SS 2005

Multigrid Methods

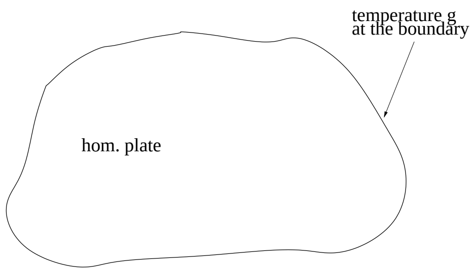
Professor Dr. Christoph Pflaum

Contents

1 Linear Equation Systems in the Numerical Solution of PDE's

1.1 Examples of PDE's

1. Heat Equation



Let us assume that there is a heat source f in the interior of the plate and that the temperature at the boundary is given by g. Question: What is the temperature inside of the plate?

Poisson Problem (P)

Let $\Omega \subset \mathbb{R}^n$ open, bounded, $f \in C(\overline{\Omega})$, $g \in C(\delta\Omega)$.

Find $u \in C^2(\overline{\Omega})$ such that

$$\begin{aligned} -\Delta u &= f \quad \text{on } \Omega \\ u\big|_{\delta\Omega} &= g \\ \text{where} \quad \Delta &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \end{aligned}$$

2. Poisson's equation with pure Neumann boundary conditions

Poisson Problem (P) with pure Neumann boundary conditions

Let $\Omega \subset \mathbb{R}^n$ an open and bounded domain and $f \in C(\overline{\Omega})$ such that $\int_\Omega u \; d(x,y) = 0$. Find $u \in C^2(\overline{\Omega})$ such that

$$\begin{aligned} -\Delta u &= f \quad \text{on } \Omega \\ \int_\Omega u \; d(x,y) &= 0. \end{aligned}$$

3. Let $\Omega \subset \mathbb{R}^2$ be an open domain. An anisotropic elliptic differential equation is an equation of the form

$$\begin{aligned} L(u) &:= -\operatorname{div} A \operatorname{grad} u + cu = f \quad \text{on} \quad \Omega \subset \mathbb{R}^2, \quad \text{where} \qquad (1) \\ A &= \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \in (L^1_{\text{loc}}(\Omega))^{2\times 2}, \qquad c \in L^1_{\text{loc}}(\Omega), \end{aligned}$$

and with suitable boundary conditions. Here, $A(x, y)$ is a symmetric positive semidefinite matrix and $c(x, y)$ is non-negative for almost every $(x, y) \in \Omega$. An additional assumption to the coefficients, described at the end of this section, guarantees that the stiffness matrix exists.

Anisotropic differential equations appear in several situations. For example equation (1) can describe a diffusion process with variable coefficients. Another example can be constructed by Poisson equation on a domain with a small hole (see [8]).

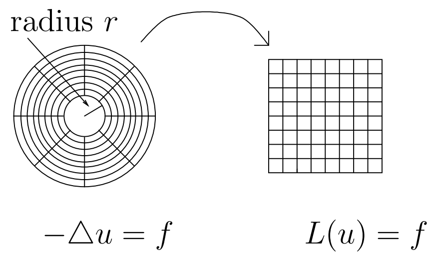


Figure 1: Transformation of a domain with a hole

Let us explain this example in more detail. Assume that the discretization grid is a tensor product grid as in Figure 22 .The bilinear finite element discretization on this grid has an equivalent formulation on the unit square. By the transformation of the curvilinear bounded domain onto the unit square one obtains an anisotropic elliptic differential equation on the unit square. If the radius r of the hole tends to zero, then the coefficients of this anisotropic elliptic equation become singular. For example they can tend to the following coefficients

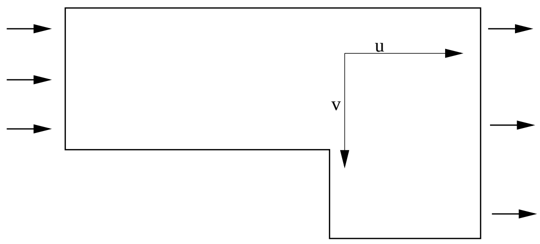
$$A = \begin{pmatrix} y^{-1} & 0 \\ 0 & y \end{pmatrix}. \tag{2}$$

4. Convection-Diffusion-Problem

   Find $u \in C^2(\overline{\Omega})$ such that

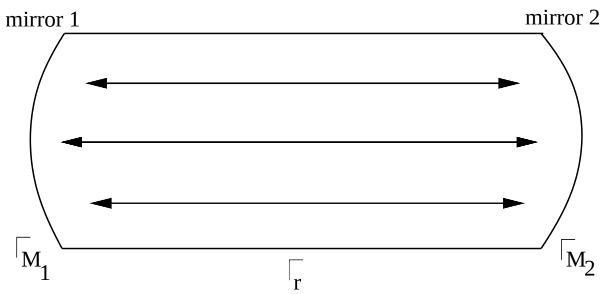
$$\begin{aligned} -\Delta u + \vec{b} \cdot \nabla u + c &= f \quad \text{on } \Omega \\ u\big|_{\delta\Omega} &= 0 \\ \text{where} \quad \vec{b} &\in (C(\Omega))^2, \quad f, c \in C(\Omega) \end{aligned}$$

5. Navier-Stokes-Equation



$$\begin{aligned}\frac{\partial u}{\partial t}+\frac{\partial p}{\partial x}+\frac{\partial (u^2)}{\partial x}+\frac{\partial (uv)}{\partial y} &= \frac{1}{\mathrm{Re}}\Delta u\\ \frac{\partial u}{\partial t}+\frac{\partial p}{\partial y}+\frac{\partial (uv)}{\partial x}+\frac{\partial (v^2)}{\partial y} &= \frac{1}{\mathrm{Re}}\Delta v\\ \frac{\partial u}{\partial x}+\frac{\partial v}{\partial y} &= 0\end{aligned}$$

6. Laser simulation



$$\Gamma_M = \Gamma_{M_1} \cup \Gamma_{M_2}$$

Find $u \in C^2(\overline{\Omega}),\ \lambda \in \mathbb{C}$ such that

$$\begin{aligned}-\Delta u - k^2 u &= \lambda u\\ u\big|_{\Gamma_M} &= 0\\ \frac{\partial u}{\partial \vec{n}}\Big|_{\Gamma_{\text{rest}}} &= 0 \qquad \text{(or boundary condition third kind)}\end{aligned}$$

We apply the ansatz

$$u = u_r e^{-i\tilde{k}z} + u_l e^{i\tilde{k}z}$$

where $\tilde{k}$ is an average value of k.
This leads to the equivalent eigenvalue problem:

Find u_r, u_l, λ such that

$$\begin{aligned}
-\Delta u_r \;+\; 2i\tilde{k}\;\frac{\partial u_r}{\partial z} \;+\; (\tilde{k}^2 - k^2)u_r \;&=\; \lambda u_r \\
-\Delta u_l \;-\; 2i\tilde{k}\;\frac{\partial u_l}{\partial z} \;+\; (\tilde{k}^2 - k^2)u_l \;&=\; \lambda u_l \\
u_r + u_l\Big|_{\Gamma_M} = 0, \qquad \frac{\partial u_r}{\partial z} - \frac{\partial u_l}{\partial z}\Big|_{\Gamma_M} \;&=\; 0 \\
\frac{\partial u_r}{\partial \vec{n}}\Big|_{\Gamma_{\text{rest}}} = \frac{\partial u_l}{\partial \vec{n}}\Big|_{\Gamma_{\text{rest}}} \;&=\; 0
\end{aligned}$$

1.2 Weak Formulation of Poisson's Equation

Let us first describe a physical problem which leads to Poisson's equation. Consider a thin plate with constant thermal conductivity. Figure 2 shows the geometry of such a plate described by the domain $\Omega \subset \mathbb{R}^2$.

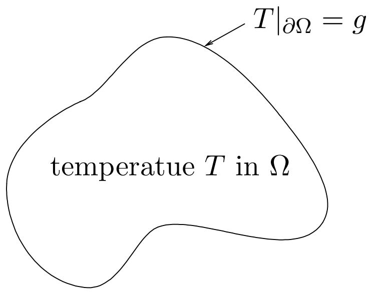


Figure 2: Temperate on a plate.

Assume that the boundary of the plate is maintained at temperature

$$T|_{\partial\Omega} = g.$$

Now, the Laplace's equation is governing the heat conduction within the plate (see [17]):

$$\begin{aligned}
\triangle T \;&=\; 0, \quad \text{where} \\
\triangle T \;&:=\; \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}.
\end{aligned}$$

Assume that $w \in \mathcal{C}^2(\bar{\Omega})$ is a function such that $w|_{\partial\Omega} = g$. Furthermore, assume $T \in \mathcal{C}^2(\bar{\Omega})$. Then, $u = T - w \in \mathcal{C}^2(\bar{\Omega})$ satisfies Poisson's equation with homogeneous Dirichlet boundary conditions

$$\begin{aligned}
-\triangle u \;&=\; f, && (3) \\
u|_{\partial\Omega} \;&=\; 0, && (4)
\end{aligned}$$

where $f := \triangle w$. For the mathematical analysis, it is more helpful to formulate this equation in a suitable Hilbert space. To this end, we multiply equation 3 by a test function $\varphi \in \tilde{\mathcal{C}}_0^1(\Omega)$ and integrate over Ω

$$-\int_\Omega \triangle u \varphi \, dz = \int_\Omega f \varphi \, dz.$$

Now, Green's formula yields

$$\int_\Omega \nabla u \nabla \varphi \, dz = \int_\Omega f \varphi \, dz. \tag{5}$$

By a continuity argument this equation holds for every function $\varphi \in H_0^1(\Omega)$. The left hand side of equation 12 is the bilinear form defined by 93. This shows that the Sobolev space $H_0^1(\Omega)$ is the right Hilbert space for the description of Poisson's equation. Furthermore, the mapping

$$H_0^1(\Omega) \to \mathbb{R}, \qquad v \mapsto \int_\Omega f v \, dz$$

is contained in the dual space $(H_0^1(\Omega))'$, since Lemma 5 implies

$$\left| \int_\Omega f v \, dz \right| \;\leq\; \|f\|_{L^2(\Omega)} \, \|v\|_{L^2(\Omega)} \;\leq\; c\,|v|_{H^1}.$$

These considerations lead to the following weak formulation of Poisson's equation:

Problem 1 (Poisson's equation)**.** *Assume that $f \in L^2(\Omega)$. Find $u \in H_0^1(\Omega)$ such that*

$$\int_\Omega \nabla u \nabla v \, dz = \int_\Omega f v \, dz \quad \textit{for every} \;\; v \in H_0^1(\Omega). \tag{6}$$

Theorem 10 guarantees the existence and uniqueness of the solution of this weak equation.

1.3 Finite-Difference-Discretization of Poisson's Equation

Assume $\Omega =]0,1[^2$ and that an exact solution of (P) exists. We are looking for an approximate solution u_h of (P) on a grid Ω_h of meshsize h. Choose $h = \frac{1}{m}$ where $m \in \mathbb{N}$.

$$\begin{aligned} \Omega_h &= \big\{(ih, jh)\big| i,j = 1,\dots,m-1\big\} \\ \overline{\Omega_h} &= \big\{(ih, jh)\big| i,j = 0,\dots,m\big\} \end{aligned}$$

Discretization by Finite Differences:
Idea: Replace second derivative by difference quotient.
Let $e_x = (1,0)$ and $e_y = (1,0)$,

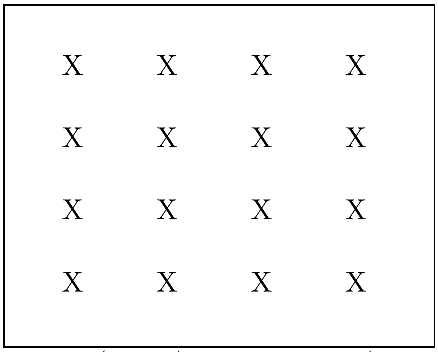
$$-\Delta u(z) = \left(-\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2}\right)(z) = f(z) \quad \text{for } z \in \Omega_h$$

$$
\begin{aligned}
-\frac{u_h(z+he_x) - 2u_h(z) + u_h(z-he_x)}{h^2} & \\
-\frac{u_h(z+he_y) - 2u_h(z) + u_h(z-he_y)}{h^2} & = f(z)
\end{aligned}
$$

$$
\begin{array}{lccc}
\text{and} & u(z) & = & g(z) \\
 & \approx & = & \text{for } z \in \overline{\Omega_h} \setminus \Omega_h \\
 & u_h(z) & = & g(z)
\end{array}
$$

This leads to a linear equation system $L_h\,U_h \;=\; F_h$ where $U_h = (u_h(z))_{z\in\Omega_h}$, L_h is $|\Omega_h| \times |\Omega_h|$ matrix. The discretization can be described by the stencil

$$
\begin{pmatrix} & -\frac{1}{h^2} & \\ -\frac{1}{h^2} & \frac{4}{h^2} & -\frac{1}{h^2} \\ & -\frac{1}{h^2} & \end{pmatrix} = \begin{pmatrix} m_{-1,1} & m_{0,1} & m_{1,1} \\ m_{-1,0} & m_{0,0} & m_{1,0} \\ m_{-1,-1} & m_{0,-1} & m_{1,-1} \end{pmatrix}
$$



Let us abbreviate $U_{i,j} := u_h(ih, jh)$ and $f_{i,j} := f(ih, jh)$. Then, in case of $g = 0$, the matrix equation $L_h U_h = F_h$ is equivalent to:

$$
\sum_{k,l=-1}^{1} m_{kl} U_{i+k,j+l} = f_{i,j}
$$

1.4 FD Discretization for Convection-Diffusion

Let Ω, Ω_h as above.

$$
-\Delta u + b\frac{\mathrm{d}u}{\mathrm{d}x} = f
$$

Assume that b is constant.

1. Discretization by central difference:

$$
\frac{\mathrm{d}u}{\mathrm{d}x}(z) \approx \frac{u_h(z+he_x) - u_h(z-he_x)}{2h}
$$

This leads to the stencil

$$\begin{pmatrix} & -\frac{1}{h^2} & \\ -\frac{1}{h^2} - \frac{b}{2h} & \frac{4}{h^2} & -\frac{1}{h^2} + \frac{b}{2h} \\ & -\frac{1}{h^2} & \end{pmatrix}$$

$\rightarrow$ unstable for large b.

