introduce a specific model problem as a simple benchmark case. To this end, we consider the Dirichlet problem: find $u \in V = H_0^1(\Omega)$ such that

$$a(u, v) \equiv \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x = \int_\Omega f v \, \mathrm{d}x \equiv f(v), \qquad \forall v \in V. \tag{3.1}$$

eq:itintro:1

The finite dimensional linear systems of equations are derived from finite element discretizations on quasi-uniform meshes of cells with maximal diameter h, yielding a sequence of spaces V_h, on which linear systems are introduced by the same weak form (3.1).

Notation 3.5. With a bilinear form $a(.,.)$ on $X \times X$ we associate the operator $A : X \to X$ by

$$\langle Au, v\rangle = a(u, v), \quad \forall v \in V, \tag{3.2}$$ eq:itintro:2

where now $V = \mathcal{D}(A)$ is the **domain** of A, that is, the subset of functions $v \in X$, such that Av is defined and in X.

We will tacitly assume that operators A, B, etc. are defined by equation (3.2) and the bilinear forms $a(.,.)$, $b(.,.)$, etc., respectively, if they are not defined otherwise.

Definition 3.6. We call the bilinear form $a(.,.)$ and its associated operator A **symmetric**, if there holds

$$a(u, v) = a(v, u) \qquad \forall u, v \in V.$$

They are called V-**elliptic**, if for there is a positive number γ such that

$$a(u, u) \geq \gamma \|u\|_V^2 \qquad \forall u \in V.$$

Definition 3.7. For positive definite symmetric operators, we obtain the possibly infinite bounds of the spectrum

$$\Lambda(A) = \sup_{u\in V} \frac{a(u, u)}{\|u\|_X^2}, \qquad \lambda(A) = \inf_{u\in V} \frac{a(u, u)}{\|u\|_X^2}, \tag{3.3}$$ eq:richardson:8

as well as the possibly infinite **spectral condition number**

$$\kappa(A) = \frac{\Lambda(A)}{\lambda(A)}.$$

Note 3.8. Note that the spectral condition number depends on the norm of the space X. It is bounded, if and only if A is bounded with respect to this norm.

Example 3.9. Let $X = H_0^1(\Omega)$ with the inner product

$$\langle u, v\rangle_1 = \int_\Omega \nabla u \cdot \nabla v \, dx.$$

If A is the operator associated with the bilinear form $a(.,.)$ in (3.1), then

$$\Lambda(A) = \lambda(A) = \kappa(A) = 1.$$

If on the other hand $X = L^2(\Omega)$ equipped with the usual L^2-inner product, then A is unbounded and thus $\kappa(A) = \infty$. $\lambda(A)$ is the constant in Friedrichs' inequality.

Notation 3.10. After choosing a basis for a finite dimensional space X_n or a Schauder basis for the space X (assuming X separable), say $\{\varphi_i\}$, we can define a (possibly infinite-dimensional) matrix $\mathfrak{A}$ associated with the bilinear form $a(.,.)$ with the entries

$$a_{ij} = a(\varphi_j, \varphi_i).$$

If we restrict the bilinear forms to a finite dimensional subspace X_n, we denote the matrices $\mathfrak{A}$ restricted to this subspace by $\mathfrak{A}_n$.

Definition 3.11. The two extremal eigenvalues of the matrix $\mathfrak{A}_n$ can be obtained by the maximum and minimum of the **Rayleigh quotient**

$$\Lambda(\mathfrak{A}) = \max_{\mathfrak{x}\in\mathbb{R}^n} \frac{\mathfrak{x}^T\mathfrak{A}\mathfrak{x}}{\mathfrak{x}^T\mathfrak{x}}, \qquad \lambda(\mathfrak{A}) = \min_{\mathfrak{x}\in\mathbb{R}^n} \frac{\mathfrak{x}^T\mathfrak{A}\mathfrak{x}}{\mathfrak{x}^T\mathfrak{x}}. \tag{3.4}$$

eq:itintro:3

The spectral condition number is

$$\kappa_n(\mathfrak{A}) = \frac{\Lambda(\mathfrak{A})}{\lambda(\mathfrak{A})}.$$

Note 3.12. The spectral condition number of the operator A depends on the bilinear form $a(.,.)$ and the choice of the norm in X. On the other hand, the spectral condition number of the matrix $\mathfrak{A}$ depends on the choice of a basis of the space X_n.

lemma:itintro:1

Lemma 3.13. *Let $\{\varphi_i\}$ be the standard, piecewise linear, finite element basis on a quasi-uniform triangulation of mesh size h. Let $\mathfrak{M}$, the so called **mass matrix** be the matrix associated with the L^2-inner product with entries*

$$m_{ij} = \int_\Omega \varphi_i(x)\varphi_j(x)\,\mathrm{d}x.$$

Then,

$$\Lambda(\mathfrak{M}) \simeq h^d \simeq \lambda(\mathfrak{M})$$

Therefore, the condition number is

$$\kappa(\mathfrak{M}) = \frac{\mathcal{O}(h^d)}{\mathcal{O}(h^d)} = \mathcal{O}(1).$$

Proof. It is easy to verify, that $m_{ii} > 0$, and that not more entries in each row as edges of the triangulation meet in one vertex are different from zero. Furthermore, that the size of those entries is of order h^d, where d is the space dimension. From these two facts we immediately obtain

$$\Lambda(\mathfrak{M}) = \mathcal{O}(h^d).$$

The argument for $\lambda(\mathfrak{M})$ is more subtle. For any mesh cell T, let $\mathfrak{x}_T$ be the entries of the vector $\mathfrak{x}$ which belong to node values of the cell T. Let $\mathfrak{M}_T$ be the cell mass matrix obtained by restricting the L^2-inner product to T. Then,

$$\mathfrak{x}^T\mathfrak{M}\mathfrak{x} = \sum_{T\in\mathbb{T}_h} \mathfrak{x}_T^T\mathfrak{M}_T\mathfrak{x}_T \geq \min_{T\in\mathbb{T}_h} \frac{\mathfrak{x}_T^T\mathfrak{M}_T\mathfrak{x}_T}{\mathfrak{x}_T^T\mathfrak{x}_T} \sum_{T\in\mathbb{T}_h} |\mathfrak{x}_T|^2 \geq \lambda(\mathfrak{M}_T)|\mathfrak{x}|^2.$$

In order to estimate the eigenvalues of $\mathfrak{M}_T$, we note that for a unisolvent element, the norms $|\mathfrak{x}_T|$ and $u_{0,T}$ are equivalent on the reference cell, and the L^2-norm scales with h^d when transforming to the real cell T. Thus, we have $\lambda(\mathfrak{M}) = \mathcal{O}(h^d)$. □

Note 3.14. A closer inspection of the proof yields, that on a shape regular triangulation with cell radii ranging between the minimum h and the maximum H, we have

$$\begin{aligned}\Lambda(\mathfrak{M}) &= \mathcal{O}(H^d)\\ \lambda(\mathfrak{M}) &= \mathcal{O}(h^d)\\ \kappa(\mathfrak{M}) &= \mathcal{O}\left(\left(\frac{H}{h}\right)^d\right)\end{aligned}$$

This can be fixed by using weighted norms in $\mathbb{R}^n$.

3.1 The Richardson iteration

Introduction 3.15. As a first example and prototype for all other iterative methods we consider Richardson's method, which for matrices and vectors in $\mathbb{R}^n$ reads

$$\mathfrak{x}^{(k+1)} = \mathfrak{x}^{(k)} - \omega_k\left(\mathfrak{A}\mathfrak{x}^{(k)} - \mathfrak{b}\right). \tag{3.5}$$

eq:richardson:1

ω_k is a relaxation parameter, which can be chosen a priori or can be changed in every step. We will for simplicity assume $\omega_k = \omega$.

lemma:richardson:1

Lemma 3.16. *The error after one step of the Richardson method is given by*

$$\mathfrak{x}^{(k+1)} - x = \mathfrak{E}\left(\mathfrak{x}^{(k)} - x\right), \tag{3.6}$$

eq:richardson:14

where the error propagation operator is

$$\mathfrak{E} = \mathfrak{I} - \omega\mathfrak{A}. \tag{3.7}$$

eq:richardson:15

Proof. Using the fact that $\mathfrak{x} = \mathfrak{A}^{-1}\mathfrak{b}$, we write

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

□

theorem:richardson:1

Theorem 3.17. *If $\mathfrak{A}$ is symmetric, positive definite, with extremal eigenvalues $\lambda > 0$ and $\Lambda > 0$, then Richardson's method converges if and only if $0 < \omega < 2/\Lambda$. The optimal relaxation parameter is*

$$\omega_{opt} = \frac{2}{\lambda + \Lambda}, \tag{3.8}$$

eq:richardson:2

which yields an optimal contraction rate of

$$\varrho_{opt} = 1 - \frac{2\lambda}{\lambda + \Lambda} = \frac{\Lambda - \lambda}{\Lambda + \lambda} = \frac{\kappa - 1}{\kappa + 1} = 1 - \frac{2}{\kappa} + \mathcal{O}\left(\kappa^{-2}\right), \tag{3.9}$$

eq:richardson:4

*where $\kappa = \Lambda/\lambda$ is the so called **spectral condition number**.*

Proof. Convergence of this method is analyzed through the Banach fixed-point theorem, which requires contraction property of the matrix $\mathfrak{M} = \mathfrak{I} - \omega\mathfrak{A}$. Alternatively, we studied a theorem that states, that a matrix iteration converges if and only if the spectral radius

$$\varrho(\mathfrak{M}) = \max|\lambda(\mathfrak{M})| < 1,$$

the maximum absolute value of the eigenvalues of $\mathfrak{M}$ is strictly less than one.

If $\mathfrak{A}$ is symmetric, positive definite, with eigenvalues $\lambda_i > 0$, we have that

$$\varrho(\mathfrak{M}) = \max_i |1 - \omega\lambda_i| . \tag{3.10}$$ eq:richardson:13

Let the extremal eigenvalues be determined by the minimum and maximum of the Rayleigh quotient,

$$\lambda = \min_{x\in\mathbb{R}^n} \frac{\mathfrak{x}^T\mathfrak{A}\mathfrak{x}}{\mathfrak{x}^T\mathfrak{x}}, \qquad \text{and} \qquad \Lambda = \max_{x\in\mathbb{R}^n} \frac{\mathfrak{x}^T\mathfrak{A}\mathfrak{x}}{\mathfrak{x}^T\mathfrak{x}}. \tag{3.11}$$ eq:richardson:3

Then, equation (3.10) yields that the method converges for $0 < \omega < 2/\Lambda$. Furthermore, for $1/\Lambda \leq \omega \leq 2/\Lambda$ we have

$$\varrho(\mathfrak{M}) = \max\{-1 + \omega\Lambda, 1 - \omega\lambda\}.$$

The optimal parameter ω is the one where both values are equal and thus (3.8) and (3.9) hold. □

Introduction 3.18. The analysis of finite element methods shows that it is beneficial to give up the focus on finite dimensional spaces and rather use theory that applies to separable Hilbert spaces. If results can obtained in this context, they can easily be restricted to finite dimensional subspaces and thus become uniform with respect to the mesh parameter. Thus, we will first reformulate Richardson's method for this case and then derive convergence estimates.