2. Upwind discretization:

$$\frac{\mathrm{d}u}{\mathrm{d}x}(z) \approx \frac{u_h(z) - u_h(z - he_x)}{h}$$

This leads to the stencil

$$\begin{pmatrix} & -\frac{1}{h^2} & \\ -\frac{1}{h^2} - \frac{b}{h} & \frac{4}{h^2} + \frac{b}{h} & -\frac{1}{h^2} \\ & -\frac{1}{h^2} & \end{pmatrix}$$

1.5 Irreducible and Diagonal Dominant Matrices

Definition 1. *A $n \times n$ matrix A is called* strong diagonal dominant*, if*

$$|a_{ii}| > \sum_{i \neq j} |a_{ij}| \qquad 1 \leq i \leq n \tag{7}$$

A is called weak diagonal dominant, if there exists at least one i such that (7) holds and such that

$$|a_{ii}| \geq \sum_{i \neq j} |a_{ij}| \qquad 1 \leq i \leq n$$

Definition 2. *A is called reducible, if there exists a subset $J \subset \{1, 2, \ldots, n\}$, such that*

$$a_{ij} = 0 \qquad \text{for all } i \notin J,\ j \in J$$

A not reducible matrix is called irreducible.

Remark. An reducible matrix has the form

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

$\rightarrow$ The equation system separates in two parts.

Example:

1. Poisson FD:
   diagonal: $a_{ii} = \frac{4}{h^2}$
   non-diagonal: $a_{ij} = \begin{cases} -\frac{1}{h^2} & \text{if } i \text{ is N,S,W,O of } j \\ 0 & \text{else} \end{cases}$

   - A is not strong diagonal dominant, but weak diagonal dominant. To see this, consider a point i such that j is N of i. Then

   $$a_{ij} = \begin{cases} -\frac{1}{h^2} & \text{if } i \text{ is S,W,O of } j \\ 0 & \text{else} \end{cases}$$

   - A is irreducible.
     Proof: If A is reducible, then, $\{1, 2, ..., n\}$ is the union of two different sets of colored points, where one set is J. Then, there is a point $j \in J$ such that one of the points i=N,W,S,E is not contained in J, but i is contained in $\{1, 2, ..., n\}$. This implies $a_{j,i} \neq 0$. $\Rightarrow$ contradiction.

2. Convection-Diffusion-Equation

   - centered difference

   $$\begin{aligned} |a_{ii}| &= \frac{4}{h^2} \\ \sum_{i \neq j} |a_{ij}| &= \frac{4}{h^2} 2 \cdot \frac{1}{h^2} + \left(\frac{1}{h^2} + \frac{b}{2h}\right) + \left|\frac{1}{h^2} - \frac{b}{2h}\right| \\ &= 3\frac{1}{h^2} + \frac{b}{2h} + \left|\frac{1}{h^2} - \frac{b}{2h}\right| \end{aligned}$$

   Thus, $|a_{ii}| \geq \sum_{i \neq j} |a_{ij}|$, if and only if $\frac{1}{h^2} - \frac{b}{2h} \leq 0$.
   This shows $|a_{ii}| \geq \sum_{i \neq j} |a_{ij}|$, if and only if $h < \frac{2}{b}$
   - upwind

   $$\begin{aligned} |a_{ii}| &= \frac{4}{h^2} + \frac{b}{h} \\ &\geq \\ \frac{4}{h^2} + \frac{b}{h} &\geq \sum_{i \neq j} |a_{ij}| \quad \text{for all } h, b > 0 \end{aligned}$$

   - Conclusion
     central: A is weak diagonal dominant if and only if $h < \frac{2}{b}$.
     upwind: A is weak diagonal dominant.
     A is irreducible in both cases.

1.6 FE (Finite Element) Discretization

Definition 3. *$\mathcal{T} = \{T_1, \ldots, T_M\}$ is a conform triangulation of Ω if*

- $\overline{\Omega} = \bigcup_{i=1}^{M} T_i$, $\qquad$ *T_i is triangle or square*
- *$T_i \cap T_j$ is either*
  - *empty or*
  - *one common corner or*
  - *one common edge.*

Remark.

- Let us write $\mathcal{T}_h$, if the diameter h_T of every element $T \in \mathcal{T}_h$ is less or equal h:

$$h_T \le h.$$

- A family of triangulations $\{\mathcal{T}_h\}$ is called quasi-uniform, if there exists a constant $\rho > 0$ such that the radius ρ_T of the largest inner ball of every triangle $T \in \mathcal{T}_h$ satisfies

$$\rho_T > \rho h.$$

Definition 4.

- *Let $\mathcal{T}_h$ be a triangulation of Ω. Then, let V_h be the space of linear finite elements defined as follows:*

$$\begin{aligned} V_h &= \left\{ v \in C^0(\overline{\Omega}) \;\middle|\; v\big|_T \text{ is linear for every } T \in \mathcal{T}_H \right\} \\ \overset{0}{V}_h &= V_h \cap H_0^1(\Omega) \end{aligned}$$

$v\big|_T$ is linear means that $v\big|_T(x,y) = a + bx + cy$.

- *Let $\Omega =]0,1[$, $h = \frac{1}{m}$ and*

$$\mathcal{T}_h = \left\{ [ih, (i+1)h] \times [jh, (j+1)h] \;\middle|\; i, j = 0, \ldots, m-1 \right\}$$

The space of bilinear finite elements on Ω is defined as follows

$$V_h = \left\{ v \in C^0(\overline{\Omega}) \;\middle|\; v\big|_T \text{ is bilinear for every } T \in \mathcal{T}_H \right\}$$

$v\big|_T$ is bilinear means that $v\big|_T(x,y) = a + bx + cy + dxy$.

- *Let V_h be the space of linear or bilinear finite elements on $\mathcal{T}_h$ and $\mathcal{N}_h$ the set of corners of $\mathcal{T}_h$. Then, define the nodal basis function $v_p \in V_h$ at the point p by:*

$$v_p(x) = \begin{cases} 1 & \text{if } x = p \\ 0 & \text{if } x \neq p \end{cases} \qquad \text{for } x \in \mathcal{N}_h$$

Observe that

$$V_h = span\left\{ v_p \;\middle|\; p \in \mathcal{N}_h \right\}$$

This means that every function $u_h \in V_h$ can be represented as

$$u_h = \sum_{p \in \mathcal{N}_h} \lambda_p v_p$$

Finite Element Discretization of Poisson's equation:

$$\begin{aligned} -\Delta u &= f \\ u\big|_{\delta\Omega} &= 0 \end{aligned}$$

Thus, for every $v_h \in \overset{0}{V}_h$, we get:

$$\begin{aligned} -\Delta u \, v_h &= f \, v_h \\ &\Downarrow \\ \int_\Omega \nabla u \, \nabla v_h \, \mathrm{d}(x,y) + \int_\Gamma \frac{\partial u}{\partial \vec{n}} \, v_h \, \mathrm{d}(x,y) &= \int_\Omega f \, v_h \, \mathrm{d}(x,y) \\ &\Downarrow \\ \int_\Omega \nabla u \, \nabla v_h \, \mathrm{d}(x,y) &= \int_\Omega f \, v_h \, \mathrm{d}(x,y) \qquad \forall v_h \in \overset{0}{V}_h \end{aligned}$$

FE Discretization: Find $u_h \in \overset{0}{V}_h$ such that

$$\int_\Omega \nabla u \, \nabla v_h \, \mathrm{d}(x,y) = \int_\Omega f \, v_h \, \mathrm{d}(x,y) \qquad \forall v_h \in \overset{0}{V}_h \tag{8}$$

Stiffness matrix.

$$\begin{aligned} a_{p.q} &:= \int_\Omega \nabla v_p \, \nabla v_q \, \mathrm{d}(x,y), \qquad f_q := \int_\Omega f \, v_q \, \mathrm{d}(x,y) \\ A &:= (a_{p,q})_{p,q \in \overset{0}{\mathcal{N}}_h}, \qquad \overset{0}{\mathcal{N}}_h := \mathcal{N}_h \cap \Omega \\ u_h &= \sum_{p \in \overset{0}{\mathcal{N}}_h} \lambda_p \, v_p \end{aligned}$$

Then, (8) implies

$$\begin{aligned}
\sum_{p\in\overset{0}{\mathcal{N}}_h} \lambda_p \int_\Omega \nabla v_p \, \nabla v_q \; \mathrm{d}(x,y) \quad &= \quad \int_\Omega f \; v_q \; \mathrm{d}(x,y) \qquad \text{for all } q \in\overset{0}{\mathcal{N}}_h \\
&\Downarrow \\
\sum_{p\in\overset{0}{\mathcal{N}}_h} \lambda_p \; a_{p,q} \quad &= \quad f_q \qquad \forall q \in\overset{0}{\mathcal{N}}_h \\
&\Downarrow \\
A\, U_h \quad &= \quad F_h \qquad \text{where} \qquad \begin{array}{l} U_h = (\lambda_p)_{p\in\overset{0}{\mathcal{N}}_h} \\ F_h = (f_q)_{q\in\overset{0}{\mathcal{N}}_h} \end{array}
\end{aligned}$$

The matrix A is called the stiffness matrix of the FE discretization.

1.7 Discretization Error and Algebraic Error

Let $||\cdot||$ be a suitable norm. Then, $||U_h - U||$ is called discretization error, with respect to this norm.

Example 1. *Poisson on a square*

- *FD,* $u \in C^4(\overline{\Omega})$*, then*

$$||U_h - U||_{L^\infty(\Omega_h)} = O(h^2)$$

- *FE,* $u \in H^2(\overline{\Omega})$*, then*

$$\begin{aligned} ||U_h - U||_{L^2(\Omega)} &= O(h^2) \\ ||U_h - U||_{H^1(\Omega)} &= O(h) \end{aligned}$$

Problem. The solution u_h cannot be calculated exactly, since L_h (or A) is a very large matrix and

$$A\, U_h = F_h.$$

Therefore, we need iterative solvers if $n > 10.000$ (or $n > 100.000$). By such an iterative solver, we get an approximation $\tilde{u}_h$ of u_h. $||\tilde{u}_h - u_h||$ is called algebraic error.

1.8 Basic Theory

Let A be a non singular $n \times n$ matrix and b a vector, $b \in \mathbb{R}^n$.

Problem:

Find $x \in \mathbb{R}^n$ such that $A\, x = b$.

A linear iterative method to solve this equation system is:

Algorithm:

Let x^0 be the start guess. Then
$x^{k+1} := Cx^k + d$

Here x must be a fixed point of $x := Cx + d$.

Theorem 1. x^k converges to x for every start vector x^0 if and only if

$$\rho(C) < 1$$

Here $\rho(C)$ is the spectral radius of C,

$$\rho(C) = \max \left\{ |\lambda| \middle| \lambda \text{ is eigenvalue of } C \right\}$$

(Observe the eigenvalues may be complex.)

Furthermore, the following convergence result holds:

$$||x^k - x|| \leq ||C^k||\, ||x^0 - x|| \tag{9}$$

If C is a normal matrix, then

$$||x^k - x||_2 \leq (\rho(C))^k \; ||x^0 - x||_2 \tag{10}$$

There exist start vectors x^0, such that the equal sign holds in the above inequality.

1.9 Aim of a Multigrid Algorithm

Let us assume that the linear equation system comes from the discretization of a partial differential equation. The the iteration method depends on the meshsize h. The aim is to construct a (linear) iterative method such

- that the computational amount of one iteration is proportional to the number of unknowns and
- such that

  $$\rho_h(C) < \rho < 1$$

  where ρ is a fixed constant.

1.10 Jacobi and Gauss-Seidel Iteration

The Jacobi-iteration is a „one-step“ method. The Gauss-Seidel-iteration is a successive relaxation method.

1.10.1 Ideas of Both Methods

Relaxation of the i-th unknown x_i:
Correct x_i^{old} by x_i^{new} such that the i-th equation of the equation system

$$A \cdot x = b$$

is correct.

Jacobi-iteration:
„Calculate the relaxations simultaneously for all $i = 1, \dots, n$“
This means: If $x^{old} = x^k$, then
let $x^{k+1} = x^{new}$

Gauss-Seidel-iteration:
„Calculate relaxation for $i = 1, \dots, n$ and use the new values“
This means: $x^{old,1} = x^k$
Iterate for $i = 1, \dots, n$:
Calculate $x^{new,i}$ by relaxation of the i-th component

Put $x^{old,i+1} = x^{new,i}$
$x^{k+1} = x^{new,n}$

Teh iteration matrix of the Gauss-Seidel iteration is

$$C_{GS} = (D - L)^{-1} R$$

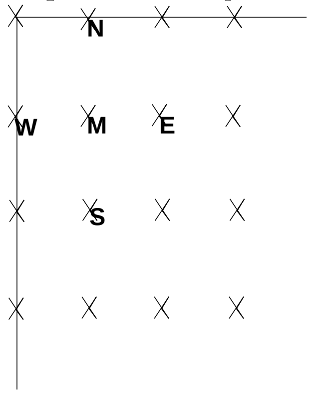
and the iteration matrix of the Jacobi iteration is

$$C_{GS} = D^{-1}(L + R)$$

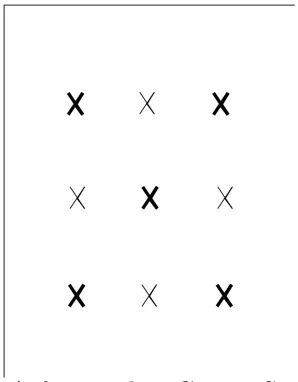
Remark.