Introduction 3.19. Elements of an abstract Hilbert space X will be denoted by u, v, w, etc. On the other hand, coefficient vectors in $\mathbb{R}^n$ are denoted by letters $\mathfrak{x}, \mathfrak{y}, \mathfrak{z}$, etc.

Definition 3.20. Let X be a Hilbert space with inner product $\langle ., . \rangle$. Let $a(., .)$ be a second bilinear form on X and the domain of its operator is V. Then, for any right hand side $f \in V$ and any start vector $u^{(0)} \in V$, **Richardson's method** is defined by the iteration

$$\left\langle u^{(k+1)}, v \right\rangle = \left\langle u^{(k)}, v \right\rangle - \omega_k \left(a(u^{(k)}, v) - \langle f, v \rangle \right), \qquad \forall v \in X. \tag{3.12}$$ eq:richardson:5

ω_k is a suitable relaxation parameter, chosen such that the method converges.

Note 3.21. The scalar products in (3.12) become necessary, since different from the case in $\mathbb{R}^n$, the result of applying the bilinear form $a(., .)$ to $u^{(k)}$ in the first argument yields a linear form on X. In order to convert this to a vector in X, we have to apply the isomorphism induced by the Riesz representation theorem.

theorem:richardson:2

Theorem 3.22. *Let the bilinear form $a(.,.)$ be bounded and elliptic on $X \times X$, namely, let there exist positive constants Λ and λ such that for all $u, v \in X$ there holds*

$$a(u,v) \leq \Lambda \|u\| \|v\|, \qquad a(u,u) \geq \lambda \|u\|^2. \tag{3.13}$$

eq:richardson:6

Then, Richardson's iteration converges for $\omega_k = \omega$ for any $\omega \in (0, 2\lambda/\Lambda^2)$.

Proof. We define the iteration operator T as the solution operator of equation (3.12), namely $Tu^{(k)} := u^{(k+1)}$. We have to prove that T is a contraction on X under the assumptions of the theorem.

For two arbitrary vectors $u^1, u^2 \in X$, let $w = u^1 - u^2$ be their difference. Due to linearity, we have $Tw = Tu^1 - Tu^2$ and

$$\langle Tw, v \rangle = \langle w, v \rangle - \omega a(w,v) = \langle w - \omega Aw, v \rangle .$$

Using $v = Tw$ as a test function, we obtain

$$\begin{aligned} \|Tw\|^2 &= \langle w - \omega Aw, w - \omega Aw \rangle \\ &= \|w\|^2 - 2\omega a(w,w) + \omega^2 \|Aw\|^2 \\ &\leq \|w\|^2 - 2\lambda\omega \|w\|^2 + \Lambda^2\omega^2 \|w\|^2 \\ &= \underbrace{\left(1 - 2\lambda\omega + \Lambda^2\omega^2\right)}_{=:\varrho(\omega)} \|w\|^2 . \end{aligned}$$

The function $\varrho(\omega)$ is a parabola open to the top, which at zero equals one and has a negative derivative. Thus, it is less than one for small positive valuers of ω. The other point where $\varrho(\omega) = 1$ is $\omega = 2\lambda/\Lambda^2$. □

Note 3.23. The condition on ω in Theorem 3.22 is more restrictive than in Theorem 3.17, since $\lambda/\Lambda \leq 1$. This is due to the fact, that in Theorem 3.17 we assume symmetry, and thus orthogonal diagonalizability of the matrix $\mathfrak{A}$. With similar assumptions, Theorem 3.22 could be made sharper.

Note 3.24. It is clear that the boundedness and ellipticity estimates (3.13) hold for any finite dimensional subspace $X_n \subset X$, and thus the convergence estimate (3.9) becomes independent of n.

More interesting and also more common is the case where the bilinear form $a(.,.)$ is unbounded on X. While it is still bounded on each finite subspace X_n, this bound cannot be independent of n if the sequence $\{X_n\}$ approximates X.

Note 3.25. We define an operator $B : X \to X^*$ such that $Bu = b(u,.) := \langle u, . \rangle$. By the Riesz representation theorem, there is a continuous inverse operator $B^{-1} : X^* \to X$, which is often called **Riesz isomorphism**.

definition:richardson:2

Definition 3.26. When we apply Richardson's method as in (3.12) on a computer, each step involves a multiplication with the matrix $\mathfrak{A}$, but an inversion of the matrix $\mathfrak{B}$, corresponding to the iteration

$$\mathfrak{B}\mathfrak{x}^{(k+1)} = \mathfrak{B}\mathfrak{x}^{(k)} - \omega_k\big(\mathfrak{A}\mathfrak{x}^{(k)} - \mathfrak{b}\big),$$

or equivalently,

$$\mathfrak{x}^{(k+1)} = \mathfrak{x}^{(k)} - \omega_k\mathfrak{B}^{-1}\big(\mathfrak{A}\mathfrak{x}^{(k)} - \mathfrak{b}\big). \tag{3.14}$$

eq:richardson:7

The iteration in (3.14) is commonly referred to as **preconditioned Richardson iteration** and $\mathfrak{B}^{-1}$ as the **preconditioner**. Note that by introducing the iteration in its weak form (3.12), the preconditioner arrives naturally and with necessity.

The goal of this chapter is finding preconditioners $\mathfrak{B}^{-1}$, or equivalently inner products $\langle .,. \rangle$, such that the bilinear form $a(.,.)$ is bounded and the condition number $\kappa = \Lambda/\lambda$ is small.

In order to reduce (or increase) confusion, we will refer to the inner product that we search in order to bound the condition number as $b(.,.)$ instead of $\langle .,. \rangle$, this way separating the Hilbert space X more clearly from the task of preconditioning. Thus, the operator B and the matrix $\mathfrak{B}$ will be associated with a bilinear form $b(.,.)$ and the final version of the preconditioned Richardson iteration is

$$b(u^{(k+1)}, v) = b(u^{(k)}, v) - \omega_k\big(a(u^{(k)}, v) - f(v)\big), \qquad \forall v \in X, \tag{3.15}$$

eq:richardson:10

or in operator form

$$u^{(k+1)} = u^{(k)} - \omega_k B^{-1}(Au^{(k)} - f). \tag{3.16}$$

eq:richardson:11

Remark 3.27. The space X and is inner product does not appear anymore in this formulation, since the bilinear form $b(.,.)$ has replaced it. Thus, finding a preconditioner also amounds to changing the space in which we iterate. This is reflected by the following:

Corollary 3.28. *Let the symmetric bilinear forms $a(.,.)$ and $b(.,.)$ in the Richardson iteration* (3.15) *be both bounded and positive definite on the same space V and fulfill the **spectral equivalence** relation*

$$\lambda b(u,u) \leq a(u,u) \leq \Lambda b(u,u), \quad \forall u \in V. \tag{3.17}$$

eq:richardson:12

Then, if $\omega_k \equiv \omega \in (0, 2\Lambda)$, the iteration is a contraction on V. The optimal contraction number is ϱ according to equation (3.9) *for ω chosen as in* (3.8).

Proof. This corollary is equivalent to Theorem 3.22 if the inner product $\langle .,. \rangle$ is replaced by the bilinear form $b(.,.)$. □

Remark 3.29. Originally, the space V was chosen as the domain of A which essentially meant $V \subset H^2(\Omega)$, since we required $Av \in X$. An additional benefit of the preconditioned version is, that now $V \subset H^1(\Omega)$ is sufficient and at least here no regularity assumption is required.

Notation 3.30. In order to distinguish different preconditioners, we will also us the notation $\lambda(B, A)$ and $\Lambda(B, A)$ to refer to the constants in the norm equivalence (3.17).

Example 3.31. Let us take the example (3.1). By the Poincaré-Friedrichs inequality, $a(.,.)$ is an inner product on X and thus we can choose $\langle .,. \rangle = a(.,.)$. In particular, $\lambda = \Lambda = 1$ and the optimal choice is $\omega = 1$. Then, Richardson's iteration becomes

$$a(u^{(k+1)}, v) = a(u^{(k)}, v) - \big(a(u^{(k)}, v) - f(v)\big) = f(v), \qquad \forall v \in X,$$

which converges in a single step, but we have to solve the original equation for u. Thus, either the inversion of the matrix A_n is trivial on each finite dimensional subspace X_n, or the method is useless. With usual finite element bases, the latter is true.

Example 3.32. In the other extreme, we would like to use the $\mathbb{R}^n$ or L^2 inner product on X_n or X, such that the Riesz isomorphism is easily computable. But then, the bilinear form $a(.,.)$ is unbounded on X. Thus, while for each finite n, the condition number $\kappa_n = \Lambda_n/\lambda_n$ exists, it converges to infinity if $n \to \infty$.

3.2 The conjugate gradient method

Distinguish between V and X.

Introduction 3.33. Relying on Hilbert space structure more than Richardson's iteration is the conjugate gradient method (cg), since it uses orthogonal search directions. Nevertheless, it also relies on constructing search directions from residuals, such that a Riesz isomorphism enters the same way as before and can then be used for preconditioning.

The beauty of the conjugate gradient method is, that it is parameter and tuning free, and it converges considerably faster than a linear iteration method.

Definition 3.34 (Conjugate gradient method)**.** Let V be a Hilbert space and V^* its dual. The **conjugate gradient method** for an iteration vector $u^{(k)} \in V$ involves the residuals $r^{(k)} \in V^*$ as well as the update direction $p^{(k)} \in V$ and the auxiliary vector $w^{(k)} \in V$. It consists of the steps

1. Initialization: for f and $u^{(0)}$ given, compute

$$\begin{aligned} r^{(0)} &= f - a(u^{(0)}, .) \\ \left\langle w^{(0)}, v \right\rangle &= r^{(0)}(v) \qquad\qquad \forall v \in V \\ p^{(0)} &= w^{(0)}. \end{aligned}$$

2. Iteration step: for $u^{(k)}$, $r^{(k)}$, $w^{(k)}$, and $p^{(k)}$ given, compute

$$\begin{aligned}
\alpha_k &= \frac{r^{(k)}(w^{(k)})}{a(p^{(k)}, p^{(k)})} \\
u^{(k+1)} &= u^{(k)} + \alpha_k p^{(k)} \\
r^{(k+1)} &= r^{(k)} - \alpha_k a(p^{(k)}, .) \\
\left\langle w^{(k+1)}, v \right\rangle &= r^{(k+1)}(v) \qquad\qquad \forall v \in V \\
\beta_k &= \frac{r^{(k+1)}(w^{(k+1)})}{r^{(k)}(w^{(k)})} \\
p^{(k+1)} &= w^{(k+1)} + \beta_k p^{(k)}
\end{aligned}$$

Remark 3.35. The results on orthogonality and minimization properties of the cg method in [GRS07] or [Saa00] remain valid in this context. Differences occur in the interpretation of these properties. The conjugate gradient method does not necessarily converge in a finite number of steps, and if the bilinear form is unbounded, no convergence rate is guaranteed.

Definition 3.36. The **preconditioned cg method** is obtained from above algorithm by reinterpreting the Riesz isomorphism in the computation of $w^{(k+1)}$ as a preconditioning operation, much alike Definition 3.26 of the preconditioned Richardson iteration. Thus, the line defining $w^{(k+1)}$ is replaced by

$$b(w^{(k+1)}, v) = r^{(k+1)}(v) \qquad\qquad \forall v \in V.$$

Here, like there, the preconditioner enters naturally from the weak form of the algorithm.

Definition 3.37. The nth **Krylov space** as subspace of the Hilbert space V with inner product $b(.,.)$ of the operator A and seed vector $w \in V$ is

$$\mathcal{K}_n = \mathcal{K}_n(B^{-1}A, w) = \operatorname{span}\left\{w, B^{-1}Aw, (B^{-1}A)^2 w, \ldots, (B^{-1}Aw)^{n-1}\right\}. \tag{3.18}$$ eq:cg:2

Lemma 3.38. *The iterates of the cg method have the following minimization properties:*

$$\begin{aligned}
\left\|u^{(k)} - u\right\|_A &= \min_{v \in \mathcal{K}_k} \left\|u^{(0)} + v - u\right\|_A \\
&= \min_{\substack{p \in P_{n-1} \\ p(0)=1}} \left\|u^{(0)} + p(B^{-1}A)w - u\right\|_A .
\end{aligned} \tag{3.19}$$ eq:cg:3

Theorem 3.39. *Let the bilinear form $a(.,.)$ be symmetric, and let the spectral equivalence* (3.17) *hold. Then, the preconditioned cg method converges and we have the estimate*

$$u^{(k)} - u_A \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k u^{(0)} - u_A. \tag{3.20}$$ eq:cg:1

Here, $\kappa = \Lambda/\lambda$ is the spectral condition number of the preconditioned problem.

Chapter 4

Schwarz methods

iteration:schwarz-methods

4.1 Additive Schwarz methods

Introduction 4.1. In this section, we study preconditioners, which are related to subspace decompositions of the space V or its finite dimensional subspaces. We will develop the theory in an abstract way, but always keep the model problem (3.1) in mind when we do so. In particular, the subspaces chosen will be associated with either coarser mesh levels or with meshes on subdomains of Ω.

This section follows in part [BS02, Chapter 7]. A more detailed discussion with extension of the methods developed here can be found in [TW05]

4.1.1 The abstract framework

Introduction 4.2. Let V be a Hilbert space with inner product $\langle ., . \rangle$ and let $a(., .) : V \times V$ be a symmetric and V-elliptic not necessarily bounded bilinear form. Let a set of auxiliary subspaces $\{V_j\}_{j=1,\dots,J}$ of V be chosen such that

$$V = \sum_{j=1}^{J} V_j.$$

The sum is not required to be direct, that is, a vector $v \in V$ may have several decompositions $v = \sum \alpha_j v_j$ with $v_j \in V_j$.

lemma:schwarz:1

Lemma 4.3. *Let the form $a(., .)$ be bounded and V_j-elliptic. Then, the weak formulation: find $u_j \in V_j$ such that*

$$a(u_j, v_j) = f(v_j), \quad \forall v_j \in V_j, \tag{4.1}$$

eq:schwarz:1

has a unique solution for all $f \in V^*$.

Proof. This is the Lax-Milgram lemma for V_j. □

definition:schwarz:1

Definition 4.4 (Ritz-projection)**.** Let the operator $P_j : V \to V_j$ be defined such that $P_j u \in V_j$ is the unique (Lemma 4.3) solution to the problem

$$a(P_j u, v_j) = a(u, v_j), \quad \forall v_j \in V_j. \tag{4.2}$$

eq:schwarz:2

We call P_j the A-orthogonal projection or **Ritz projection** to V_j. Since V_j is a subspace of V, we will also understand P_j as an endomorphism of V. The left hand side of this equation induces an operator $A_j : V_j \to V_j^*$ by $A_j u_j = a(u_j, .)$.

lemma:schwarz:ritz

Lemma 4.5. *The projections P_j as mappings from V to itself are self-adjoint with respect to the $a(.,.)$-inner product and positive semi-definite. Furthermore, P_j acts as identity on P_j and there holds $P_j^2 = P_j$.*

Proof. This is a well-known fact about orthogonal projections, which we will prove shortly. First, we note that by the uniqueness in Lemma 4.3 $P_j u_j = u_j$ for all $u_j \in V_j$. Thus, for all $u \in V$: $P_j P_j u = P_j u$. Let now $u, v \in V$ arbitrary. Then, there holds

$$a(u, P_j v) = a(P_j u, P_j v) = a(P_j v, P_j u) = a(v, P_j u) = a(P_j u, v).$$

Furthermore,

$$a(P_j u, u) = a(u, P_j u) = a(P_j u, P_j u) \geq 0,$$

since $a(.,.)$ is positive definite. □

definition:schwarz:1a

Definition 4.6. We define the orthogonal projection operator $\Pi_j : V \to V_j$ such that $\Pi_j u \in V_j$ is the solution to the problem

$$\left\langle \Pi_j u_j, v_j \right\rangle = \left\langle u, v_j \right\rangle, \quad \forall v_j \in V_j. \tag{4.3}$$

We define its dual $\Pi_j^T : V_j^* \to V^*$ by

$$\left\langle \Pi_j^T \varphi_j, v \right\rangle_{V^* \times V} = \left\langle \varphi_j, \Pi_j v \right\rangle_{V_j^* \times V_i} \tag{4.4}$$

eq:schwarz:3

Show that Π^T is an orthogonal projection, and onto which space.

lemma:schwarz:2

Lemma 4.7. *There holds*

$$A_j P_j = \Pi_j^T A. \tag{4.5}$$

eq:schwarz:15

Proof. Let $u, v \in V$ arbitrary. Let $v_j = \Pi_j v$. We rewrite equation (4.2) as

$$\left\langle A_j P_j u, v \right\rangle_{V^* \times V} = \left\langle A_j P_j u, v_j \right\rangle_{V^* \times V} = \left\langle A u, v_j \right\rangle_{V^* \times V} = \left\langle \Pi_j^T A u, v \right\rangle_{V^* \times V}.$$

□

Definition 4.8. The additive Schwarz preconditioner for the operator A associated with the symmetric, and V-elliptic bilinear form $a(.,.)$ with respect to the subspace decomposition V_j is the mapping $B : V \to V^*$ such that

$$B^{-1} = \sum_{j=1}^{J} P_j A^{-1}. \tag{4.6}$$ eq:schwarz:4

example:schwarz:Jacobi **Example 4.9.** The Jacobi method may serve as a guiding example for the definition of these methods. To this end, let $V = \mathbb{R}^n$ with its Euclidean inner product $\langle .,. \rangle$. let $V_j = \operatorname{span}\{e_j\}$ be the space spanned by the jth unit vector. Let A be a symmetric, positive definite matrix and $a(u,v) = v^T A u$. Then, equation (4.2) becomes

$$e_j^T A u_j = e_j^T A u \quad \Leftrightarrow \quad P_j u = u_j = \frac{1}{a_{jj}} (Au)_j. \tag{4.7}$$ eq:schwarz:27

Since for this decomposition, the sum $V = \bigoplus V_j$ is direct, we obtain with $D = \operatorname{diag}(a_{11}, \ldots, a_{nn})$ the matrix representation

$$(B^{-1} v)_j = \frac{1}{a_{jj}} (A A^{-1} v)_j = \frac{1}{a_{jj}} v_j \quad \Leftrightarrow \quad B^{-1} = D^{-1}.$$

We enter this preconditioner into the Richardson method in operator form (3.16) to obtain the iteration

$$\begin{aligned} u^{(k+1)} &= u^{(k)} - \omega_k \sum_{j=1}^{J} P_j \left(u^{(k)} - A^{-1} f \right) \\ &= u^{(k)} - \omega_k D^{-1} \left(A u^{(k)} - f \right). \end{aligned} \tag{4.8}$$ eq:schwarz:28

lemma:schwarz:3 **Lemma 4.10.** *If A is symmetric and positive definite, so is B^{-1} as defined in* (4.6).

Proof. By Lemma 4.7 and the fact that P_j maps into V_j, we have that

$$B^{-1} = \sum_{j=1}^{J} A_j^{-1} \Pi_j^T. \tag{4.9}$$ eq:schwarz:16

Due to equation (4.1), A_j inherits its symmetry and positive definiteness from A, and thus A_j^{-1} is s.p.d. Therefore, for each term in this sum and arbitrary elements $\varphi, \psi \in V^*$, we have

$$\left\langle A_j^{-1} \Pi_j^T \varphi, \psi \right\rangle_{V \times V^*} = \left\langle A_j^{-1} \Pi_j^T \varphi, \Pi_j^T \psi \right\rangle_{V \times V^*} = \left\langle \Pi_j^T \varphi, A_j^{-1} \Pi_j^T \psi \right\rangle = \left\langle \varphi, A_j^{-1} \Pi_j^T \psi \right\rangle.$$

The result now follows by linearity. □

lemma:schwarz:5 **Lemma 4.11.** *For $v \in V$ holds*

$$b(v,v) \equiv \langle Bv, v \rangle = \min_{v = \sum v_j} \sum_{j=1}^{J} a(v_j, v_j), \tag{4.10}$$ eq:schwarz:5

where the minimum is taken over all possible decompositions of v into a sum of elements $v_j \in V_j$ with $j = 1, \ldots, J$.

Proof. Since B^{-1} is s.p.d., so is B. Therefore, $\left\langle ., A_j^{-1}. \right\rangle$ is an inner product on V_j^*, for which the Bunyakovsky-Cauchy-Schwarz inequality holds. Thus, for an arbitrary decomposition $v = \sum v_j$ with $v_j \in V_j$, the computation

$$\begin{aligned}
b(v,v) &= \sum_{j=1}^{J} b(v, A_j^{-1} A_j v_j) = \sum_{j=1}^{J} \left\langle \Pi_j^T B v, A_j^{-1} A_j v_j \right\rangle \\
&\leq \sum_{j=1}^{J} \sqrt{\left\langle \Pi_j^T B v, A_j^{-1} \Pi_j^T B v \right\rangle} \sqrt{\left\langle A_j v_j, A_j^{-1} A_j v_j \right\rangle} \\
&\leq \sqrt{\sum_{j=1}^{J} \left\langle \Pi_j^T B v, A_j^{-1} \Pi_j^T B v \right\rangle} \sqrt{\sum_{j=1}^{J} \left\langle A_j v_j, A_j^{-1} A_j v_j \right\rangle} \\
&= \sqrt{\left\langle B v, \sum A_j^{-1} \Pi_j^T B v \right\rangle} \sqrt{\sum_{j=1}^{J} \left\langle A_j v_j, v_j \right\rangle} \\
&= \sqrt{b(v,v)} \sqrt{\sum_{j=1}^{J} a(v_j, v_j)},
\end{aligned}$$

yields for arbitrary decompositions

$$b(v,v) \leq \sum_{j=1}^{J} a(v_j, v_j), \tag{4.11}$$ eq:schwarz:17

and thus in particular, that the left hand side is bounded by the minimum of the right. Now we choose a special decomposition, showing that it cannot be less than the minimum. To this end, let

$$v_j = A_j^{-1} \Pi_j^T B v. \tag{4.12}$$ eq:schwarz:18

By Lemma 4.7, we have

$$\sum v_j = \sum A_j^{-1} \Pi_j^T B v = B^{-1} B v = v.$$

Furthermore,

$$\begin{aligned}
\sum_{j=1}^{J} \left\langle A_j v_j, v_j \right\rangle &= \sum_{j=1}^{J} \left\langle A_j A_j^{-1} \Pi_j^T B v, A_j^{-1} \Pi_j^T B v \right\rangle \\
&= \sum_{j=1}^{J} \left\langle \Pi_j^T B v, A_j^{-1} \Pi_j^T v \right\rangle \\
&= \left\langle B v, \sum A_j^{-1} \Pi_j^T B v \right\rangle = b(v,v).
\end{aligned}$$