- Jacobi iteration is independent of the numbering of the grid points
- The convergence rate of the Gauss-Seidel iteration depends on the numbering of the grid points

Example 2. *Model problem, FD for Poisson*

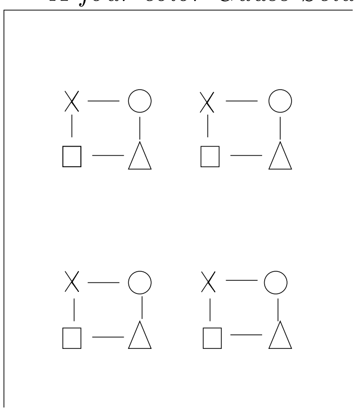


$$u_M^{new} = \frac{1}{4}\left(u_N^{old} + u_S^{old} + u_E^{old} + u_W^{old}\right) + f_M$$



red-black Gauss-Seidel

A four color Gauss-Seidel-relaxation is used for a 8-point stencil



$$\begin{matrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{matrix}$$

- better relaxation property
- after relaxation of one color all equations at those points are correct

<u>Relaxation</u> for the Convection-Diffusion:
A convection-diffusion problem is a so-called singular perturbed problem. To see this write the convection-diffusion problem in the form:

$$-\epsilon\Delta u + \frac{\partial u}{\partial x} = \tilde{f} \quad , \quad \epsilon > 0$$

$\epsilon \to 0$ is the difficult case.

(Hackbusch's) rule for relaxing singular perturbed problems: Construct the iteration such that it is an exact solver for $\epsilon = 0$

For $\epsilon = 0$ we get the stencil (for upwind FD):

$$\begin{pmatrix} & 0 & \\ -\frac{1}{h} & \frac{1}{h} & 0 \\ & 0 & \end{pmatrix}$$

Thus a Gauss-Seidel relaxation with a numbering of the grid points from left to right leads to an exact solver

1	2	3
4	5	6
7	8	9

This can be done also for more complicated convection directions. Exception: Circles!

1.11 Convergence Rate of Jacobi and Gauss-Seidel Iteration

1.11.1 Analysis of the Convergence of the Jacobi Method

Let us consider Poisson's equation on a unit square. Let $Ax = b$ the corresponding linear system and

$$A = D - L - R,$$

where D is the diagonal matrix.

Then, the iteration matrix of the Jacobi method is $C_j = D^{-1}(L + R)$. I case of the model problem Poisson's equation, we get

$$A = D-L-R \Longrightarrow C_j = D^{-1}(D-A) = -D^{-1}A+E+E-\frac{h^2}{4}A = E-\frac{h^2}{4}L_h \tag{11}$$

Let $e_{\nu\mu}$ be the eigenfunctions of A and $\lambda_{\nu\mu}$ the corresponding eigenvalues. This is

$$e_{\nu\mu} = \Big(\sin(\nu\pi hi)\sin(\mu\pi hj)\Big)_{i,j=1,\ldots,m-1}$$

Then, we get

$$C_J e_{\nu\mu} = \left(1 - \frac{h^2}{4}\lambda_{\nu\mu}\right) e_{\nu\mu}. \tag{12}$$

Here $\lambda_{\nu,\mu}$ are the eigenvalues

$$\lambda_{\nu,\mu} = \frac{4}{h^2}\left(\sin^2\left(\frac{\pi\nu h}{2}\right) + \sin^2\left(\frac{\pi\mu h}{2}\right)\right)$$

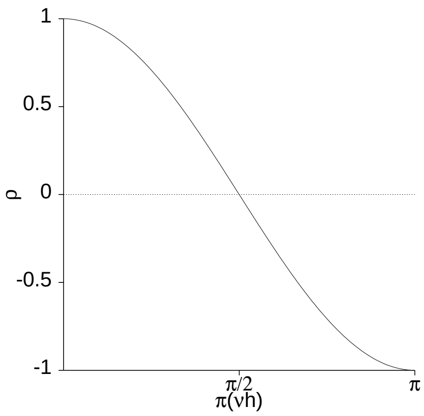
for $\nu, \mu = 1 \ldots (m-1)$, where $h = \frac{1}{m}$. Thus, the iteration matrix C_j has the eigenvalues

$$(\rho_J)_{\nu\mu} = 1 - \sin^2\left(\frac{\pi\nu h}{2}\right) - \sin^2\left(\frac{\pi\mu h}{2}\right) \tag{13}$$

Here, J denotes the Jacobi method. In case of $\nu = \mu$ we have,

$$(\rho_J)_{\nu\nu} = 1-\sin^2\left(\frac{\pi\nu h}{2}\right)-\sin^2\left(\frac{\pi\nu h}{2}\right) = 1-2\sin^2\left(\frac{\pi\nu h}{2}\right) = \cos(\pi\nu h) \tag{14}$$

The following graph depicts the eigenvalues $(\rho_J)_{\nu\nu}$ with respect to the parameter $\pi\nu h$ in (14).



The eigenvalues $\rho_{\nu\mu}$ of the matrix C describe how the algebraic error

$$x^k - x = \sum c_{\nu\mu} e_{\nu\mu}$$

is reduced by one iteration, since

$$x^{k+1} - x = \sum (c_{\nu\mu}\rho_{\nu\mu}) e_{\nu\mu}.$$

$\Longrightarrow$ Bad convergence for high and low frequencies.

$\Longrightarrow$ Good convergence for middle frequencies.

In particular, one can show that the spectral radius of the iteration matrix is

$$\rho(C) = 1 - O(h^2) \tag{15}$$

1.11.2 Iteration Method with Damping Parameter

Let us assume that $x^k \longrightarrow x^{k+1}$ is an iteration. The iteration can be written as $x^k \longrightarrow x^k + (x^{k+1} - x^k)$. The term $(x^{k+1} - x^k)$ can be treated as a correction term. Now a damped iteration is $x^k \longrightarrow \omega(x^{k+1} - x^k)$, where

- ω is called the damping factor or the relaxation parameter and $\omega \in]0, 2[$.

- $\omega > 1$ is called over relaxation.

- $\omega < 1$ is called under relaxation.

SOR(Successive Over Relaxation) method is obtained by performing the Gauss-Seidel method with over relaxation. But SOR has disadvantages for e.g like,

- It is very difficult to find ω for certain class of problems.

1.11.3 Damped Jacobi Method

The Jacobi method with relaxation parameter $\omega = 1$ is

$$x^{k+1}_{Jacobi} = D^{-1}(L + R)x^k_{Jacobi} + D^{-1}b \tag{16}$$

The Jacobi method with damping parameter ω is

$$\begin{aligned} x^{k+1}_\omega &= x^k_\omega + \omega(D^{-1}(L+R)x^k_\omega + D^{-1}b - x^k_\omega) \\ &= \left\{E(1-\omega) + \omega D^{-1}(L+R)\right\} x^k_\omega + \omega D^{-1}b & (17) \\ \Longrightarrow\ & C_\omega = E(1-\omega) + \omega D^{-1}(L+R) & (18) \end{aligned}$$

This is the iteration matrix of the damped Jacobi method.

1.11.4 Analysis of the Damped Jacobi method

The iteration matrix of the damped Jacobi method can be written as

$$C_{J,\omega} = E(1-\omega) + wD^{-1}(D - A) = E - \omega D^{-1}A = E - \omega\frac{h^2}{4}A \tag{19}$$

Furthermore, by (18), the iteration matrix of the damped Jacobi method is

$$C_{J,\omega} = [E + \omega C_j - \omega E] = (1-\omega)E + \omega C_j \tag{20}$$

where C_j is the iteration matrix of the Jacobi method. The eigenvalues of the iteration matrix of the Jacobi method are

$$(\rho_J)_{\nu,\mu} = 1 - \left[\sin^2\left(\frac{\pi\nu h}{2}\right) + \sin^2\left(\frac{\pi\mu h}{2}\right)\right]$$

Thus, the eigenvalues of the iteration matrix of the damped Jacobi method are

$$(\rho_{J,\omega})_{\nu,\mu} = 1 - \omega\left[\sin^2\left(\frac{\pi\nu h}{2}\right) + \sin^2\left(\frac{\pi\mu h}{2}\right)\right] \tag{21}$$

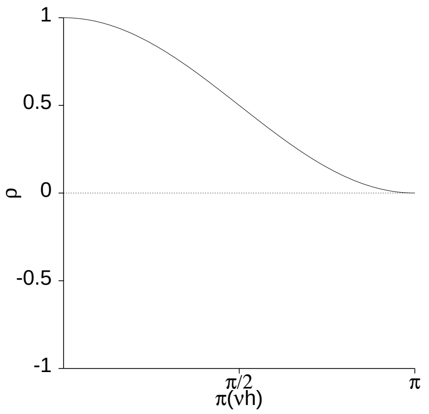
Now, for $\nu = \mu$, we have

$$(\rho_{J,\omega})_{\nu,\nu} = 1 - 2\omega\left[\sin^2\left(\frac{\pi\nu h}{2}\right)\right] \tag{22}$$

Thus, if $\omega = \frac{1}{2}$

$$(\rho_{J,\omega})_{\nu,\nu} = 1 - \left[\sin^2\left(\frac{\pi\nu h}{2}\right)\right] \tag{23}$$

The following graph depicts the eigenvalues $(\rho_{J,\omega})_{\nu\nu}$ with respect to the parameter $\pi\nu h$ in (23).

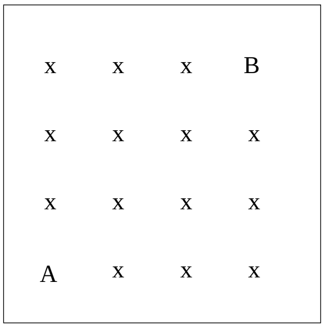


This shows that the damped Jacobi method with $\omega = \frac{1}{2}$ has the properties

- Bad convergence for low frequencies.
- Good convergence for high frequencies.

The Gauss–Seidel method has similar properties as the damped Jacobi method with $\omega = \frac{1}{2}$.

1.11.5 Heuristic approach



By single step methods we require $O(\sqrt{n}) = O(h^{-1})$ operations for a correction in B due to a change in A. The idea is to achieve faster correction by using a coarser grid.

2 Classical Multigrid Algorithm

2.1 Multigrid algorithm on a Simple Structured Grid

2.1.1 Multigrid

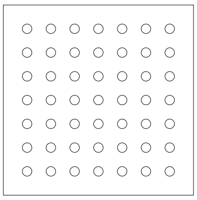


Figure 3: l=3

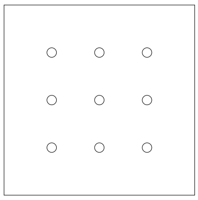


Figure 4: l=2

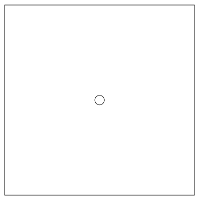


Figure 5: l=1

The classical multigrid algorithm is described in [9], [10], [6], [7] or [18]. Let l be the number of levels such that $l_{max} \in \mathbb{N}$ and

$$m_l = 2^l$$
$$n_l = (m_l - 1)^2$$
$$h_l = 2^{-l}$$

for $l = 1 \dots l_{max}$.

Let us assume that a PDE (e.g. Poisson's equation) is given. Discretize this equation by the grids $\Omega_l := \Omega_{h_l}$ where $l = 1, \dots, l_{max}$. This leads to the discrete matrix equations

$$A_l x_l = b_l \tag{24}$$

where $b_l, x_l \in S_l$ and $S_l = \mathbb{R}^{n_l}$. The matrix A_l is an invertible matrix of order $n_l \times n_l$.

Let an iterative solution for (24) be given as

$$x_l^{k+1} \quad = \quad C_l^{relax} x_l^k + N_l b_l \quad = \quad \mathcal{S}_{l,b_l}(x_l^k) \tag{25}$$

2.1.2 Idea of Multigrid Algorithm

Let $\widetilde{x}_l$ be an approximate solution for (24). The algebraic $\widetilde{e}_l$ is defined as

$$\widetilde{e}_l = x_l - \widetilde{x}_l. \tag{26}$$

Now $\widetilde{e}_l$ has to be calculated in order to find x_l. The following residual equation is valid for $\widetilde{e}_l$,

$$A_l \widetilde{e}_l = r_l \tag{27}$$

where r_l is called the residual and is given by

$$r_l = b_l - A_l \widetilde{x}_l \tag{28}$$

The aim is to find an approximate solution of the residual equation by solving the equation approximately on a course grid Ω_{l-1}. To this end, we need the following matrix operators

- Restriction operator
$$I_l^{\ l-1} : S_l \mapsto S_{l-1}$$
- Prolongation operator
$$I_{l-1}^{l} : S_{l-1} \mapsto S_l$$

2.1.3 Two–grid Multigrid Algorithm

Two–grid Multigrid algorithm with parameters v_1 and v_2

Let x_l^k be an approximate solution of (24) and v_1 and v_2 the parameters of pre–smoothing and post–smoothing.

1. Step 1 (Pre–smoothing)

$$x_l^{k,1} = \mathcal{S}_{l,b_l}^{\,v_1} x_l^k \tag{29}$$

2. Step 2 (Coarse grid correction)

   Residual calculation :

$$r_l = b_l - A_l x_l^{k,1} \tag{30}$$

   Restriction :

$$r_{l-1} = I_l^{\,l-1} r_l \tag{31}$$

   Solve on coarse grid:

$$e_{l-1} = A_{l-1}{}^{-1} r_{l-1} \tag{32}$$

   Prolongation :

$$e_l = I_{l-1}^l e_{l-1} \tag{33}$$

   Correction :

$$x_l^{k,2} = x_l^{k,1} + e_l \tag{34}$$

3. Step 3 (Post–smoothing)

$$x_l^{k+1} = \mathcal{S}_{l,b_l}^{\,v_2} (x_l^{k,2}) \tag{35}$$

2.1.4 Restriction and Prolongation Operators

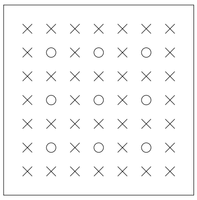


Figure 6: O–Coarse grid point and X–Fine grid point

Let us abbreviate $x_{i,j} = x_{(ih_{l-1}, jhl-1)}$ and set $x_{i,j} = 0$ for $i = 0$ or $j = 0$ or $i = m_{l-1}$ or $j = m_{l-1}$.