□

theorem:schwarz:1

Theorem 4.12. *Let A be s.p.d. and B defined by equation* (4.6)*. Then, the spectral equivalence* (3.17) *holds with positive constants*

$$\Lambda(B,A) = \max_{v\in V} \frac{a(v,v)}{\min\limits_{v=\sum v_j} \sum\limits_{j=1}^{J} a(v_j,v_j)}, \qquad \lambda(B,A) = \min_{v\in V} \frac{a(v,v)}{\min\limits_{v=\sum v_j} \sum\limits_{j=1}^{J} a(v_j,v_j)}. \tag{4.13}$$

eq:schwarz:19

Proof. Here we use the fact, that $b(.,.)$ is an inner product on V and that by

$$b(B^{-1}Av,v) = a(v,v) = b(v,B^{-1}Av),$$

the operator $B^{-1}A$ is symmetric with respect to this inner product. Thus, the Rayleigh quotient qualifies to estimate the extremal eigenvalues, for instance,

$$\Lambda(B^{-1}A) = \max_{v\in V} \frac{b(B^{-1}Av,v)}{b(v,v)} = \max_{v\in V} \frac{a(v,v)}{\min\limits_{v=\sum v_j} \sum\limits_{j=1}^{J} a(v_j,v_j)},$$

and the same for the minimum. □

note:schwarz:1

Note 4.13. In order to estimate the condition number $\Lambda(B,A)/\lambda(B,A)$ of a Schwarz preconditioner, it is now sufficient to bound the two quotients in (4.13) from above and below. In particular, in order to find a bound for $\Lambda(B,A)$, we have to find an estimate of the form

$$a(v,v) \lesssim \min_{v=\sum v_j} \sum_{j=1}^{J} a(v_j,v_j), \tag{4.14}$$

eq:schwarz:23

or in other words, $a(v,v)$ has to be bounded by the sum on the right for any decomposition $v = \sum v_j$. On the other hand, in order to bound $1/\lambda(B,A)$, we need an estimate in the opposite direction, where it is sufficient to find one decomposition $v = \sum v_j$ such that it holds. We reduce these conditions to the following two abstract assumptions, which guarantee that Theorem 4.12 holds true.

hwarz:stable-decomposition

Assumption 4.14 (Stable decomposition). *For each $v \in V$ there is a decomposition*

$$v = \sum_{j=1^J} v_j, \qquad v_j \in V_j,$$

such that there holds

$$\min_{v=\sum v_j} \sum_{j=1}^{J} a(v_j,v_j) \lesssim a(v,v). \tag{4.15}$$

eq:schwarz:24

assumption:schwarz:1

Assumption 4.15 (Strengthened Cauchy-Schwarz inequalities)**.** *There is a symmetric $J \times J$-matrix $\mathcal{E}$ with entries $\varepsilon_{ij} \in [0,1]$ and a constant C independent of J, such that for the spectral radius $\varrho(\mathcal{E})$ there holds*

$$\varrho(\mathcal{E}) \leq C$$

and for all $1 \leq i, j \leq J$, $v_i \in V_i$ and $v_j \in V_j$ there holds

$$\left|a(v_i, v_j)\right| \leq \varepsilon_{ij} \sqrt{a(v_i, v_i)} \sqrt{a(v_j, v_j)}. \tag{4.16}$$

eq:schwarz:25

Note 4.16. Inequality (4.16) with $\varepsilon_{ij} \equiv 1$ holds by the regular Bunyakovsky-Cauchy-Schwarz inequality. But for such a matrix, the spectral radius is J. As the following lemma will reveal, it is necessary to obtain $\varrho(\mathcal{E})$ independent of J to obtain estimate (4.14).

lemma:schwarz:7

Lemma 4.17. *Let the estimate* (4.16) *hold. Then, estimate* (4.14) *holds with the constant $\varrho(\mathcal{E})$.*

Proof. Let $v \in V$ and its decomposition $v = \sum v_j$ with $v_j \in V_j$ be chosen arbitrarily. Then,

$$a(v,v) \;=\; a\left(\sum_i v_i, \sum_j v_j\right) \;=\; \sum_{i,j=1}^{J} a(v_i, v_j) \;\leq\; \sum_{i,j=1}^{J} \varepsilon_{ij} \sqrt{a(v_i, v_i)} \sqrt{a(v_j, v_j)}.$$

The latter sum corresponds to a matrix-vector product of the form $\mathfrak{x}^T \mathcal{E} \mathfrak{x}$, where the entries of x are of the form $\sqrt{a(v_i, v_i)}$. Since $\mathcal{E}$ is symmetric positive definite, this product can be estimated by $\varrho(\mathcal{E})|\mathfrak{x}|^2$, and thus

$$a(v,v) \leq \varrho(\mathcal{E}) \sum_{j=1}^{J} a(v_j, v_j). \tag{4.17}$$

eq:schwarz:38

□

4.2 Two-level additive Schwarz preconditioner

Introduction 4.18. This preconditioner is in the class of domain decomposition methods. The attribute two-level refers to the fact that we are considering finite element discretizations of (3.1) on two finite element meshes, the fine mesh $\mathbb{T}_h$ on which we desire to compute the solution, and the auxiliary coarse mesh $\mathbb{T}_H$. Both meshes cover the whole domain Ω (see Figure 4.1), and each cell of the coarse mesh is the union of cells of the fine mesh (4×4 fine cells in the figure).

In addition to these two meshes, we introduce subdomains $\Omega_1, \Omega_2, \ldots, \Omega_J$ of Ω such that each Ω_j is the union of cells in $\mathbb{T}_h$. We require that those subdomains overlap each other like the three examples in Figure 4.1 on the right. A more precise definition of the required overlap follows.

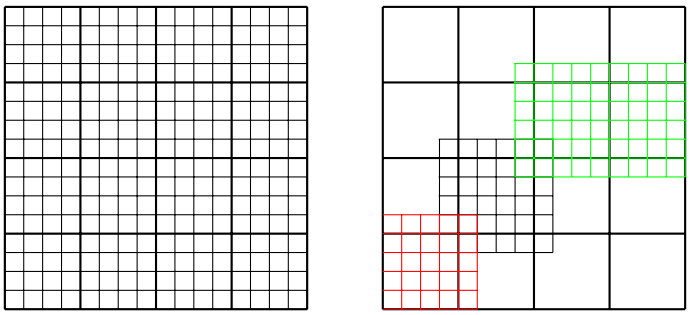

Figure 4.1: Fine mesh and coarse mesh (left) for overlapping domain decomposition. Examples for a subdomain decomposition on the right. fig:schwarz:ddmeshes

definition:schwarz:overlap

Definition 4.19. A covering of Ω with subdomains Ω_j is called **overlapping** with minimal **overlap** δ, if for each Ω_j and all $x \in \Omega_j$ holds:

$$\operatorname{dist}(x, \partial\Omega_J \setminus \partial\Omega) < \delta \;\Rightarrow\; \exists k \neq j : x \in \Omega_k.$$

on:schwarz:finite-covering

Definition 4.20. We say that a family of coverings is finite, if there is a constant N_O independent of $\mathbb{T}_h$ and the number of subdomains, such that for each j the intersection $\Omega_j \cap \Omega_k$ is nonempty for at most N_O subdomains Ω_k.

Definition 4.21. A smooth **partition of unity** with respect to the subdomains $\Omega_1, \Omega_2, \ldots, \Omega_J$ of Ω is a set of nonnegative functions $\{\varphi_1, \ldots, \varphi_J\} \subset C^\infty(\overline{\Omega})$ such that

$$\varphi_j(x) = 0 \qquad \forall x \in \Omega \setminus \Omega_j, \quad j = 1, \ldots, J \tag{4.18}$$ eq:schwarz:6

$$\sum_{j=1}^{J} \varphi_j(x) = 1 \qquad \forall x \in \overline{\Omega}. \tag{4.19}$$ eq:schwarz:7

Similarly, we can define partitions of unity in $H^1(\Omega)$ or partitions of unity which are piecewise C^1.

Furthermore, we assume that there is a positive constant δ, called **overlap**, such that for all $j = 1, \ldots, J$ there holds

$$\left\|\nabla \varphi j\right\|_{L^\infty(\Omega)} \lesssim \frac{1}{\delta}, \tag{4.20}$$ eq:schwarz:8

where the implicit constant is independent of h, δ and J.

Note 4.22. The term overlap for δ is justified by the following consideration. Let $x \in \Omega_j$ be a point which is not in any other Ω_i. Then, $\varphi_j(x) = 1$. If (4.20) is to hold, then it is necessary that $\operatorname{dist}(x, \partial\Omega_j) \geq \delta$ (up to a constant, but this constant is already in (4.20)). Thus, the points of distance less than δ from $\partial\Omega_j$ must be elements of another subdomain as well, which is then said to overlap with Ω_j.

example:schwarz:2

Example 4.23. On uniform meshes, overlapping subdomains with an overlap of $\delta = nh$ with $n = 1, 2, \ldots$ can be achieved easily by the following procedure:

1. Begin with a non overlapping subdivision $\{\Omega_j^0\}$, for instance aligned with the mesh cells of $\mathbb{T}_H$.
2. Add all cells that share at least a vertex with a cell in Ω_j^0 to obtain a domain Ω_j^1. After this procedure, two neighboring domains will overlap by two cells, resulting in $\delta = 2h$.
3. Repeat this procedure to obtain larger overlaps.

On the resulting partitions, a partition of unity in H^1 can be constructed with piecewise linear (bilinear on quadrilaterals) functions. For instance for Ω_j^1 this function is constructed as follows:

1. Choose $\varphi_j(x) = 1/2$ in all vertices on $\partial\Omega_j^0$.
2. Choose $\varphi_j(x) = 0$ in all vertices on $\partial\Omega_j^1$ and outside Ω_j^1.
3. Choose $\varphi_j(x) = 1$ in all remaining vertices inside Ω_j^0.
4. Connect these values by linear (on simplicial meshes), bilinear (quadrilateral meshes) of trilinear (hexahedral meshes) polynomials inside each mesh cell $T \in \mathbb{T}_h$.

This partition of unity achieves the estimate (4.20) with a constant of 1/2.

par:schwarz:1

Notation 4.24. The solution space of our problem is the space $V = V_h$ given by the finite element space on the mesh $\mathbb{T}_h$. We define finite element spaces on Ω_j by

$$V_j = \left\{ v \in V_h \middle| \forall x \in \Omega \setminus \Omega_j : v(x) = 0 \right\}. \tag{4.21}$$

eq:schwarz:9

Additionally, we define the space $V_0 \equiv V_H$ as the finite element space on the coarse mesh $\mathbb{T}_H$. Since the meshes are nested, V_H is indeed a subspace of V_h. Thus, we obtain a decomposition of V_h into $J + 1$ subspaces

$$V_h = V_0 + \sum_{j=1}^{J} V_j,$$

where the last J are associated with the subdomains. In fact, Lemma 4.27 below states that already the spaces V_1 to V_J are sufficient to span V_h. Nevertheless, the coarse grid space plays a crucial role in the efficiency of the method due to Lemma 4.30.

Definition 4.25. Let the spaces V_j, $j = 0, \dots, J$ be defined as above. Then, the **two-level additive Schwarz preconditioner** is defined as

$$B_{\mathrm{TLS}}^{-1} = \sum_{j=0}^{J} P_j A^{-1} = \sum_{j=0}^{J} A_j^{-1} \Pi_j^T, \tag{4.22}$$ eq:schwarz:10

where P_j is defined according to (4.2) and $A_j : V_j \to V_j^*$ by

$$\left\langle A_j u_j, v_j \right\rangle_V = a(u_j, v_j), \quad \forall u_j, v_j \in V_j. \tag{4.23}$$

Note 4.26. In order to simplify notation, we have assigned index zero to V_H. Thus, sums in future terms may either start at one, summing over subdomains, or at zero, summing over all subspaces.

lemma:schwarz:4 **Lemma 4.27.** *There holds*

$$V_h = \sum_{j=1}^{J} V_j. \tag{4.24}$$ eq:schwarz:11

Proof. Let $I_h : C(\overline{\Omega}) \to V_h$ be the interpolation operator of the finite element space. Then, for any given $v \in V_h$ define $v_j = I_h(\varphi_j v)$, where φ_j is the function associated to Ω_j of a partition of unity for $\Omega_1, \dots, \Omega_J$.

By definition of φ_j, there holds $\varphi_j v = 0$ on $\Omega \setminus \Omega_j$. Furthermore, we assumed that a mesh cell of $\mathbb{T}_h$ is either completely in Ω_j or completely in its complement. Since nodal values of a cell are located in the cell itself, this implies that $I_h(\varphi_j v) = 0$ on $\Omega \setminus \Omega_j$. Therefore, $I_h(\varphi_j v) \in V_j$.

On the other hand, we use the linearity of the interpolation operator to obtain

$$\sum_{j=1}^{J} v_j = \sum_{j=1}^{J} I_h(\varphi_j v) = I_h \left(v \sum_{j=1}^{J} \varphi_h \right) = I_h v = v,$$

thus, the v_j are indeed a decomposition of v. Since $v \in V_h$ was chosen arbitrarily, the lemma is proven. □

lemma:schwarz:8 **Lemma 4.28.** *Let the covering $\{\Omega_j\}_{j=1,\dots,J}$ for Ω be finite according to Definition 4.20. Then, the strengthened Cauchy-Schwarz inequalities* (4.16) *hold with a spectral radius*

$$\varrho(\mathcal{E}) \leq N_O. \tag{4.25}$$ eq:schwarz:26

Proof. The term $a(v_i, v_j)$ is nonzero only if the supports of the two functions have a nonempty intersection. Accordingly, for each index i only a maximum of N_O of the coefficients ε_{ij} are nonzero. We set these equal to one and use Gershgorin's theorem to estimate the greatest eigenvalue. □

lemma:schwarz:6

Lemma 4.29. *Let $v_j \in V_j$ for $j = 0, \dots, J$ be a composition of $v \in V_h$ such that*

$$v = \sum_{j=0}^{J} v_j.$$

Then,

$$a(v, v) \lesssim \sum_{j=0}^{J} a(v_j, v_j), \tag{4.26}$$

eq:schwarz:12

where the implicit constant does not depend on h, H, or J.

Proof. First a note: the inequality would be obvious, if V_h was a direct sum of the spaces V_j, and it would hold with a constant of one if they were mutually orthogonal. Thus, we have to show some kind of orthogonality between the spaces.

We start out by stating that

$$\begin{aligned} a(v, v) &= a\left(v_0 + \sum_{j=1}^{J} v_j, v_0 + \sum_{j=1}^{J} v_j\right) \\ &\leq 2\left(a(v_0, v_0) + a\left(\sum_{j=1}^{J} v_j, \sum_{j=1}^{J} v_j\right)\right) \\ &= 2\left(a(v_0, v_0) + \sum_{j,k=1}^{J} a(v_j, v_k)\right) \\ &\leq 2a(v_0, v_0) + 2N_O \sum_{j=0}^{J} a(v_j, v_j), \end{aligned}$$

where the last inequality is due to Lemma 4.17 and Lemma 4.28. Since N_O is assumed independent of h, H, and J, the lemma is proven. □

hwarz:stable-decomposition

Lemma 4.30 (Stable decomposition). *For each $v \in V_h$ there exists a decomposition $v = \sum_{j=0}^{J} v_j$ with $v_j \in V_j$, such that*

$$\sum_{j=0}^{J} a(v_j, v_j) \lesssim \left(1 + \frac{H}{\delta}\right)^2 a(v, v). \tag{4.27}$$

eq:schwarz:13

Proof. Let $\tilde{I}_H : H_0^1(\Omega) \to V_H$ be an interpolation operator continuous on $H^1(\Omega)$, for instance the interpolation operator by Clement or the one by Scott and Zhang. For a given function $v \in V_h$, let $v_H = \tilde{I}_H v$. Then

$$\begin{aligned} |v_H|_1 &\lesssim |v|_1 \\ v - v_{H0} &\lesssim H|v|_1. \end{aligned} \tag{4.28}$$

eq:schwarz:20

From (4.28) there holds

$$a(v_H, v_H) = |v_H|_1^2 \lesssim |v|_1^2 = a(v, v). \tag{4.29}$$ eq:schwarz:22

Let $w = v - v_H$. Using a partition of unity $\{\varphi_j\}$ for the subdomains $\{\Omega_j\}$ and the nodal interpolant I_h as in the proof of Lemma 4.27, and let

$$v_j = I_h(\varphi_j w).$$

Thus, $v = \sum v_j$. We point out, that we can use the nodal interpolant, since w is a finite element function on $\mathbb{T}_h$ and φ_j is either smooth or piecewise polynomial. For the remainder of this proof, we will assume the piecewise polynomial case (see Example 4.23) and leave the arguments for a smooth function φ_j to the reader.

The interpolation operator is exact for polynomials of degree k (assuming such an order for the finite element being used). Therefore,

$$a(v_j, v_j) = |v_j|_1^2 \lesssim |\varphi_j w|_1^2 \lesssim \nabla\varphi_j w_0^2 + \varphi_j \nabla w_0^2.$$

Using the properties of the partition of unity, we obtain

$$a(v_j, v_j) \lesssim \frac{1}{\delta^2}\chi(\Omega_j) w_0^2 + |\chi(\Omega_j) w|_1^2.$$

Summing up yields

$$\begin{aligned}
\sum_{j=1}^{J} a(v_j, v_j) &\lesssim \sum_{j=1}^{J} \left(\frac{1}{\delta^2}\chi(\Omega_j) w_0^2 + |\chi(\Omega_j) w|_1^2 \right) \\
&\leq N_O \left(\frac{1}{\delta^2} w_0^2 + |w|_1^2 \right) \\
&= N_O \left(\frac{1}{\delta^2} v - v_{H0}^2 + |v - v_H|_1^2 \right) \qquad (4.30) \quad \texttt{eq:schwarz:21} \\
&\lesssim \frac{H^2}{\delta^2} |v|_1^2 + |v|_1^2 \\
&= \left(1 + \frac{H^2}{\delta^2} \right) a(v, v).
\end{aligned}$$

The estimate (4.27) now follows from (4.29) and (4.30). □

warz:two-level-convergence **Theorem 4.31.** *Under the assumptions made so far in this section, there holds*

$$\kappa(B_{TLS}^{-1} A_h) = \frac{\Lambda(B_{TLS}, A_h)}{\lambda(B_{TLS}, A_h)} \lesssim \left(1 + \frac{H}{\delta} \right)^2, \tag{4.31}$$ eq:schwarz:14

where the implicit constant is independent of h, δ, H, and J.

Proof. The proof follows Note 4.13. Indeed, Lemma 4.29 proves inequality (4.14) and Lemma 4.30 proves (4.15). □

Remark 4.32. We have constructed a preconditioner B_{TLS} for the finite element discretization of the Poisson problem (3.1) such that the preconditioned system has a bounded condition number independent of the mesh size h. Thus, a Richardson iteration or a conjugate gradient method using this preconditioner will reduce the error by a certain amount within a fixed number of steps.

Closer inspection of the estimate (4.31) in view of Example 4.23 reveals a problem though: typically, δ is of the order of h, such that the estimate becomes

$$\kappa(B_{\mathrm{TLS}}^{-1}A_h) \lesssim \left(1 + \frac{H}{h}\right)^2.$$

If we choose H constant, this is exactly as bad as the condition number of A itself. Under mild further assumptions, the square on the right hand side can be avoided [DW94], which is an improvement compared to the operator without preconditioning, but is not uniform with respect to h. Therefore, we are left with two options:

1. Increase the overlap such that it is $\mathcal{O}(H)$. This procedure yields a uniform preconditioner, but it introduces a problem: when refining h, more and more cells belong to several subdomains and thus computations on them have to be performed several times. Therefore, the effort per preconditioning step is increased considerably.
2. Keep the overlap at a small multiple of h and choose H such that H/h is bounded by a constant. Then, the preconditioner remains uniform, the overlap remains small. This way, the difficulty has been transferred to the coarse grid problem on $\mathbb{T}_H$, since now this problem becomes more and more difficult to solve, when h decreases.

While the problems of the first option above are inherent and unavoidable, the second option is at least seemingly optimal and more creativity may be invested into the solution of the coarse grid problem. Therefore, the latter is usually preferred. The coarse grid problem then leads to the idea of multigrid methods, which will be dealt with in Chapter 5.

4.3 Multiplicative Schwarz methods

Example 4.33. If we write the Jacobi method in equation (4.8) for each line, we get for each $j = 1, \dots, J$:

$$u_j^{(k+1)} = u_j^{(k)} - \omega_k P_j\left(u^{(k)} - A^{-1}f\right). \tag{4.32}$$ eq:schwarz:29

Since the updates are orthogonal, this is equivalent to the sum in (4.8). In the Gauß-Seidel method, these projections are done consecutively. For instance, if we introduce broken indices, we can write it in the form

$$u_j^{(k+\frac{j}{J})} = u_j^{(k+\frac{j-1}{J})} - \omega_k P_j \left(u^{(k+\frac{j-1}{J})} - A^{-1} f \right). \tag{4.33}$$ eq:schwarz:30

This means, we apply the corrections consecutively one after the other. In order to understand convergence of this method, we are bringing it into a different form and study the propagation of the error. Let $uA^{-1}f$ be the solution of the problem. Then, the error propagates like

$$\begin{aligned} u^{(k+\frac{j}{J})} - u = u^{(k+\frac{j-1}{J})} - u - \omega_k P_j \left(u^{(k+\frac{j-1}{J})} - u \right) \\ = (I - \omega_k P_j) \left(u^{(k+\frac{j-1}{J})} - u \right). \end{aligned} \tag{4.34}$$ eq:schwarz:31

The error after a whole Gauß-Seidel step is

$$u^{(k+1)} - u = (I - \omega_k P_J)(I - \omega_k P_{J-1}) \dots (I - \omega_k P_1) \left(u^{(k)} - u \right). \tag{4.35}$$ eq:schwarz:32

The corresponding error equation for the Jacobi method is

$$u^{(k+1)} - u = \left(I - \omega_k \sum_{j=1}^{J} P_j \right) \left(u^{(k)} - u \right). \tag{4.36}$$ eq:schwarz:33

Thus the notions of multiplicative and additive methods.