2.1.5 Prolongation or Interpolation

The interpolation or prolongation of $x_{i,j}$ given by $w_{i,j} = \{I_{l-1}^{l}(x)\}_{(ih_l, jh_l)}$ is defined by the following equations

$$w_{2i,2j} = \frac{1}{2} x_{i,j} \tag{36}$$

$$w_{2i+1,2j} = \frac{1}{4}(x_{i,j} + x_{i+1,j}) \tag{37}$$

$$w_{2i,2j+1} = \frac{1}{4}(x_{i,j} + x_{i,j+1}) \tag{38}$$

$$w_{2i+1,2j+1} = \frac{1}{8}(x_{i,j} + x_{i+1,j} + x_{i,j+1} + x_{i+1,j+1}) \tag{39}$$

2.1.6 Pointwise Restriction

Piecewise restriction is rarely applied and defined by

$$\{\dot{I}_l^{\,l-1}(x)\}_{(ih_{l-1}, jh_{l-1})} = x_{2i,2j} \tag{40}$$

The quality of this restriction operator is not very good.

2.1.7 Weighted Restriction

Weighted restriction or full weighting is defined by

$$\begin{aligned}\{I_l^{\,l-1}(x)\}_{(ih_{l-1}, jh_{l-1})} \quad = \quad & \frac{1}{8}(x_{2i+1,2j+1} + x_{2i-1,2j+1} + x_{2i+1,2j-1} + x_{2i-1,2j-1}) + \\ & \frac{1}{4}(x_{2i+1,2j} + x_{2i-1,2j} + x_{2i,2j+1} + x_{2i,2j-1}) + \\ & \frac{1}{2} x_{2i,2j}\end{aligned}$$

Remark

$$\left(I_l{}^{l-1}\right)^T = I^l_{l-1} \tag{41}$$

2.2 Iteration Matrix of the Two–Grid Multigrid Algorithm

Theorem 1. *The iteration matrix of a two–grid Multigrid algorithm is*

$$C_l^{two\_grid} = \left(C_l{}^{relax}\right)^{v_2}\left(E - I^l_{l-1}(A_{l-1})^{-1}I_l{}^{l-1}A_l\right)\left(C_l{}^{relax}\right)^{v_1} \tag{42}$$

Proof

The coarse grid correction is

$$\begin{aligned} x_l{}^{k,2} &= x_l{}^{k,1} + I^l_{l-1}(A_{l-1})^{-1}I_l{}^{l-1}(b_l - A_l x_l{}^{k,1}) \\ &= \left(E - I^l_{l-1}(A_{l-1})^{-1}I_l{}^{l-1}A_l\right)x_l{}^{k,1} + I^l_{l-1}(A_{l-1})^{-1}I_l{}^{l-1}b_l \end{aligned}$$

Therefore the iteration matrix of the coarse grid correction of the two–grid Multigrid algorithm is

$$\left(E - I^l_{l-1}(A_{l-1})^{-1}I_l{}^{l-1}A_l\right)$$

A short calculation shows that the iteration matrix of two linear iteration algorithms is the product of the iteration matrices of these algorithms.

2.3 Multigrid Algorithm

Multigrid algorithm $MGM(x_l^k, b_l, l)$ with parameters (v_1,v_2,μ)

Let x^k_{lmax} be an approximate solution of (24). Then,

$$x^{k+1}_{lmax} = MGM(x^k_{lmax})$$

is the approximate solution of (24) by the multigrid algorithm with an initial vector x^k_{lmax}. The multigrid algorithm can then be described as

If $l = 1$ then $MGM(x_l^k, b_l, l) = A_l^{-1}b_l$

If $l > 1$ then

Step 1 (v_1-pre–smoothing)

$$x_l^{k,1} = \mathcal{S}_{l,b_l}{}^{v_1}(x_l^k)$$

Step 2 (Coarse grid correction)

Residual : $r_l = b_l - A_l x_l^{k,1}$

Restriction : $r_{l-1} = I_l{}^{l-1} r_l$

Recursive call:

$$e_{l-1}^0 = 0$$
for $i = 1 \dots \mu$
$$e_{l-1}^i = MGM(e_{l-1}^{i-1}, r_{l-1}, l-1)$$
$$e_{l-1} = e_{l-1}^{\mu}$$
Prolongation : $e_l = I_{l-1}^l e_{l-1}$

Correction : $x_l^{k,2} = x_l^{k,1} + e_l$

Step 3 (v_2-post–smoothing)

$$MGM(x_l^k, b_l, l) = \mathcal{S}_{l,b_l}^{v_2}(x_l^{k,2})$$

The algorithm $\mu = 1$ is called V-cycle (see Figure 7). The algorithm $\mu = 2$ is called W-cycle (see Figure 8).
To obtain a good start approximation for a multigrid algorithm, we apply the F-cycle (see Figure 9).

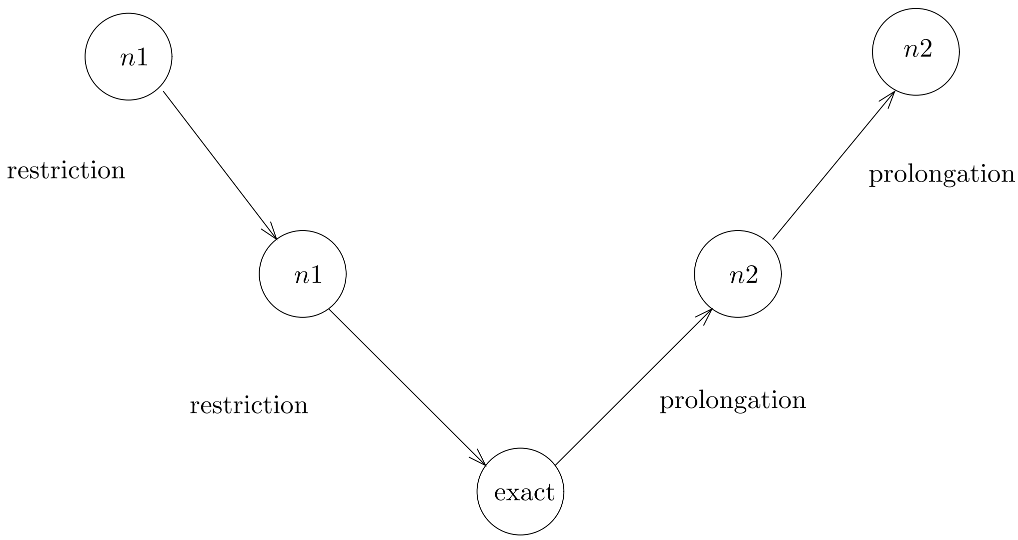


Figure 7: V-cycle

Homework: Describe the multigrid algorithm as a finite state machine, where every state is smoothing step and an operation is a restriction or prolongation. Then, the finite state machine of a V-cycle looks like a “V” and the finite state machine of a W-cycle looks like a “W”.
Let N be the number of unknowns. The computational amount of the V-cycle and W-cycle is $O(N)$.
The theory of multigrid algorithms shows that there is a constant ρ such that the convergence rate of the multigrid algorithm satisfies

$$\rho(C_{MGM,l}) \leq \rho < 1$$

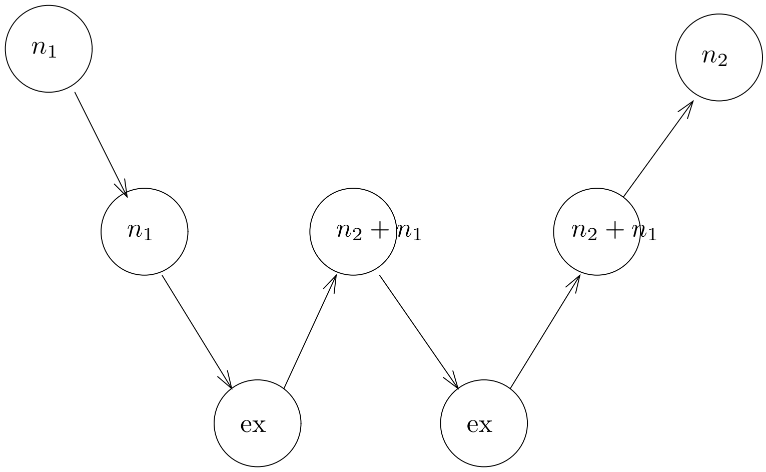


Figure 8: W-cycle

independent of l. This shows that the multigrid algorithm on a unit square for Poisson's equation is optimal with respect to the asymptotic computational amount.

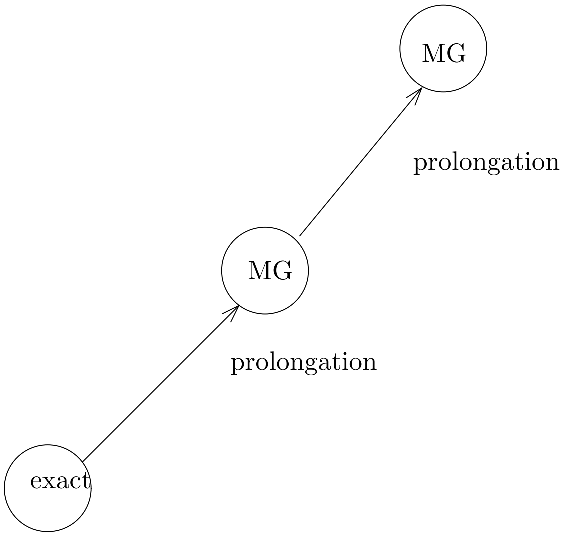


Figure 9: F-cycle

2.4 Local Mode Analysis of the Multigrid method

2.4.1 1D Model Problem

The local mode analysis is a method to analyze the convergence rate of a multigrid method. It is not an exact mathematical analysis of the the multigrid method, but an analysis which gives a rather good hint about the convergence properties of a multigrid method. To explain this problem, let us consider the following Poisson's equation in 1D:

Problem 2. *Let* $f \in \mathcal{C}([0,1])$*. Find* $u \in cC^2([0,1])$ *such that*

$$-\triangle u = f \quad on \quad [0,1].$$

To solve this problem let us apply the multigrid method with damped Jacobi iteration as a smoother.
But, obviously, the local mode analysis can also be applied for more complicated PDE's and multigrid algorithms in 2D and 3D. But the local mode analysis cannot be applied for every multigrid algorithm as FE-discretizations on unstructured grids.

2.4.2 Extension of Operators

A multigrid algorithm consists of several parameters that have to be properly tuned such that the algorithm converges rapidly.The parameters are,

μ	:	recursion parameter.
ν_1, ν_2	:	smoothing parameter.
$\mathcal{S}_{l,b_l}$	:	choice of smoother.
I_{l-1}^{l}	:	choice of the prolongation operator.
I_{l}^{l-1}	:	choice of the restriction operator
A_l for $l < l_{max}$	:	choice of the stiffness matrix on the courser grid.
		($A_{l_{max}}$is determined by the discretisation.)

To simplify the analysis of the convergence of the two–grid method we omit the boundary conditions and study all operators on an infinite dimensional grid!

Instead of the finite grid

$$\Omega_h^d := \left\{ (j_1 h, j_2 h, \ldots, j_d h) \quad | \quad j_1, j_2, \ldots, j_d \in \left\{ 0, \ldots, \frac{1}{h} \right\} \right\} \tag{43}$$

we apply an infinite grid

$$\overset{\infty}{\Omega}{}_h^d := \{ (j_1 h, j_2 h, \ldots, j_d h) \quad | \quad j_1, j_2, \ldots, j_d \in \mathbb{Z} \} \tag{44}$$

The operators A_l, I_l^{l-1}, $\mathcal{S}_{l,b_l}$ have to be extended to the infinite dimensional grid in a suitable manner.

Remark

- The operators A_l etc. are stencil operators, e.g a nine point stencil.
- The operators A_l etc. depend on the spatial coordinates.

Therefore, we define the operators on the infinite grid as follows:

> Let Q_h^d be a stencil operator on the grid Ω_h^d. Furthermore, let x_0 be an interior point of the grid Ω_h^d. Now, define the stencil operator $\overset{\infty}{Q}{}_h^d$ to be the operator with stencil $\mathcal{S}(x_0)$ for every grid point $x \in \overset{\infty}{\Omega}{}_h^d$.

Example

Let $d = 1$. The stiffness matrix obtained by the finite difference discretization of the operator $-\frac{d^2}{dx^2}$ on the grid Ω_h^1 is

$$\mathbf{A_h^1} = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \frac{1}{h^2} \tag{45}$$

Now, the operator on the corresponding infinite grid $\overset{\infty}{A}{}_h^1$ is:

$$\overset{\infty}{A}{}_h^1 = \begin{pmatrix} \ddots & & \ddots & & \ddots & \\ & -1 & & 2 & & -1 \\ & & \ddots & & \ddots & & \ddots \end{pmatrix} \frac{1}{h^2} \tag{46}$$

which implies

$$\overset{\infty}{A}{}_h^1(u)(x) = (-u(x-h) + 2u(x) - u(x+h)) \frac{1}{h^2} \qquad \forall x \in \overset{\infty}{\Omega}{}_h^1 \tag{47}$$

By the extension of the above operators on the infinite dimensional grid, we can construct a two–grid method on the infinite dimensional grid $\overset{\infty}{\Omega}{}_h^d$. To analyze the convergence of the two–grid method, we need to know the iteration matrix of the method. By Lemma 3, the iteration matrix for the two–grid method is

$$\left(C_h^{relax}\right)^{\nu_2} \left(E_h - I_H^h (A_H)^{-1} I_h^H A_h\right) \left(C_h^{relax}\right)^{\nu_1}, \text{where } H = 2h. \tag{48}$$

where,

C_h^{relax}	:	iteration matrix of the smoothening step.
E_h	:	extended unit matrix.
I_H^h	:	extended prolongation operator.
I_h^H	:	extended Restriction operator.
A_h, A_H	:	extended stiffness matrices on the coarser grid.

For reasons of simplicity, let us write A_h instead of $\overset{\infty}{A}{}_h$.