Introduction 4.34. The example of the Jacobi and Gauß-Seidel methods is generic in the way that for any subspace decomposition of V into the sum of V_j, we can define an additive and a multiplicative method. Not surprisingly, their analysis also rests on the same ingredients.

Definition 4.35. The **multiplicative Schwarz preconditioner** for the operator A associated with the symmetric, and V-elliptic bilinear form $a(.,.)$ with respect to the subspace decomposition V_j is the mapping $B_m : V \to V^*$ such that

$$B_m^{-1} = (I - E_J) A^{-1}, \tag{4.37}$$ eq:schwarz:34

where E_J is the multiplicative **error propagation operator**

$$E_J = (I - P_J)(I - P_{J-1}) \dots (I - P_1). \tag{4.38}$$ eq:schwarz:35

Lemma 4.36. *The error after one step of the iteration*

$$u^{(k+1)} = u^{(k)} - B_m^{-1} \left(A u^{(k)} - f \right)$$

is given by

$$u^{(k+1)} - u = E_J \left(u^{(k)} - u \right).$$

Proof. First, we use the definition of B_m to obtain

$$u^{(k+1)} = u^{(k)} - (I - E_J)\left(u^{(k)} - u\right).$$

Therefore,

$$u^{(k+1)} - u = u^{(k)} - u - (I - E_J)\left(u^{(k)} - u\right) = (I - I + E_J)\left(u^{(k)} - u\right).$$

□

Remark 4.37. We can define the error propagation operator E_j through the recursion

$$E_0 = I, \qquad E_k = (I - P_k)E_{k-1}, \quad k = 1, \dots, J. \tag{4.39}$$ eq:schwarz:45

lemma:schwarz:9 **Lemma 4.38.** *The error propagation operator E_j has the following porperties:*

$$E_{j-1}^* E_{j-1} - E_j^* E_j = E_{j-1}^* P_j E_{j-1} \tag{4.40}$$ eq:schwarz:44

$$I - E_j = \sum_{k=1}^{j} P_k E_{k-1}. \tag{4.41}$$ eq:schwarz:46

Here, E^* is the $a(.,.)$-adjoint of E.

Proof. In order to prove the first identity, we use the recursion formula (4.39) to obtain (using $P_j = P_j^*$ and $P_j^2 = P_j$)

$$\begin{aligned} E_j^* E_j &= E_{j-1}^*(I - P_j^*)(I - P_j)E_{j-1} \\ &= E_{j-1}^* E_{j-1} - E_{j-1}^* P_j E_{j-1}. \end{aligned}$$

The second identity is proven by induction with

$$(I - E_0) = I - I = 0,$$

$$(I - E_j) = I - (I - P_j)E_{j-1} = P_j E_{j-1} + I - E_{j-1} = P_j E_{j-1} + \sum_{k=1}^{j-1} P_k E_{k-1}.$$

□

lemma:schwarz:10 **Lemma 4.39.** *Let Assumption 4.15 (strengthened Cauchy-Schwarz inequalities) be satisfied. Then the following inequalities hold for $0 \le j, k \le J$ and for $u, v \in V$:*

$$a(P_j u, v) \le \sqrt{a(P_j u, u)}\sqrt{a(P_j v, v)}, \tag{4.42}$$

$$a(P_j u, P_k v) \le \varepsilon_{jk}\sqrt{a(P_j u, u)}\sqrt{a(P_k v, v)}. \tag{4.43}$$

Proof. By the definition of P_j, the fact that P_j is $a(.,.)$-self adjoint (Lemma 4.5) and the Bunyakovsky-Cauchy-Schwarz inequality, we have

$$\begin{aligned} a(P_j u, v) = a(u, P_j v) &= a(P_j u, P_j v) \\ &\leq \sqrt{a(P_j u, P_j u)}\sqrt{a(P_j v, P_j v)} \leq \sqrt{a(P_j u, u)}\sqrt{a(P_j v, v)}. \end{aligned}$$

The second inequality follows readily by

$$a(P_j u, P_k v) \leq \varepsilon_{jk}\sqrt{a(P_j u, P_j u)}\sqrt{a(P_j v, P_j v)} = \varepsilon_{jk}\sqrt{a(P_j u, u)}\sqrt{a(P_j v, v)}.$$

□

Theorem 4.40. *Let Assumptions 4.15 and the estimate* (4.14) *be satisfied. Then, the error propagation operator satisfies*

$$E_{JA}^2 \leq 1 - \frac{1}{\varrho(\mathcal{E})C_{4.15}} < 1, \tag{4.44}$$

eq:schwarz:47

Proof. We use equation (4.40) of Lemma 4.38 to obtain

$$I - E_J^* E_J = \sum_{j=1}^{j} \left(E_{j-1}^* E_{j-1} - E_j^* E_j\right) = \sum_{j=1}^{J} E_{j-1}^* P_j E_{j-1}.$$

Since the operators P_j are positive semi-definite, by rearranging we obtain for all $v \in V$ the estimate

$$a(E_J v, E_J v) \leq a(v, v) - \sum_{j=1}^{J} a(E_{j-1} v, P_j E_{j-1} v). \tag{4.45}$$

eq:schwarz:49

Thus, $E_{JA}^2 \leq 1$, but we have to show that the sum on the right is sufficiently positive to get an estimate of the convergence rate. Therefore, we start with equation (4.41) of Lemma 4.38, yielding by Lemma 4.39

$$\begin{aligned} a(P_j v, v) &= a(P_j v, E_{j-1} v) + \sum_{k=1}^{j-1} a(P_j v, P_k E_{k-1} v) \\ &\leq \sqrt{a(P_j v, v)} \left(\sqrt{a(P_j E_{j-1} v, E_{j-1} v)} + \sum_{k=1}^{j-1} \varepsilon_{jk} \sqrt{a(P_k E_{k-1} v, E_{k-1} v)} \right) \\ &\leq \sqrt{a(P_j v, v)} \left(\sum_{k=1}^{j} \varepsilon_{jk} \sqrt{a(P_k E_{k-1} v, E_{k-1} v)} \right) \end{aligned}$$

Now let $z \in \mathbb{R}^J$ be the vector with entries $z_k = \sqrt{a(P_k E_{k-1} v, E_{k-1} v)}$. Then, we rewrite the previous estimate as

$$a(P_j v, v) \leq (\mathcal{E} z)_j^2.$$

Summing over j yields

$$\sum_{j=1}^{J} a(P_j v, v) \le \mathcal{E} z^2 \le \varrho(\mathcal{E})^2 z^2 = \varrho(\mathcal{E})^2 \sum_{k=1}^{j} a(P_k E_{k-1} v, E_{k-1} v). \tag{4.46}$$ eq:schwarz:48

At this point, we use Lemma 4.41 to estimate

$$\frac{1}{C_{4.15}^2} a(v,v) \le \sum_{j=1}^{J} a(P_j v, v) \le \varrho(\mathcal{E})^2 \sum_{k=1}^{j} a(P_k E_{k-1} v, E_{k-1} v)$$

Finally, entering this estimate into (4.45), we obtain

$$a(E_J v, E_J v) \le a(v,v) \left(1 - \frac{1}{\varrho(\mathcal{E})^2 C_{4.15}^2}\right),$$

which is the statement of the theorem. □

lemma:schwarz:11 **Lemma 4.41.** *Let inequality* (4.15) *hold for some stable decomposition* $v = \sum v_j$ *with the constant* $C_{4.15}$*. Then,*

$$\sum_{j=1}^{J} a(P_j v, v) \ge \frac{1}{C_{4.15}^2} a(v,v). \tag{4.47}$$ eq:schwarz:50

Proof. From inequality (4.15) for the decomposition $v = \sum v_j$, and the definition of P_j, and the Bunyakovsky-Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} a(v,v) &= \sum_{j=1}^{J} a(v, v_j) \\ &= \sum_{j=1}^{J} a(P_j v, v_j) \\ &\le \sqrt{\sum_{j=1}^{J} a(P_j v, P_j v)} \sqrt{\sum_{j=1}^{J} a(v_j, v_j)} \\ &\le \sqrt{\sum_{j=1}^{J} a(P_j v, P_j v)}\ C_{4.15} \sqrt{a(v,v)}. \end{aligned}$$

Thus,

$$a(v,v) \le C_{4.15}^2 \sum_{j=1}^{J} a(P_j v, P_j v) = C_{4.15}^2 \sum_{j=1}^{J} a(v, P_j v).$$

This proves the lemma. □

4.4 Extensions

Introduction 4.42. In order to keep the presentation and the proofs simple, we have presented Schwarz methods in their most straight forward form. The framework allows for extensions, which each add minor complications to the proofs and yield slightly modified results. We are going to mention a few of them.

Remark 4.43. The definition of the additive Schwarz operator in equation (4.6) and the multiplicative Schwarz operator (4.37) rely on the Ritz projections P_j, which in turn, through their definition in (4.2) require solving the local problems with the operators A_j exactly.

In some cases, it might be advantageous either to solve a different local problem, or to solve the local problem approximately. In both cases, we can rewrite the algorithm as using a different local projection $\widetilde{P}_j$ which instead of the Ritz projection (4.2) is defined by the equation

$$\widetilde{a}_j(\widetilde{P}_j u_j, v_j) = a(u, v_j), \quad \forall v_j \in V_j, \tag{4.48}$$ eq:schwarz:42

with a corresponding operator $\widetilde{A}_j : V_j \to V_j^*$. We will continue to assume that $\widetilde{a}_j(.,.)$ is symmetric and elliptic in the same way as $a(.,.)$ is, but possibly with different constants.

It is obvious, that we will need assumptions on the modified bilinear forms $\widetilde{a}_j(.,.)$and their relationship with $a(.,.)$. But it turns out, that if we make these additional assumptions, we can make the replacements of P_j by $\widetilde{P}_j$ in the algorithm and the analysis carries through with just one additional parameter involved.

The modifications in the analysis are as follows: first, replace the stable decomposition lemma 4.30 by assumption 4.44. Then, introduce the additional assumption 4.45, which is ellipticity of the modified forms in V_j with respect to the norm established by the original bilinear form. Both assumptions together establish a relaxed form (only for the sum over j) of spectral equivalence for $\widetilde{a}_j(.,.)$ and $a(.,.)$ on V_j. The strengthened Cauchy-Schwarz inequalities in Assumption 4.15 remain the same.

rz:stable-decomposition-2

Assumption 4.44 (Stable decomposition). *For each $v \in V_h$ there exists a decomposition $v = \sum_{j=0}^{J} v_j$ with $v_j \in V_j$, such that*

$$\sum_{j=0}^{J} \widetilde{a}_j(v_j, v_j) \lesssim a(v, v), \tag{4.49}$$ eq:schwarz:36

where the implicit constant is independent of the number of subdomains J.

on:schwarz:local-stability

Assumption 4.45 (Local stability). *There is a constant $\omega > 0$ such that*

$$a(v_j, v_j) \leq \omega \widetilde{a}_j(v_j, v_j) \quad \forall j = 1, \ldots, J \; \forall v_j \in V_j. \tag{4.50}$$ eq:schwarz:37

Remark 4.46. With these two assumptions, the convergence estimates change as follows: first, using the strengthened Cauchy-Schwarz inequalities (4.16), we extend (4.17) to

$$a(v,v) \leq \omega\varrho(\mathcal{E}) \sum_{j=1}^{J} \widetilde{a}_j(v_j, v_j). \tag{4.51}$$ eq:schwarz:39

Then, the proof of Theorem 4.12 can be conducted in the very same way as before, using (4.49) and (4.51).

Example 4.47. A simple example is the introduction of a relaxation parameter ω such that

$$\widetilde{a}_j(v_j, v_j) = \frac{1}{\omega} a(v_j, v_j).$$

Then, obviously, the local stability (4.50) holds. Furthermore, the constant in the stable decomposition estimate changes by a factor $1/\omega$. Thus, in this case, the upper and lower bounds $\Lambda(B,A)$ and $\lambda(B,A)$ change by the same factor and the condition number stays the same.

Remark 4.48. The second extension is, that we can replace the subspaces V_j by auxiliary spaces X_j, which are not subspaces of V. In such a situation, we have to require the existence of a prolongation or extension operator $R_j^T : X_j \to V$. This situation is adapted most easily to our existing framework by introducing subspaces V_j as the range of R_j^T, namely,

$$V_j = \left\{ R_j^T x \in V \middle| x \in X_j \right\}. \tag{4.52}$$ eq:schwarz:40

Then, the local forms are defined on the auxiliary spaces,

$$\widetilde{a}_j(.,.) : X_j \times X_j \to \mathbb{R}, \quad j = 1, \ldots, J, \tag{4.53}$$ eq:schwarz:41

and wherever we need a vector in V_j, we use the interpolation operator. Thus, we introduce decompositions of the form $v = \sum \alpha_j R_j^T x_j$ and the operator $\widetilde{P}_j : V \to X_j$ is defined by

$$\widetilde{a}_j(\widetilde{P}_j u, y_j) = a(u, R_j^T y_j), \quad \forall y_j \in X_j. \tag{4.54}$$ eq:schwarz:43

These modifications introduce new operators, but they do not affect the analysis.

Example 4.49. When we implement a block-Jacobi method, we have $V = \mathbb{R}^n$ and m-dimensional subspaces with $J = n/m$ (we assume the quotient is integer). In the standard formulation with subspaces V_j, the operators A_j, which have to be inverted, are $n \times n$-matrices with only m rows and columns different from zero. This is not a useful description of the local problems. Instead, we want an invertible

matrix $A_j \in \mathbb{R}^{m\times m}$. This can be achieved by choosing the extension operator

$$R_j^T : \mathbb{R}^m \to \mathbb{R}_n, \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ \vdots \\ 0 \\ v_{mj+1} = x_1 \\ \vdots \\ v_{m(j+1)} = x_m \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{4.55}$$

Example 4.50. Other examples are finite element methods with nonnested spaces on different levels, for instance, when the meshes of a hierarchy are not nested.

Remark 4.51. We have discussed additive and multiplicative Schwarz methods, but we are not forced to consider only methods organized strictly in one way or the other. Instead, some of the operations can be performed in the additive, some in the multiplicative way.

This is advantageous for instance on multicore hardware: an inspection of the algorithms shows, that application of all operators P_j can be implemented in parallel, while the applications of the operators $(I - P_j)$ in the multiplicative method is sequential.

Example 4.52. Let the indices $j = 1, \ldots, J$ be grouped into M subsets I_m, such that

$$P_i P_k = P_k P_i = 0, \qquad \forall i, k \in I_m.$$

Such a distribution of indices is also called **coloring**. Then,

$$(I - P_i)(I - P_k) = I - (P_i + P_k),$$

and by application to the whole subset and all subsets, the error propagation operator of the multiplicative method can be rewritten as

$$E_J = \left(I - \sum_{j\in I_1} P_j\right) \ldots \left(I - \sum_{j\in I_m} P_j\right).$$

While the operations inside each "color" are arranged in an additive and thus parallelizable manner, the colers themselves are arranged in a multiplicative way.

Example 4.53. Another example is a rearrangement of the two-level Schwarz method in a way, that the domain decomposition subspaces are still dealt with in

an additive fashion, while the coarse space is connected multiplicatively. An example for a symmetric version of this method is described by the error propagation operator

$$E_{TL} = \left(I - \sum_{j=1}^{J} P_j \right) (I - P_0) \left(I - \sum_{j=1}^{J} P_j \right), \tag{4.56}$$

eq:schwarz:51

which is the two-level operator with Schwarz smoother discussed in the next chapter.

Chapter 5

Multigrid methods

teration:multigrid-methods

Introduction 5.1. multigrid Multigrid methods avoid the problems discussed in the section on two-level Schwarz methods by using not only two, but a whole hierarchy of mesh levels. On each level, an approximate solver, a so called smoother is employed, which improves the error somewhat, and then an approximation on a coarser level is used to improve further. This is done down to the coarsest level, where we assume that the solution process is cheap.

Definition 5.2. A **hierarchy of spaces** $\{V_\ell\}_0 \leq \ell \leq L$ is a sequence of the form

$$V_0 \subset V_1 \subset \cdots \subset V_L. \tag{5.1}$$

We assume that $V_L = V$ is the high resolution space on which we want to solve (3.1), but where the condition number of the matrix $\mathfrak{A}$ is bad. On the other end of the spectrum, we assume that the solution of (3.1) on V_0 is easily possible.

Definition 5.3. A **multigrid method** consists of the following components:

1. A **smoother** R_ℓ acting on the level space V_ℓ, usually an iterative method like Richardson, Jacobi, Gauß-Seidel or a Schwarz method.
2. A **coarse grid solver** solving the problem on V_0 exactly.
3. Transfer operators between the levels V_ℓ and $V_{\ell+1}$. For standard finite element methods, this is typically the embedding operator. The transfer in opposite direction is achieved by the L^2-projection.

On a given level V_ℓ, the multigrid level consists of an alternating sequence of smoothing steps and coarse grid corrections, where the latter consist of a projection of the residual to the space $V_{\ell-1}$ and then recursive application of the same sequence. This is easiest described by the function

eq:mg:5

$$u_1 = MG_\ell(u^{(0)}, g), \tag{5.2a}$$

eq:mg:1

which takes an initial value $u^{(0)}$ and computes an approximation u_1 to the solution to $A_\ell u = g$ by the following scheme: first, for $\ell = 0$ let

$$MG_0(u^{(0)}, g) = A_0^{-1} g_0.$$

On levels $\ell \neq 0$, perform the three steps

- Pre-smoothing: apply m_{pre} steps of a Richardson iteration preconditioned with the smoother R_ℓ:

$$u^{(k+1)} = u^{(k)} - R_\ell^{-1} \left(A_\ell u^{(k)} - g_\ell \right), \qquad 0 \leq k < m_{\text{pre}}. \tag{5.2b}$$

eq:mg:2

- Coarse grid correction: let $v^{(0)} \in V_{\ell-1}$ and $g_{\ell-1} \in V_{\ell-1}^*$ such that

$$g_{\ell-1} = \Pi_{\ell-1}^T \left(g_\ell - A_\ell u^{(m_{\text{pre}})} \right), \qquad v^{(0)} = 0. \tag{5.2c}$$

eq:mg:3

  Then, compute

$$v^{(k+1)} = MG_{\ell-1}(v^{(k)}, g_{\ell-1}), \qquad 0 \leq k < m_{\text{coarse}}. \tag{5.2d}$$

eq:mg:4

  Let $w^{(0)} \in V_\ell$ be given by $w^{(0)} = u^{(m_{\text{pre}})} + v^{(m_{\text{coarse}})}$.

- Post-smoothing: apply m_{post} steps of a Richardson iteration preconditioned with the smoother R_ℓ:

$$w^{(k+1)} = w^{(k)} - R_\ell^{-1} \left(A_\ell w^{(k)} - g_\ell \right), \qquad 0 \leq k < m_{\text{post}}. \tag{5.2e}$$

eq:mg:2a

  Assign $MG(u^{(0)}, g_\ell) = w^{(m_{\text{post}})}$.

This method has three parameters, the numbers of pre- and post smoothing steps m_{pre} and m_{post} as well as the number of coarse grid iterations m_{coarse}. Here, it is the last one which has a strong impact on the structure of the iteration. It defines what is called the **cycle type**, which is either **V-cycle** for $m_{\text{coarse}} = 1$ or **W-cycle** for $m_{\text{coarse}} = 2$. The structure of the cycles can be seen in Figure 5.1.

Remark 5.4. Figure 5.1 shows that the recursive structure of the W-cycle is much more complex than that of the V-cycle. The complexity analysis below will show that higher values of m_{coarse} do not lead to efficient algorithms.

Definition 5.5. If the numbers of pre- and post smoothing steps in the V-cycle are dependent on the level ℓ, we speak of the **variable V-cycle**. A typical choice is $m_\ell = 2^{L-\ell} m_L$, thus doubling the number of smoothing steps whenever stepping down one level.

Note 5.6. The variable V-cycle with m_ℓ as mentioned in the previous definition has as many smoothing steps per iteration as the W-cycle.

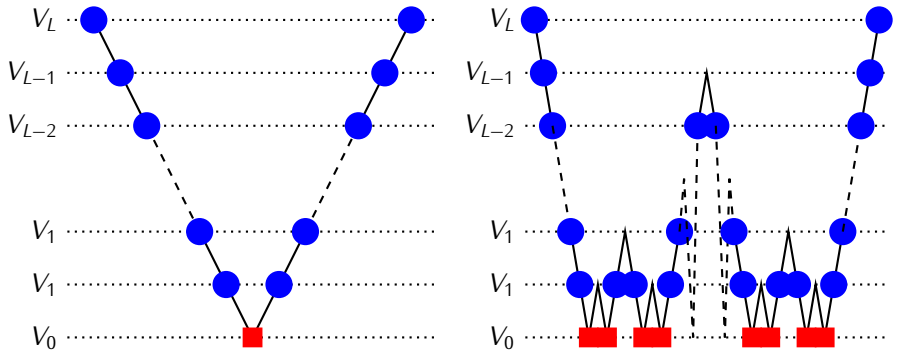

Figure 5.1: Smoothing and grid transfer of the V-cycle (left) and W-cycle (right). Black lines indicate grid transfer, blue dots are smoothing operations and red squares are coarse grid solvers. "Time" is left to right. fig:mg:1

Remark 5.7. If we use an additive or multiplicative Schwarz method (omitting the coarse grid) as our smoother R_ℓ, it should be possible in principle to use the analytical tools of Chapter 4. The difficulty then consists in ensuring that the spectral radius of the iteration matrix does not grow towards one if we proceed upwards on our scale of spaces V_ℓ. This remark is a todo for the author and an encouragement for the reader. Hints may be found in [GO95; Xu92].

Remark 5.8. It turns out that the techniques used for the analysis of the V-cycle and the W-cycle, respectively are quite different. Therefore, we separate them into two sections.

Theorem 5.9. *Let n_ℓ be the dimension of V_ℓ. Assume that the effort needed to for the operations in equations* (5.2)*b*/*c*/*e is linear in n_ℓ and assume that $n_{\ell+1}/n_\ell \approx 2^d$, where d is the space dimension of the grid. Assume that the effort for the coarse grid solver is negligible. Then, the effort for one step of the V-cycle is of order n_L. The effort for one step of the W-cycle is of order n_L for $d \geq 2$, while it is of order $n_L \log(n_L)$ in one dimension.*

Proof. Start the recursion on level L with the function $MG_L(0, g)$. This function calls $MG_{L-1}(\ldots)$ m_{coarse} times. Thus, by recursion, $MG_\ell(\ldots)$ is executed $m_{\text{coarse}}^{L-\ell}$ times.