Example: Operators for the model problem

The operators for d=1 are as follows.

$$A_h = \begin{pmatrix} \ddots & \ddots & \ddots & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \end{pmatrix} \frac{1}{h^2} \tag{49}$$

$$A_H = \begin{pmatrix} \ddots & \ddots & \ddots & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \end{pmatrix} \frac{1}{4h^2} \tag{50}$$

$$I_h^H = \begin{pmatrix} \ddots & & & & & & \\ & 1 & 2 & 1 & & & \\ & & & 1 & 2 & 1 & \\ & & & & & & \ddots \end{pmatrix} \frac{1}{4} \qquad \left(or \;\; factor \;\; \frac{1}{2\sqrt{2}}\right) \tag{51}$$

$$I_H^h = \begin{pmatrix} \ddots & & & \\ & 1 & & \\ & 2 & & \\ & 1 & 1 & \\ & & 2 & \\ & & 1 & \\ & & & \ddots \end{pmatrix} \frac{1}{2} \qquad \left(or \;\; factor \;\; \frac{1}{2\sqrt{2}}\right) \tag{52}$$

$$\begin{aligned} C_h^{relax} &= \begin{pmatrix} \ddots & \ddots & \ddots & \\ \frac{1}{2}\omega & 1-\omega & \frac{1}{2}\omega & \\ & \ddots & \ddots & \ddots \end{pmatrix} \\ &\overset{\omega=\frac{1}{2}}{=} \begin{pmatrix} \ddots & \ddots & \ddots & \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \\ & \ddots & \ddots & \ddots \end{pmatrix} \end{aligned} \tag{53}$$

We allow these operators to act on the following functional spaces.

$$V_h := \operatorname{span}\left\{ exp\left(i\theta\frac{x}{h}\right)_{x\in \overset{\infty}{\Omega}{}^{d}_{h}} \;\; \middle| \;\; -\pi \le \theta \le \pi \right\} \tag{54}$$

$$V_H := \operatorname{span}\left\{ exp\left(i\theta\frac{x}{H}\right)_{x\in \overset{\infty}{\Omega}{}^{d}_{H}} \;\; \middle| \;\; -\pi \le \theta \le \pi \right\} \tag{55}$$

For reasons of simplicity, let us restrict ourselves to the 1–D case. The harmonic frequency of $\exp\left(i\theta\frac{x}{h}\right)$ is $\exp\left(i\tilde{\theta}\frac{x}{h}\right)$ where,

$$\begin{aligned} \tilde{\theta} &:= \theta - \pi \quad \text{for} \quad \theta \geq 0 \\ \tilde{\theta} &:= \theta + \pi \quad \text{for} \quad \theta < 0 \end{aligned}$$

2.4.3 Local Mode Analysis of the Smoother

Definition 5. *Let us assume that the functions* $\exp\left(i\theta\frac{x}{h}\right)$ *are the eigenfunctions of the iteration matrix* C *of the smoother* $\mathcal{S}$*with eigenvalues* $\mu(\theta)$*. This means*

$$C \exp\left(i\theta\frac{x}{h}\right) = \mu(\theta) \exp\left(i\theta\frac{x}{h}\right)$$

Then, let us define the smoothening factor of $\mathcal{S}$ *by*

$$\bar{\mu} := \max_{\frac{\pi}{2} \leq |\Theta| \leq \pi} |\mu(\Theta)|$$

Figure 10 depicts the local mode analysis of the Jacobi smoother with relaxation parameter $\omega = \frac{1}{2}$ and $\omega = 1$ for the 1D model problem. The corresponding smoothing factors are 0.5 and 1. This shows that the Jacobi iteration without relaxation ($\omega = 1$) is not suitable for a multigrid method.

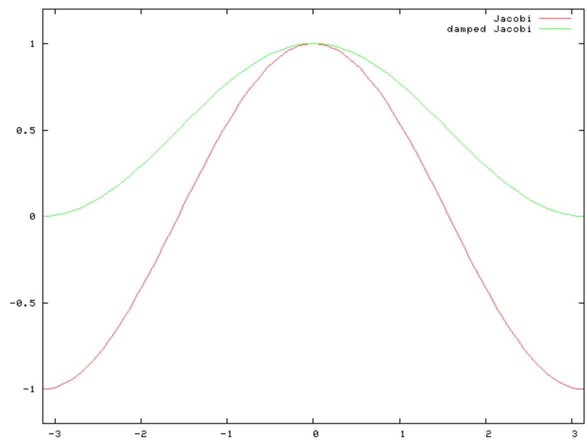


Figure 10: Local mode analysis of the Jacobi iteration

2.4.4 Local Mode Analysis of the Restriction and Prolongation

The local mode analysis of the restriction of the 1D-model problem shows (see Figure 11)

$$I_h^H \exp(i\Theta\frac{x}{h}) = \cos^2(\frac{\Theta}{2}) \exp(i2\Theta\frac{x}{H}).$$

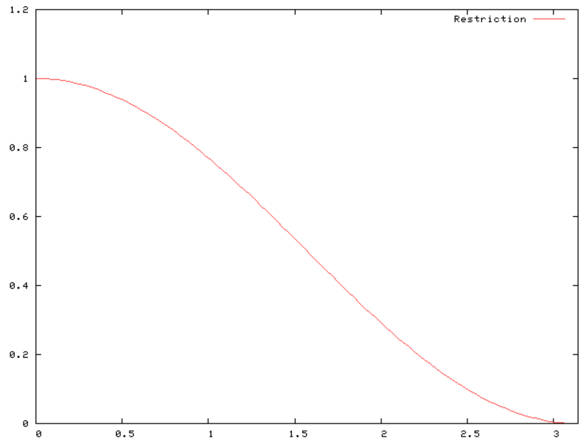


Figure 11: Local mode analysis of the restriction

The local mode analysis of the prolongation of the 1D-model problem shows (see Figure 12)

$$I_H^h \exp(i\Theta \frac{x}{H}) = \cos^2(\frac{\Theta}{4}) \exp(i \frac{\Theta}{2} \frac{x}{h}) + \sin^2(\frac{\Theta}{4}) \exp(i \widetilde{\left(\frac{\Theta}{2}\right)} \frac{x}{h}). \qquad (56)$$

To prove Equation (56), observe that

$$\begin{aligned} I_H^h \exp(i\Theta \frac{x}{H}) &= \exp(i\tfrac{\Theta}{2}\tfrac{x}{h}) && \text{if } x = h2k \in \Omega_H \\ I_H^h \exp(i\Theta \frac{x}{H}) &= \tfrac{1}{2}\left(\exp(i\Theta\tfrac{x-h}{H}) + \exp(i\Theta\tfrac{x+h}{H})\right) && \\ &= \exp(i\tfrac{\Theta}{2}\tfrac{x}{h})\cos(\tfrac{\Theta}{2}) && \text{if } x = h(2k+1) \in \Omega_H + h \end{aligned}$$

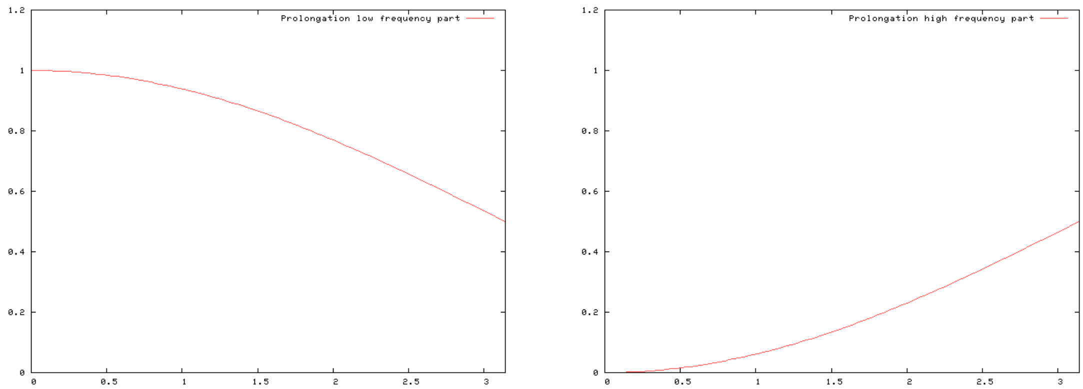
Furthermore, observe that

$$\begin{aligned} \exp(i\widetilde{\left(\frac{\Theta}{2}\right)}\frac{x}{h}) &= \exp(i\left(\tfrac{\Theta}{2} - \pi\right)2k) && \\ &= \exp(i\tfrac{\Theta}{2}\tfrac{x}{h}) && \text{if } x = h2k \in \Omega_H \\ \exp(i\widetilde{\left(\frac{\Theta}{2}\right)}\frac{x}{h}) &= \exp(i\left(\tfrac{\Theta}{2} - \pi\right)(2k+1)) && \\ &= -\exp(i\tfrac{\Theta}{2}\tfrac{x}{h}) && \text{if } x = h(2k+1) \in \Omega_H + h \end{aligned}$$

Using the formulas

$$\begin{aligned} \cos^2(\phi) + \sin^2(\phi) &= 1 \\ \cos^2(\phi) - \sin^2(\phi) &= \cos(2\phi) \end{aligned}$$

completes the prove of (56).



low frequency part   high frequency part

Figure 12: Local mode analysis of the prolongation

2.4.5 Local Mode Analysis of the Two-Grid-Algorithm

Let consider the 1D model problem. The the two-grid iteration matrix is:

$$C_h^{two-grid}\left(C_h^{relax}\right)^{\nu_2}\left(E_h - I_H^h(A_H)^{-1}I_h^H A_h\right)\left(C_h^{relax}\right)^{\nu_1} \tag{57}$$

The local mode analysis of the two-grid iteration can be described by a matrix. In case of the 1D model problem, this is a 2×2 matrix $M(\Theta) = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$ such that:

$$\begin{aligned} C_h^{two-grid}(\exp(i\Theta\frac{x}{h})) &= m_{11}\exp(i\Theta\frac{x}{h}) + m_{21}\exp(i\tilde{\Theta}\frac{x}{h}) \\ C_h^{two-grid}(\exp(i\tilde{\Theta}\frac{x}{h})) &= m_{21}\exp(i\Theta\frac{x}{h}) + m_{22}\exp(i\tilde{\Theta}\frac{x}{h}). \end{aligned}$$

The matrix $M(\Theta)$ is called two grid amplification matrix.

Definition 6. *The asymptotic two-grid convergence rate is*

$$\bar{\lambda} := \rho(C_h^{two-grid}) = \max_{|\Theta|\leq\frac{\pi}{2}} \rho(M(\Theta)).$$

In case of the 1D model problem, we obtain

$$M(\Theta) = \begin{pmatrix} sc^{\nu_1+\nu_2} & -cc^{\nu_1}s^{\nu_2} \\ -ss^{\nu_2}c^{\nu_1} & cs^{\nu_1+\nu_2} \end{pmatrix},$$

where $s = \sin^2(\frac{\Theta}{2})$ and $c = \cos^2(\frac{\Theta}{2})$. A short calculation shows

$$\rho(M(\Theta)) = \frac{p+q+\sqrt{(p+q)^2+8pq}}{2},$$

where $p = cs^{\nu_1+\nu_2}$ and $q = sc^{\nu_1+\nu_2}$.

Example 3. $\nu_1 + \nu_2 = \nu = 2$. *Then, we get*

$$\rho(M(\Theta)) = \max_{|\Theta| \leq \frac{\pi}{2}} \sin^2(\Theta)(1 + \sqrt{1 + 2\sin^2(\Theta)})$$

A numerical calculation shows the asymptotic two-grid convergence rate is

$$\bar{\lambda} \approx 0.3415$$

Since the smoothening factor is $\bar{\mu} = 0.5$ *we obtain*

$$\bar{\mu}^2 < \bar{\lambda} < \bar{\mu}.$$

Additionally, one can show that

$$\bar{\lambda}_{\nu=2}^{\frac{3}{2}} < \bar{\lambda}_{\nu=3}$$

This shows, that the choice $\nu = 2$ *is an optimal choice.*

2.5 Multigrid Algorithm for Finite Elements

2.5.1 Sequence of Subgrids and Subspaces

Let $\tau_{h_1} \cdots, \tau_{h_{l_{max}}}$ be a sequence of quasi-uniform subdivisions of a polygon domain Ω, where $h_l = 2^{-l}$. Let

- $\mathring{\Omega}_{h_l} = \mathring{\Omega}_l$ be the set of interior grid points of τ_h.
- $\Omega_{h_l} = \Omega_l$ be the set of all grid points of τ_h.

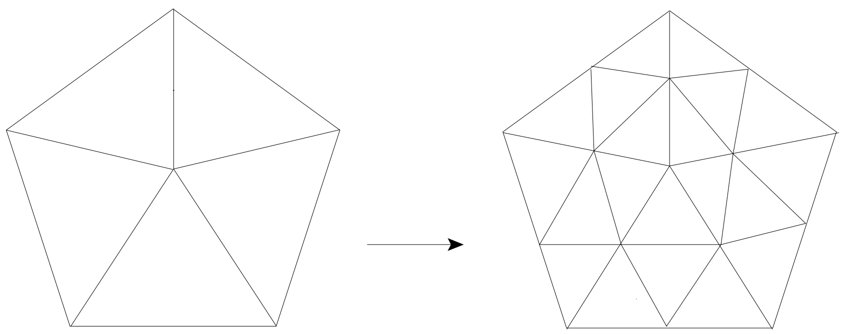
Obviously,

$$\Omega_{l-1} \subset \Omega_l.$$

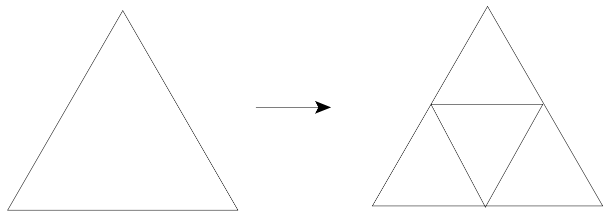
Furthermore, let us assume, that V_{h_l} is a finite element space on the grid $\Omega_{h_l} = \Omega_l$ such that

$$V_{h_l} \subset V_{h_{l+1}} \quad \text{(This means} \quad V_{2h} \subset V_h).$$

Example 4. *Let us assume that* τ_{h_l} *is a triangulation. Then, let* $V_{h_l} \subset H^1(\Omega)$ *be the finite element space of linear finite elements on* τ_{h_l}.



Every triangle is divided into four triangles



Let $a(u,v)$ be a symmetric positive definite bilinear form on $V_{h_{l_{max}}}$.
Furthermore, let $f \in V'_{h_{l_{max}}}$.

Example 5. *An example is*

$$a(u,v) = \int_\Omega \nabla u \nabla v + uv \; d(x,y)$$

and $f(v) = \int_\Omega \tilde{f} v d$.

We want to solve the problem

Find $u_{h_{l_{max}}} \in V_{h_{l_{max}}}$ such that

$$a(u_{h_{l_{max}}}, v_h) = f(v_h) \quad \forall \quad v_h \in V_{h_{l_{max}}}. \tag{58}$$

(In case of the above example, this problem is equivalent to
$-\triangle u + u = \tilde{f}, \; \frac{\partial u}{\partial \vec{n}}\big|_{\partial\Omega} = 0.$)
To this end, let us study the problems

Find $u_{h_l} \in V_{h_l}$ such that

$$a(u_{h_l}, v_h) = f_l(v_h) \quad \forall \quad v_h \in V_{h_l} \tag{59}$$

for every $l = 0, \cdots, l_{max}$

where f_l is a suitable coarse grid right hand side.
Remark: In case of Dirichlet boundary conditions, one has to replace the space V_{h_l} by the space $\overset{\circ}{V}_{h_l} := V_{h_l} \cap H^1_0(\Omega)$.

2.5.2 The Nodal Basis

Let $(v_l^k)_{k \in \Omega_{h_l}}$ be the nodal basis for V_{h_l}.
(In case of Dirichlet boundary conditions consider $(v_l^k)_{k \in \overset{\circ}{\Omega}_{h_l}}$)
Now (59) can be defined in matrix form as follows:

$$A_l x_l = b_l \tag{60}$$

where

$$A_l = (a_{kj})_{kj\in\Omega_{h_l}}, a_{kj} = a(v_l^k, v_l^j) \quad (61)$$

$$x_l = (x_l^k)_{k\in\Omega_{h_l}} \quad (62)$$

$$b_l = (b_l^k)_{k\in\Omega_{h_l}} \quad (63)$$

and the solution vector u_{h_l} is given by

$$u_{h_l} = \sum k \in \Omega_{h_l} x_l^k v_l^k \quad (64)$$

2.5.3 Prolongation Operator for Finite Elements

The natural inclusion is the prolongation operator

$$\begin{array}{lll} u & \in & V_{h_i} \\ \downarrow & & \\ u & \in & V_{h_{i+1}} \end{array}$$

To implement this operator, we have to describe this operator in a matrix form.

By $V_{h_i} \subset V_{h_{i+1}}$, there are coefficients $\gamma_k^{k'}$ such that

$$v_i^{k'} = \sum_k \gamma_k^{k'} v_{i+1}^k \quad (65)$$

Thus, we get

$$u_{h_i} = \sum_{k'} x_i^{k'} v_i^{k'} = \sum_{k'} \sum_k \gamma_k^{k'} v_{i+1}^k x_i^{k'} \quad (66)$$

$$= \sum_k \left(\sum_{k'} \gamma_k^{k'} x_i^{k'} \right) v_{i+1}^k \quad (67)$$

Now the matrix version of the prolongation operator is

$$I_i^{i+1} \left(x_i^{k'} \right)_{k'} = \left(\sum_{k'} \gamma_k^{k'} x_i^{k'} \right)_k$$

$$\Downarrow$$

$$I_i^{i+1} = (\gamma_k^{k'})_{(k,k')}$$

2.5.4 Restriction Operator for Finite Elements

Observe that $F_i \in (V_{h_i})'$.
This means that $F_i : V_{h_i} \longrightarrow \mathbf{R}$ is a linear mapping. The natural inclusion is the restriction operator.

$$\begin{array}{rcl} F_{i+1} & \in & (V_{h_{i+1}})' \\ \downarrow & & \\ F_i & \in & (V_{h_i})' \\ F_i(w) & := & F_{i+1}(w) \quad \forall \quad w \in V_{h_i} \end{array}$$

The matrix version of the restriction operator can be obtained as follows

$$b_i^{k'} = F_i(v_i^{k'}) = \sum_k \gamma_k^{k'} F_i(v_{i+1}^k) \tag{68}$$

$$= \sum_k \gamma_k^{k'} b_{i+1}^k \tag{69}$$

$$I_{i+1}^i = \left(\gamma_k^{k'}\right)_{(k',k)} \tag{70}$$

3 Subspace Correction Methods

3.1 Multiplicative Subspace Correction Methods

Let us assume that V is a finite dimensional Hilbert space and that

$$a : V \times V \to \mathbb{R}$$

is a positive definite bilinear form. Furthermore, assume that $f \in V'$.

Problem 3. *Find $u \in V$ such that*

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

We want to find an iterative method to solve this problem. To this end, let $V_1, ..., V_m$ be subspaces of V such that

$$V = \sum_{i=1}^{m} V_i.$$

Definition 7. *A correction in the direction of the subspace V_i is defined as follows. Let u_{old} be an old approximation of 5. Then, let $u_{new} = u_{old} + w$ be the solution of*

$$a(u_{old} + w, v) = f(v) \quad \forall v \in V_i$$

such that $w \in V_i$. Define

$$u_{new} = \mathcal{S}_{V_i}(u_{old})$$

A multiplicative subspace correction for solving Problem 5 is

$$\mathcal{S}_{V_1} \circ \mathcal{S}_{V_2} \circ \ldots \circ \mathcal{S}_{V_{l_{\max}}}.$$

Example 6 (Gauss-Seidel Iteration)**.** *Consider the space V_h of bilinear finite elements on a grid of size h. Color the points according to Figure 13. Define the spaces spanned by the nodal points corresponding to these colors by $V_{r,h}$, $V_{b,h}$, $V_{g,h}$, $V_{y,h}$. Then,*

$$\mathcal{S}_{V_{r,h}} \circ \mathcal{S}_{V_{b,h}} \circ \mathcal{S}_{V_{g,h}} \circ \mathcal{S}_{V_{y,h}}$$

is the classical Gauss-Seidel iteration. Observe that

$$V_h = V_{r,h} \oplus V_{b,h} \oplus V_{g,h} \oplus V_{y,h}.$$

is a direct sum.

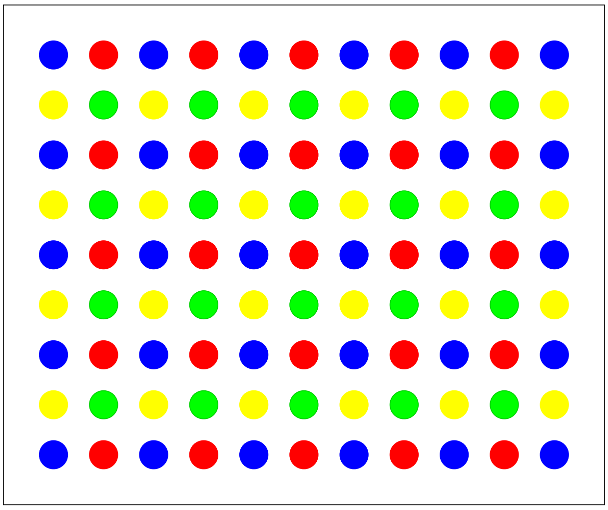


Figure 13: Four colors of Gauss-Seidel iteration

Example 7 (Classical Multigrid Algorithm). *Construct the spaces V_h and $V_{r,h}$, $V_{b,h}$, $V_{g,h}$, $V_{y,h}$ according to Example 6.*
Then, the multigrid algorithm can be described as a subspace correction method. For example the V-cycle with one pre-smoothing is

$$\begin{array}{rcl} \mathcal{S}_{V_{r,h_1}} & \circ & \mathcal{S}_{V_{b,h_1}} \circ \mathcal{S}_{V_{g,h_1}} \circ \mathcal{S}_{V_{y,h_1}} \circ \\ \mathcal{S}_{V_{r,h_2}} & \circ & \mathcal{S}_{V_{b,h_2}} \circ \mathcal{S}_{V_{g,h_2}} \circ \mathcal{S}_{V_{y,h_2}} \circ \\ & \circ & \ldots \\ & \circ & \ldots \\ \mathcal{S}_{V_{r,h_{l_{\max}}}} & \circ & \mathcal{S}_{V_{b,h_{l_{\max}}}} \circ \mathcal{S}_{V_{g,h_{l_{\max}}}} \circ \mathcal{S}_{V_{y,h_{l_{\max}}}} \end{array}$$

The general multigrid algorithm can be described as follows:

Multigrid algorithm as a subspace correction method

If $l = 1$ then perform Gauss-Seidel iterations:
$MGM(u_l^k, f_l, l) = (\mathcal{S}_{V_{r,h}} \circ \mathcal{S}_{V_{b,h}} \circ \mathcal{S}_{V_{g,h}} \circ \mathcal{S}_{V_{y,h}})^{\nu_1+\nu_2}(u_l^k)$

If $l > 1$ then

Step 1 (ν_1-pre–smoothing)
$u_l^{k,1} = (\mathcal{S}_{V_{r,h}} \circ \mathcal{S}_{V_{b,h}} \circ \mathcal{S}_{V_{g,h}} \circ \mathcal{S}_{V_{y,h}})^{\nu_1}(u_l^k)$

Step 2 (Coarse grid correction)

Define: Residual : $f_{l-1}(v) := f_l(v) - a(u_l^{k,1}, v)$ *for all* $v \in V_{l-1}$.
Recursive call:
$e_{l-1}^0 = 0$
for $i = 1 \ldots \mu$
$e_{l-1}^i = MGM(e_{l-1}^{i-1}, f_{l-1}, l-1)$
$e_{l-1} = e_{l-1}^{\mu}$
Prolongation : $e_l = I_{l-1}^l e_{l-1}$
Correction : $u_l^{k,2} = u_l^{k,1} + e_l$

Step 3 (ν_2-post–smoothing)

$MGM(x_l^k, b_l, l) = (\mathcal{S}_{V_{r,h}} \circ \mathcal{S}_{V_{b,h}} \circ \mathcal{S}_{V_{g,h}} \circ \mathcal{S}_{V_{y,h}})^{\nu_2}(u_l^k)$

Example 8 (Multigrid Algorithm with Relaxation on a Complementary Space). *Construct the spaces V_h and $V_{r,h}$, $V_{b,h}$, $V_{g,h}$, $V_{y,h}$ according to Example 6 such that*

$$V_{g,h_l} = V_{h_{l-1}}.$$

Then, the spaces

$$\begin{array}{rcl} V_{g,h_1} & \oplus & \\ V_{r,h_1} & \oplus & V_{b,h_1} \oplus V_{y,h_1} \oplus \\ V_{r,h_2} & \oplus & V_{b,h_2} \oplus V_{y,h_2} \\ & \oplus & ... \\ & \oplus & ... \\ V_{r,h_{l_{\max}}} & \oplus & V_{b,h_{l_{\max}}} \oplus V_{y,h_{l_{\max}}}. \end{array}$$

form a direct sum. The corresponding subspace correction method is the V-cycle with one pre-smoothing and without relaxation at the coarser grid points, but with a relaxation on a complementary space. The complementary spaces are

$$W_l := V_{r,h_l} \oplus V_{b,h_l} \oplus V_{y,h_l}.$$

Example 9 (Multigrid Algorithm on a Complementary Space)**.** *Define the spaces $V_{h_l} =: V_l$ and W_l according to Example 8. Then,*

$$V_l = V_{l-1} \oplus W_l$$

is a direct sum. The subspace correction method corresponding to this construction is:

$$\mathcal{S}_{V_1} \circ \mathcal{S}_{W_2} \circ \mathcal{S}_{W_3} \circ ... \circ \mathcal{S}_{W_{l_{\max}}}.$$

In Section 4, we will see that the convergence rate of this multigrid algorithm depends on the angle between W_l and V_{l-1}.

3.2 Multigrid Algorithm with Hierarchical Surplus

The classical multigrid requires one storage for every multigrid level at every grid point for each variable. This means for every variable

- 1 storage at the grid points $\Omega_{l_{\max}} \backslash \Omega_{l_{\max}-1}$.
- 2 storages at the grid points $\Omega_{l_{\max}} \backslash \Omega_{l_{\max}-1} \backslash \Omega_{l_{\max}-2}$.
- 3 storages at the grid points $\Omega_{l_{\max}} \backslash \Omega_{l_{\max}-1} \backslash \Omega_{l_{\max}-2} \backslash \Omega_{l_{\max}-3}$.
- ...

Using a "hierarchical surplus", one can implement a multigrid algorithm such that $O(1)$ storages are needed at every grid point. There are two advantages of this approach:

- Extension of the classical multigrid algorithm for non-linear PDE's.

- Implementation of multigrid algorithms on adaptive grids.

To explain this kind of multigrid algorithm, let us define the interpolation operator I_l by

$$\begin{aligned} I_l : V_{l_{\max}} &\rightarrow V_l \\ I_l(u)(x) &= u(x) \quad \forall x \in \Omega_{h_l} \end{aligned}$$

The hierarchical surplus is defined as

$$\begin{aligned} H_l : V_l &\rightarrow V_l \\ H_l(u) &= u - I_{l-1}(u). \end{aligned}$$

Observe that

$$I_l(u)(x) = 0 \quad \forall x \in \Omega_{h_{l-1}}$$

This implies, that we can store

$$H_l(u_l), \quad \text{for } l = 1, ..., l_{\max} \ .$$

by *only 1 storage* for every grid point.

Multigrid algorithm with Hierarchical Surplus (here only V-cycle)

If $l = 1$ then $MGM(u_1^k, f_1, 1) = u_1^{k,3} = u_{h_1}$

If $l > 1$ then

Step 1 (ν_1-pre–smoothing)

$u_l^{k,1} = \mathcal{S}_{l,f_l}^{\,v_1}(u_l^k)$

Step 2 (Coarse grid correction)

Store hierarchical surplus: $w_l := H(u_l^{k,1})$.
Coarse right hand side:

$$f_{l-1}(v) := f_l(v) - a(w_l, v) \quad \forall v \in V_{l-1}.$$

Recursive call: $u_{l-1}^{k,3} = MGM(u_l^{k,1}, f_{l-1}, l-1)$
Correction : $u_l^{k,2} = u_{l-1}^{k,3} + w_l$

Step 3 (ν_2-post–smoothing)

$MGM(x_l^k, f_l, l) = u_l^{k,3} = \mathcal{S}_{l,b_l}^{\,v_2}(u_l^{k,2})$

In this algorithm, the variables $u_l^{k,i}$ and w_l can be stored by *only 1 storage* for every grid point.