By our assumptions, the amount of operations $\check{N}_\ell$ in $MG_\ell(\ldots)$ without the coarse grid correction is linear in n_ℓ, say bounded by Cn_ℓ. Then, the overall effort N_L on level L is

$$N_L \leq C \sum_{\ell=1}^{L} n_\ell m_{\text{coarse}}^{L-\ell} \leq C \sum_{\ell=1}^{L} n_L 2^{d(l-L)} m_{\text{coarse}}^{L-\ell} = C n_L \sum_{\ell=1}^{L} \left(\frac{m_{\text{coarse}}}{2^d} \right)^l. \tag{5.3}$$ eq:mg:6

It remains to notice that the sum converges and is bounded independent of L if and only if $m_{\text{coarse}}/2^d < 1$. The statements of the theorem follow immediately, observing that $L \simeq \log n_L$. □

Lemma 5.10. *Let B_ℓ^{-1} be the operator associated with the action of the multigrid preconditioner on level ℓ for $\ell = 0, \dots, L$. Then, the error after one step of the multigrid method has the form*

$$u^{(k+1)} - u = E_L \left(u^{(k)} - u \right), \tag{5.4}$$ eq:mg:7

where for $\ell = 0, \dots, L$ we denote by E_ℓ the error propagation operator

$$E_\ell = \left(I - R_\ell^{-1} A_\ell \right)^{m_{post}} \left(I - B_{\ell-1}^{-1} A_{\ell-1} P_{\ell-1} \right)^{m_{coarse}} \left(I - R_\ell^{-1} A_\ell \right)^{m_{pre}} \tag{5.5}$$ eq:mg:8

Proof. For the smoother, we use the standard technique for Richardson's method outlined in Lemma 3.16. For the coarse grid correction, we use Lemma 4.7. □

Note 5.11. The structure of the error propagation operator (5.5) already suggests the course of the multigrid analysis (as well as the design of smoothers). Namely, we will have to decompose V_l into V_{l-1} and its A-orthogonal complement. Then, we use the induction argument that $I - B_{\ell-1}^{-1} A_{\ell-1}$ is is small on V_{l-1}, while bounded on its complement. Vice versa, $I - R_\ell^{-1} A_\ell$ must be bounded on all of V_ℓ, while providing good reduction properties on the complement of V_{l-1}.

5.1 The V-cycle

assumption:mg:1

Assumption 5.12. *Let the smoother R_ℓ be symmetric, positive definite and let the following two conditions hold, the second for some positive constant α independent of ℓ:*

eq:mg:9

$$a\left((I - R_\ell^{-1} A_\ell) v, v \right) \geq 0 \qquad \forall v \in V_\ell, \tag{5.6a}$$ eq:mg:10

$$r(w, w) \leq \alpha a(w, w) \qquad \forall v \in V_\ell, \quad w = (I - P_{\ell-1}) v \tag{5.6b}$$ eq:mg:20

Theorem 5.13. *Let $a(.,.)$ be symmetric, positive definite and let Assumption 5.12 hold. Then, the V-cycle operator with $m_{pre} = m_{post} = m$ admits the estimate*

$$0 \leq a(\left((I - B_\ell^{-1} A_\ell) v, v \right) \leq \delta a(v, v), \qquad \forall v \in V_\ell, \tag{5.7}$$ eq:mg:11

where

$$\delta = \frac{\alpha}{\alpha + 2m}. \tag{5.8}$$ eq:mg:12

In particular, the contraction number of the multigrid method is bounded by a number less than 1, independent of the level.

Proof. [1] First, abbreviate $K_\ell = I - R_\ell^{-1} A_\ell$, the error propagation operator of a smoothing step. Now we will prove the theorem by induction over ℓ. First, since $B_0 = A_0$, it holds on level zero. For higher levels, we derive from (5.5) the relation

$$E_\ell = I - B_\ell^{-1} A_\ell = K_\ell^m \Big((I - P_{\ell-1}) + (I - B_{\ell-1}^{-1} A_{\ell-1}) P_{\ell-1} \Big) K_\ell^m. \tag{5.9}$$ eq:mg:13

Non-negativity follows readily by the induction argument and the same properties of the smoother and the Ritz-projection. For the upper bound, let $w = K_\ell^m v$ to obtain by the induction hypothesis

$$\begin{aligned} a(E_\ell v, v) &\le a\big((I - P_{\ell-1})w, w\big) + \delta a(P_{\ell-1} w, w) \\ &= (1-\delta) a\big((I - P_{\ell-1})w, w\big) + \delta a(w, w). \end{aligned} \tag{5.10}$$ eq:mg:14

Now we use the smoothing hypothesis anf the Bunyakovsky-Cauchy-Schwarz inequality for the bilinear form $r(.,.)$ associated with the smoothing operator R_ℓ to obtain

$$\begin{aligned} a\big((I - P_{\ell-1})w, w\big) &= \langle (I - P_{\ell-1})w, A_\ell w \rangle \\ &= \big\langle R_\ell (I - P_{\ell-1})w, R_\ell^{-1} A_\ell w \big\rangle \\ &= r\big((I - P_{\ell-1})w, R_\ell^{-1} A_\ell w\big) \\ &\le \sqrt{r\big((I - P_{\ell-1})w, (I - P_{\ell-1})w\big)} \sqrt{r\big(R_\ell^{-1} A_\ell w, R_\ell^{-1} A_\ell w\big)} \\ &\le \sqrt{\alpha a\big((I - P_{\ell-1})w, (I - P_{\ell-1})w\big)} \sqrt{a\big(R_\ell^{-1} A_\ell w, w\big)} \end{aligned}$$

Using the projection property of $I - P_{\ell-1}$, we obtain

$$a\big((I - P_{\ell-1})w, w\big) \le \alpha a\big(R_\ell^{-1} A_\ell w, w\big) = \alpha a\big((I - K_\ell) K_\ell^{2m} v, v\big). \tag{5.11}$$ eq:mg:15

The smoothing assumption also guarantees that the spectrum of K_ℓ is contained in the interval $[0, 1]$. Therefore,

$$a\big((I - K_\ell) K_\ell^{2m} v, v\big) \le a\big((I - K_\ell) K_\ell^{i} v, v\big), \qquad i = 0, \ldots, 2m, \tag{5.12}$$ eq:mg:16

yielding by deflating the telescoping sum

$$a\big((I - K_\ell) K_\ell^{2m} v, v\big) \le \frac{1}{2m} \sum_{i=0}^{2m-1} a\big((I - K_\ell) K_\ell^{i} v, v\big) = \frac{1}{2m} a\big((I - K_\ell^{2m}) v, v\big) \tag{5.13}$$ eq:mg:17

Combining (5.10), (5.11), and (5.13), we obtain

$$\begin{aligned} a(E_\ell v, v) &\le (1-\delta) \frac{\alpha}{2m} a\big((I - K_\ell^{2m}) v, v\big) + \delta a(K_\ell^m v, K_\ell^m v) \\ &= (1-\delta) \frac{\alpha}{2m} a(v, v) + \Big(\delta - (1-\delta) \frac{\alpha}{2m} \Big) a(K_\ell^m v, K_\ell^m v). \end{aligned} \tag{5.14}$$

[1] This version of the proof is taken from [AFW97]. It can also be found in [BH83; Bra93].

Finally, we enter $\delta = \alpha/(\alpha + 2m)$ to obtain

$$\delta - (1 - \delta)\frac{\alpha}{2m} = 0,$$

and thus

$$a(E_\ell v, v) \leq \left(1 - \frac{\alpha}{\alpha + 2m}\right) \frac{\alpha}{2m} a(v, v) = \frac{\alpha}{\alpha + 2m} a(v, v). \tag{5.15}$$ eq:mg:18

□

Lemma 5.14. *Let R_ℓ be the scaled additive Schwarz method*

$$R_a^{-1} = \omega \sum_{j=1}^{J} P_j A^{-1}, \tag{5.16}$$ eq:mg:19

where the subspaces V_j are defined by overlapping subdomains Ω_j as in (4.21)*. Not that we do not include the coarse space here. Then, for ω sufficiently small, this smoother fulfills Assumption 5.12.*

Proof. The positive definiteness and symmetry of the smoother have been proven in an abstract way in Lemma 4.10. In order to prove estimate (5.6a), we observe that by Lemma 4.11 and Lemma 4.29

$$r(v, v) = \omega \min_{v = \sum v_j} a(v_j, v_j) \gtrsim \omega a(v, v),$$

where the implicit constant depends on the number of overlaps in Definition 4.20. Thus, we can choose ω independent of ℓ such that

$$r(v, v) \geq a(v, v) \qquad \forall v \in V_\ell.$$

Accordingly, $R_\ell - A_\ell$ is positive definite, and since R_ℓ^{-1} is as well, so is $I - R_\ell^{-1} A_\ell$. It remains to prove (5.6b), but this is exactly the second half of the proof of Lemma 4.30, if we replace the Clément interpolant into the coarse space by the Ritz projection. □

Bibliography

ArnoldFalkWinther97Hdiv [AFW97] Douglas N. Arnold, Richard S. Falk, and R. Winther. "Preconditioning in $H(\mathrm{div})$ and applications." In: *Math. Comput.* 66.219 (1997), pp. 957–984. ISSN: 0025-5718. DOI: 10.1090/S0025-5718-97-00826-0.

BraessHackbusch83 [BH83] D. Braess and W. Hackbusch. "A New Convergence Proof for the Multigrid Method Including the V-cycle." In: *SIAM J. Sci. Comput.* 20.5 (1983), pp. 967–975.

BrennerScott02 [BS02] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. 2nd edition. Springer, 2002.

Bramble93 [Bra93] J. H. Bramble. *Multigrid Methods*. Pitman research notes in mathematics series 294. Longman Scientific, 1993.

DryjaWidlund94 [DW94] M. Dryja and O. Widlund. "Domain decomposition methods with small overlap." In: *SIAM J. Sci. Comput.* 15 (1994), pp. 604–620.

GriebelOswald95 [GO95] M. Griebel and P. Oswald. "On the abstract theory of additive and multiplicative Schwarz algorithms." In: *Numer. Math.* 70 (1995), pp. 163–180.

GrossmannRoosStynes07 [GRS07] Ch. Grossmann, H.-G. Roos, and M. Stynes. *Numerical Treatment of Partial Differential Equations*. 3rd edition. Springer, 2007.

Saad00 [Saa00] Y. Saad. *Iterative Methods for Sparse Linear Systems*. 2nd edition. Oxford University Press, 2000.

ToselliWidlund05 [TW05] A. Toselli and O. Widlund. *Domain decomposition methods—algorithms and theory*. Vol. 34. Springer Series in Computational Mathematics. Berlin: Springer-Verlag, 2005, pp. xvi+450.

Xu92 [Xu92] J. Xu. "Iterative Methods by Space Decomposition and Subspace Correction." In: *SIAM Review* 34.4 (1992), pp. 581–613.

Yosida80 [Yos80] K. Yosida. *Functional Analysis*. Springer, 1980.