3.3 A Multigrid Algorithm for Non-Linear Problems

In this section, we explain a multigrid algorithm for non-linear problems as an extension of a subspace correction method. The multigrid algorithm is equivalent to the full approximation scheme in [6].
Let us assume that

$$\begin{aligned} a : V \times V \times V &\rightarrow \mathbb{R} \\ (w; u, v) &\mapsto a(w; u, v). \end{aligned}$$

is a function such that

$$(u, v) \mapsto a(w; u, v)$$

is a positive definite bilinear form for every $w \in V$. We want to solve the problem:

Problem 4. *Find $u \in V$ such that*

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

Assumptions to guarantee existence of a solution of this problem are described in [14]

Example 10. *The thermal conductivity of certain materials depends on the temperature. This property of the material can be modeled by the following "non-linear" form:*

$$(w; u, v) \mapsto a(w; u, v) = \int_\Omega \nabla u \left(1 + \frac{1}{w^2 + 1}\right) \nabla v \; d(x, y)$$

On every grid Ω_l, we can define a coarse grid equation. Let $u_l \in V_l$ be the solution of

$$a(u_l; u_l, v_l) = f_l(v_l) \quad \forall v_l \in V_l.$$

To derive a multigrid algorithm for a non-linear problem, let us first consider a two-grid problem. Let u_l^{old} be an approximation on the fine grid. By a coarse grid correction, we want to obtain a new approximation u_l^{new}. To this end, we want to find an approximation $\hat{e}_{l-1} \in V_{l-1}$ of the exact coarse grid correction $e_{l-1} \in V_{l-1}$, which is defined by:

$$a(u_l^{\text{old}} + e_{l-1}; u_l^{\text{old}} + e_{l-1}, v_{l-1}) = f_l(v_{l-1}) \quad \forall v_{l-1} \in V_{l-1}.$$

This coarse grid equation, which defines $\hat{e}_{l-1}$ must satisfy two conditions

- If $e_{l-1} = 0$, then there exists a $\hat{e}_{l-1}$ such that $\hat{e}_{l-1} = 0$.
- The term u_{l-1}^{new} in the non-linear form $a(u_{l-1}^{\text{new}}; ...)$ must be a coarse grid approximation of u_l.

If e_{l-1} is small, then an approximation of e_{l-1} can be found by:

$$a(u_l^{\text{old}}; u_l^{\text{old}} + \tilde{e}_{l-1}, v_{l-1}) = f_l(v_{l-1}) \quad \forall v_{l-1} \in V_{l-1},$$

where $\tilde{e}_{l-1} \in V_{l-1}$. This equation is equivalent to

$$a(u_l^{\text{old}}; \tilde{e}_{l-1}, v_{l-1}) = f_l(v_{l-1}) - a(u_l^{\text{old}}; u_l^{\text{old}}, v_{l-1}) =: r(v_{l-1}) \quad \forall v_{l-1} \in V_{l-1}.$$

Decompose u_l^{old} by

$$u_l^{\text{old}} = w_l + I_l(u_l^{\text{old}}) = w_l + u_{l-1}^{\text{old}}.$$

Then, an approximation of the above equation is $\tilde{\tilde{e}}_{l-1} \in V_{l-1}$

$$a(u_{l-1}^{\text{old}}; \tilde{\tilde{e}}_{l-1}, v_{l-1}) = r(v_{l-1}) \quad \forall v_{l-1} \in V_{l-1}$$

and equivalent to this equation

$$\begin{aligned} a(u_{l-1}^{\text{old}}; \tilde{\tilde{e}}_{l-1} + u_{l-1}^{\text{old}}, v_{l-1}) &= r(v_{l-1}) + a(u_{l-1}^{\text{old}}; u_{l-1}^{\text{old}}, v_{l-1}) \\ &=: f_{l-1}(v_{l-1}) \quad \forall v_{l-1} \in V_{l-1}. \end{aligned}$$

Thus, we can define the following coarse grid equation

$$a(u_{l-1}^{\text{new}}; u_{l-1}^{\text{new}}, v_{l-1}) = f_{l-1}(v_{l-1}) \quad \forall v_{l-1} \in V_{l-1},$$

where u_{l-1}^{new} is an approximation of $\tilde{\tilde{e}}_{l-1} + u_{l-1}^{\text{old}}$. Now, define

$$\hat{e}_{l-1} := u_{l-1}^{\text{new}} - u_{l-1}^{\text{old}}.$$

One can see that the above coarse grid equation describes the non-linearity of the equation. Furthermore, the following lemma holds:

Lemma 1. *If* $e_{l-1} = 0$*, then* $\hat{e}_{l-1} = 0$.

Proof. If $e_{l-1} = 0$, then $r_{l-1} = 0$. By the positive definiteness of $a(u_{l-1}^{\text{old}}; \cdot, \cdot)$, we get $\tilde{\tilde{e}}_{l-1} = 0$. This implies $u_{l-1}^{\text{new}} = u_{l-1}^{\text{old}}$. Thus, $\hat{e}_{l-1} = 0$. □

From the above equations we see that

$$\begin{aligned} f(v_{l-1}) &= r(v_{l-1}) + a(u_{l-1}^{\text{old}}; u_{l-1}^{\text{old}}, v_{l-1}) \\ &= f(v_{l-1}) - a(u_l^{\text{old}}; u_l^{\text{old}}, v_{l-1}) + a(u_{l-1}^{\text{old}}; u_{l-1}^{\text{old}}, v_{l-1}) \end{aligned}$$

and

$$\begin{aligned} u_l^{\text{new}} &= u_l^{\text{old}} + u_{l-1}^{\text{new}} - u_{l-1}^{\text{old}} \\ &= u_{l-1}^{\text{new}} + H_l(u_l^{\text{old}}). \end{aligned}$$

Multigrid algorithm for non-linear problems
(here only V-cycle)

If $l = 1$ then $MGM(u_1^k, f_1, 1) = u_1^{k,3} = u_{h_1}$

If $l > 1$ then

Step 1 (ν_1-pre–smoothing)

$$u_l^{k,1} = \mathcal{S}_{l,f_l}^{v_1}(u_l^k)$$

Step 2 (Coarse grid correction)

Store hierarchical surplus: $w_l := H(u_l^{k,1})$.
Coarse right hand side:

$$\begin{aligned} f_{l-1}(v_{l-1}) \quad := \quad & f_l(v_{l-1}) - \\ & a(u_l^{k,1}; u_l^{k,1}, v_{l-1}) + a(I_{l-1}(u_l^{k,1}); I_{l-1}(u_l^{k,1}), v_{l-1}) \\ & \quad \forall v_{l-1} \in V_{l-1}. \end{aligned}$$

Recursive call: $u_{l-1}^{k,3} = MGM(u_l^{k,1}, f_{l-1}, l-1)$
Correction : $u_l^{k,2} = u_{l-1}^{k,3} + w_l$

Step 3 (ν_2-post–smoothing)

$$MGM(x_l^k, f_l, l) = u_l^{k,3} = \mathcal{S}_{l,b_l}^{v_2}(u_l^{k,2})$$

Remark 1. *This algorithm coincides with the multigrid algorithm with hierarchical surplus in section 3.2, if a is a bilinear form. This means that a is independent of the first parameter:*

$$a(w_1; u, v) = a(w_2; u, v) \quad \forall w_1, w_2.$$

3.4 A Multigrid Algorithm on Adaptive Grids

In this section, we explain a multigrid algorithm on adaptive discretization grids.
Let us first explain an adaptive discretization for finite elements. To this end, let

$$V_{h_1} \subset V_{h_2} \subset ... \subset V_{h_{\max}}$$

be a sequence of finite element spaces V_{h_i} with respect to the meshsize h_i. Furthermore, let Ω_{h_i} be the discretization grid corresponding to V_{h_i} and $(v_p^{h_i})_{p\in\Omega_{h_i}}$ the set of nodal basis functions, such that

$$V_{h_i} = \text{span}\left\{ v_p^{h_i} \mid p \in \Omega_{h_i} \right\}$$

To obtain an adaptive discretization, choose a sequence

$$\Omega_1, \Omega_2, ..., \Omega_{\max}$$

such that

$$\Omega_l \subset \Omega_{h_l}.$$

Then, let us define the spaces

$$\begin{aligned} V_i &= \text{span}\left\{ v_p^{h_i} \mid p \in \Omega_i \right\} \\ V_{\text{adaptive}} &= \text{span} \bigcup_{i=1}^{\max} V_i. \end{aligned}$$

The adaptive discretization is defined by:

Problem 5 (Adaptive Discretization)**.** *Find* $u \in V_{l_{adaptive}}$ *such that*

$$a(u, v) = f(v) \quad \forall v \in V_{l_{adaptive}}.$$

An iterative solver for this linear equation system is the multigrid algorithm with hierarchical surplus in section 3.2. A Gauss-Seidel iteration can be constructed by the subspaces

$$V_{r,i} := V_{r,h_i} \cap V_i \quad V_{b,i} := V_{b,h_i} \cap V_i \quad V_{g,i} := V_{g,h_i} \cap V_i \quad V_{y,i} := V_{y,h_i} \cap V_i.$$

The difficulty in an efficient implementation of this multigrid algorithm is the implementation of

- the Gauss-Seidel relaxation (or the implementation of stencil operators) and
- the interpolation and restriction operators.

To avoid this problem, we construct a grid $\tilde{\Omega}_i \supset \Omega_i$, $\tilde{\Omega}_i \subset \Omega_{h_i}$, with hanging nodes and we permit only subgrids Ω_i with a certain *refinement property*. First, let us define the neighbor points $\mathcal{N}_i(p)$ on level i for a grid point $p \in \Omega_{h_i}$. $\mathcal{N}_i(p)$ is the set of points of Ω_{h_i}, which is needed to apply a stencil operator at the point p. Now, define

$$\tilde{\Omega}_i := \bigcup_{p \in \Omega_i} \mathcal{N}_i(p)$$

Using this grid $\tilde{\Omega}_i$, we can perform a Gauss-Seidel iteration on Ω_i by treating the points $\tilde{\Omega}_i \backslash \Omega_i$ as points with inhomogeneous Dirichlet boundary conditions.
For the implementation of efficient interpolation operators, we assume that the following *refinement property* holds:

Refinement Property
For every $p \in \tilde{\Omega}_i \backslash \Omega_{h_{i-1}}$ the following equation is satisfied:

$$\mathcal{N}_i(p) \cap \Omega_{h_{i-1}} \subset \Omega_{i-1}.$$

Figure 14 shows an adaptive grid with two levels and hanging nodes.

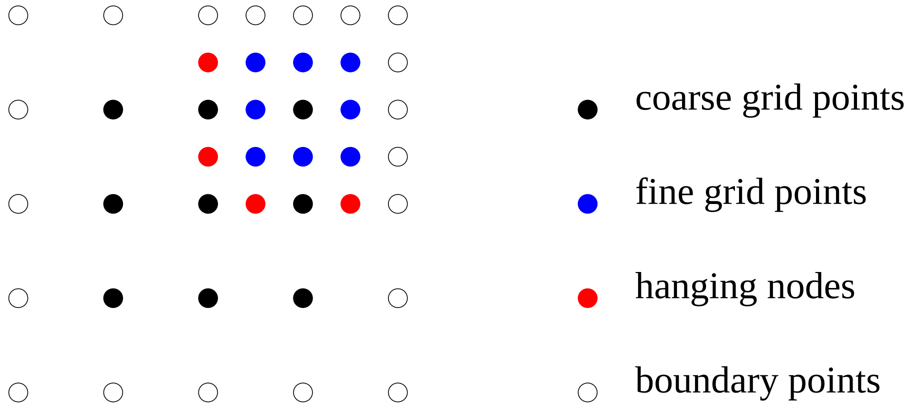


Figure 14: Adaptive grid.

4 Analysis of Multigrid Algorithms on a Complementary Space

See [3], [4], [11], [1] and [2] for further literature.

4.1 Analysis for a Symmetric Bilinear Form

Let $\mathcal{V}_1 \subset \mathcal{V}_2 \subset ... \subset \mathcal{V}_n$ be a sequence of vector spaces and let a be a symmetric positive bilinear form

$$a : \mathcal{V}_n \times \mathcal{V}_n \in \mathbb{R}.$$

Then, $\mathcal{V}_n$ is a Hilbert space with scalar product a, which induces the norm $\|\cdot\|$. For $f_i \in \mathcal{V}_i'$ consider the problem

Problem 6. *Find $u_i \in \mathcal{V}_i$ such that*

$$a(u_i, v) = f_i(v) \qquad \forall v \in V_i. \tag{71}$$

Furthermore, let us assume that $\mathcal{W}_i$ is a subspace of $\mathcal{V}_i$ such that we obtain the direct sum

$$\mathcal{V}_i = \mathcal{W}_i \oplus \mathcal{V}_{i-1}.$$

Such subspaces $\mathcal{W}_i$ are called complementary subspaces. A simple construction of a complementary subspaces $\mathcal{W}_i$ can be a obtained by the hierarchical construction as in Example 8 and 9.
The corresponding subspace correction method with recursion parameter μ can be described as follows:

Algorithm 1 (Multilevel cycle with exact subspace correction $(i, (\mu_k))$)**.**
Let $u_{i,1,0} \in \mathcal{V}_i$ be an approximate solution of equation (71).
If $i = 1$, let $u_{i,\mu_i,3}$ be the exact solution of equation (71).
Otherwise, perform the following steps.
1. A priori exact subspace correction:
Find $w_i' \in \mathcal{W}_i$ such that $a(u_{i,1,0} + w_i', w_i) = f_i(w_i) \quad \forall w_i \in \mathcal{W}_i$.
Let $u_{i,1,1} = u_{i,1,0} + w_i'$
For $\mu = 1, \ldots, \mu_i$, do:
BEGIN
2. Coarse-grid correction:
Define $f_{i-1} \in \mathcal{V}_{i-1}'$ by:

$$f_{i-1}(v_{i-1}) = f_i(v_{i-1}) - a(u_{i,\mu,1}, v_{i-1}) \;\forall\, v_{i-1} \in \mathcal{V}_{i-1}.$$

Let $\tilde{u}_{i-1} \in \mathcal{V}_{i-1}$ be the approximate solution of equation (71) obtained by Multilevel cycle with exact subspace correction $(i-1, (\mu_k), \nu)$ and initial approximation $u_{i-1,1,0} = 0$ ($\tilde{u}_{i-1} = u_{i-1,\mu_{i-1},3}$).
Let $u_{i,\mu,2} = u_{i,\mu,1} + \tilde{u}_{i-1}$.

3. A posteriori exact subspace correction:
Find $w_i' \in \mathcal{W}_i$ such that $a(u_{i,\mu,2} + w_i', w_i) = f_i(w_i) \quad \forall w_i \in \mathcal{W}_i$.
Let $u_{i,\mu,3} = u_{i,\mu,2} + w_i'$.
Let $u_{i,\mu+1,1} = u_{i,\mu,3}$.
END
Return $u_{i,\mu_i,3}$.

In case of a direct splitting $\mathcal{V}_i = \mathcal{W}_i \oplus \mathcal{V}_{i-1}$, the equation system on $\mathcal{W}_i$ usually is much easier to solve than the equation system on the complete space $\mathcal{V}_i$. Therefore, we assume that there exists a fast iterative solver $\mathcal{S}_{i,\mathrm{sm}}$ for the linear equation system on $\mathcal{W}_i$. This means, that $\mathcal{S}_{i,\mathrm{sm}}$ is a linear iterative solver for solving the problem

Problem 7.

$$a(u_i, v) = g_i(v) \qquad \forall v \in V_i. \tag{72}$$

for a given $g \in V_i'$, such that the the convergence rate of $\mathcal{S}_{i,\mathrm{sm}}$ does not depend on the number of unknowns. This is stated in the following assumption:

> **Assumption A:** Assume that there are constants $0 < C_{\mathrm{sm}}$ and $0 \leq \rho_{\mathrm{sm}} < 1$ independent of i such that
>
> $$\|(C_{i,\mathrm{sm}})^\nu(w)\| \leq C_{\mathrm{sm}}\rho_{\mathrm{sm}}^\nu\|w\| \quad \forall w \in \mathcal{W}_i, \quad 2 \leq i \leq n.$$

The convergence rate of the whole multilevel algorithm can be estimated by the constant C_{sm}, if the spaces $\mathcal{W}_i$ are a-orthogonal. This follows by the observation that, in this case, a correction in the direction of $\mathcal{W}_i$ does not influence the correction in the direction of another subspace $\mathcal{W}_j$. This leads to the conjecture that the convergence rate of the multilevel algorithm depends on the angle between the coarse-grid space $\mathcal{V}_{i-1}$ and the complementary space $\mathcal{W}_i$. By (92) this implies that the convergence of the multilevel algorithm depends on the strengthened Cauchy-Schwarz inequality between $\mathcal{V}_{i-1}$ and $\mathcal{W}_i$. The aim of this section is to study the convergence rate of the multilevel algorithm with respect to the constant

$$\gamma(\mathcal{V}_{i-1}, \mathcal{W}_i).$$

To this end, let us assume the following:

> **Assumption B:** Assume that there is a constant $0 \leq \gamma < 1$ such that
>
> $$\gamma(\mathcal{V}_{i-1}, \mathcal{W}_i) = \sup_{v \in \mathcal{V}_{i-1}, w \in \mathcal{W}_i} \frac{a(v, w)}{\|v\|\,\|w\|} \leq \gamma, \quad 2 \leq i \leq n.$$

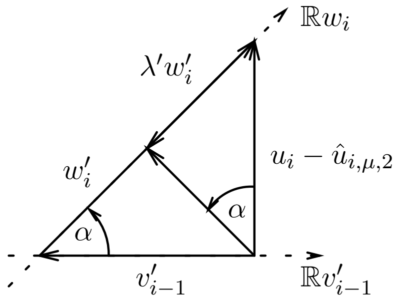


Figure 15: The vectors v_{i-1}', w_i', $u_i - \hat{u}_{i,\mu,2}$, and $u_i - \hat{u}_{i,\mu,2} + \lambda' w_i'$.

In case of the hierarchical construction and bilinear finite elements, one can prove $\gamma = \sqrt{\frac{3}{8}}$ for Poisson's equation.
Our proof of convergence of the multilevel cycle involves several steps. The first step is to analyze the two-grid algorithm with exact coarse-grid correction and exact subspace correction on $\mathcal{W}_i$. An exact subspace correction can be obtained, if $\rho_{\rm sm} = 0$;

Theorem 2 (Two-grid convergence with exact subspace correction)**.** *Assume $2 \le i \le n$ and let $f_i \in \mathcal{V}_i'$ be given. Assume that $u_{i,\mu,1} \in \mathcal{V}_i$, $\tilde{u}_{i-1} \in \mathcal{V}_{i-1}$, and $\tilde{w}_i \in \mathcal{W}_i$, such that*

$$\begin{aligned} a(u_{i,\mu,1}, w_i) &= f_i(w_i) \quad \textit{for every } w_i \in \mathcal{W}_i, \\ a(u_{i,\mu,1} + \tilde{u}_{i-1}, v_{i-1}) &= f_i(v_{i-1}) \quad \textit{for every } v_{i-1} \in \mathcal{V}_{i-1}, \qquad \textit{and} \\ a(u_{i,\mu,1} + \tilde{u}_{i-1} + \tilde{w}_i, w_i) &= f_i(w_i) \quad \textit{for every } w_i \in \mathcal{W}_i. \end{aligned}$$

Define

$$\hat{u}_{i,\mu,2} = u_{i,\mu,1} + \tilde{u}_{i-1} \quad \textit{and} \quad \hat{u}_{i,\mu,3} = u_{i,\mu,1} + \tilde{u}_{i-1} + \tilde{w}_i.$$

Note that $\hat{u}_{i,\mu,3}$ is the solution of the two-grid algorithm with an exact subspace correction corresponding to the spaces $\mathcal{V}_i$ and $\mathcal{V}_{i-1}$. For this algorithm, we obtain

$$\|u_i - \hat{u}_{i,\mu,3}\| \le \gamma^2 \|u_i - u_{i,\mu,1}\|.$$

Proof. First, we prove

$$\|u_i - \hat{u}_{i,\mu,3}\| \le \gamma \|u_i - \hat{u}_{i,\mu,2}\|.$$

Let us write $u_i - \hat{u}_{i,\mu,2} = v_{i-1}' + w_i'$ where $v_{i-1}' \in \mathcal{V}_{i-1}$ and $w_i' \in \mathcal{W}_i$.
Obviously, it is

$$\|u_i - \hat{u}_{i,\mu,3}\| = \min_{w_i \in \mathcal{W}_i} \|u_i - (\hat{u}_{i,\mu,2} + w_i)\| \le \min_{\lambda \in \mathbb{R}} \|u_i - (\hat{u}_{i,\mu,2} - \lambda\, w_i')\|. \quad (73)$$

Let $\lambda' \in \mathbb{R}$ such that

$$\min_{\lambda\in\mathbb{R}} \|u_i - \hat{u}_{i,\mu,2} + \lambda\, w_i'\| = \|u_i - \hat{u}_{i,\mu,2} + \lambda'\, w_i'\|. \tag{74}$$

The vectors v_{i-1}' and w_i' span a 2-dimensional Hilbert space with scalar product $\langle\,\cdot\,,\,\cdot\,\rangle$.
Therefore, (74) is equivalent to

$$a(u_i - \hat{u}_{i,\mu,2} + \lambda'\, w_i', w_i') = 0.$$

This means that $u_i - \hat{u}_{i,\mu,2} + \lambda'\, w_i'$ is orthogonal to w_i'. By equation (71), we get

$$a(u_i - \hat{u}_{i,\mu,2}, v_{i-1}') = 0.$$

This means that $u_i - \hat{u}_{i,\mu,2}$ is orthogonal to v_{i-1}'. Figure 15 depicts this geometric behavior of the vectors v_{i-1}', w_i', $u_i - \hat{u}_{i,\mu,2}$, and $u_i - \hat{u}_{i,\mu,2} + \lambda'\, w_i'$. One can see that the angle α between $-v_{i-1}'$ and w_i' is the angle between $u_i - \hat{u}_{i,\mu,2}$ and $u_i - \hat{u}_{i,\mu,2} + \lambda'\, w_i'$. Therefore, by assumption **B**, we get

$$\|u_i - \hat{u}_{i,\mu,2} + \lambda'\, w_i'\| = \cos(\alpha)\|u_i - \hat{u}_{i,\mu,2}\| = \frac{\langle -w_i', v_{i-1}'\rangle}{\|w_i'\|\,\|v_{i-1}'\|}\,\|u_i - \hat{u}_{i,\mu,2}\| \le \gamma\|u_i - \hat{u}_{i,\mu,2}\|.$$

Thus by (74) and (73), we obtain

$$\|u_i - \hat{u}_{i,\mu,3}\| \le \gamma\|u_i - \hat{u}_{i,\mu,2}\|. \tag{75}$$

Analogously, we get

$$\|u_i - \hat{u}_{i,\mu,2}\| \le \gamma\|u_i - u_{i,\mu,1}\|.$$

The last two inequalities complete the proof. q.e.d.

Now, we generalize this theorem for the case of a recursive coarse-grid correction and exact subspace correction (Algorithm 1).

Theorem 3 (Convergence of the multilevel cycle with exact subspace correction)**.**
Assume that $\rho_{sm} = 0$ and let u_i be the solution of equation (71). Define

$$\rho_i = \sup_{u_i - u_{i,1,0}\in\mathcal{V}_i} \frac{\|u_i - u_{i,\mu_i,3}\|}{\|u_i - u_{i,1,0}\|}, \quad 1 \le i \le n,$$

which is the sharp convergence factor bound for the multilevel cycle with exact subspace correction. Then the following recursion formula holds:

$$\begin{aligned} \rho_i &\le (\gamma^2 + \rho_{i-1}(1-\gamma^2))^{\mu_i}, \quad 2 \le i \le n, \\ \rho_1 &= 0. \end{aligned}$$

Proof. Assume $1 \leq \mu \leq \mu_i$. Let $\hat{u}_{i,\mu,2}$ be the result of the exact coarse-grid correction in Step 2 of the multilevel cycle. Furthermore, let $\hat{u}_{i,\mu,3}$ be the result of the exact subspace correction in Step 3 using the exact coarse-grid correction $\hat{u}_{i,\mu,2}$. This means

$$\begin{aligned} a(\hat{u}_{i,\mu,3}, w_i) &= f_i(w_i) \quad \text{for every } w_i \in \mathcal{W}_i \quad \text{and} \\ \hat{u}_{i,\mu,3} - \hat{u}_{i,\mu,2} &= \hat{w}_i \in \mathcal{W}_i. \end{aligned}$$

Let us introduce the following auxiliary function

$$w_{\text{aux}} := (1 - \rho_{i-1})(\hat{u}_{i,\mu,3} - \hat{u}_{i,\mu,2}).$$

Obviously, it is $w_{\text{aux}} \in \mathcal{W}_i$. Therefore, we get

$$\begin{aligned} \|u_{i,\mu,3} - u_i\| &= \min_{w \in \mathcal{W}_i} \|(u_{i,\mu,2} + w) - u_i\| \leq \\ &\leq \|(u_{i,\mu,2} + w_{\text{aux}}) - u_i\| \leq \\ &\leq \|\rho_{i-1}(\hat{u}_{i,\mu,2} - u_i) + u_{i,\mu,2} - \hat{u}_{i,\mu,2}\| + (1 - \rho_{i-1})\|\hat{u}_{i,\mu,3} - u_i\|. \end{aligned} \tag{76}$$

$\rho_{i-1}(\hat{u}_{i,\mu,2} - u_i)$ is orthogonal to $u_{i,\mu,2} - \hat{u}_{i,\mu,2} \in \mathcal{V}_{i-1}$. By Phytagoras' Theorem, we get

$$\rho_{i-1}^2\|\hat{u}_{i,\mu,2} - u_i\|^2 + \|u_{i,\mu,2} - \hat{u}_{i,\mu,2}\|^2 = \|\rho_{i-1}(\hat{u}_{i,\mu,2} - u_i) + u_{i,\mu,2} - \hat{u}_{i,\mu,2}\|^2. \tag{77}$$

Furthermore, $\hat{u}_{i,\mu,2} - u_i$ is orthogonal to $u_{i,\mu,1} - \hat{u}_{i,\mu,2} \in \mathcal{V}_{i-1}$. This implies

$$\|\hat{u}_{i,\mu,2} - u_i\|^2 + \|u_{i,\mu,1} - \hat{u}_{i,\mu,2}\|^2 = \|u_{i,\mu,1} - u_i\|^2. \tag{78}$$

By Theorem 2, we obtain

$$\|u_i - \hat{u}_{i,\mu,3}\| \leq \gamma^2 \|u_{i,\mu,1} - u_i\|. \tag{79}$$

By the error reduction of the coarse-grid correction, we get

$$\|u_{i,\mu,2} - \hat{u}_{i,\mu,2}\| \leq \rho_{i-1}\|u_{i,\mu,1} - \hat{u}_{i,\mu,2}\|. \tag{80}$$

By (76), (77), (78), (79), and (80), we obtain

$$\begin{aligned} \|u_{i,\mu,3} - u_i\| &\leq \left(\rho_{i-1} + (1 - \rho_{i-1})\gamma^2\right) \|u_{i,\mu,1} - u_i\| \leq \\ &\leq \left(\gamma^2 + \rho_{i-1}(1 - \gamma^2)\right) \|u_{i,\mu,1} - u_i\|. \end{aligned} \tag{81}$$

This implies

$$\|u_{i,\mu_i,3} - u_i\| \leq (\gamma^2 + \rho_{i-1}(1 - \gamma^2))^{\mu_i} \|u_{i,1,1} - u_i\|.$$

Obviously, it is

$$\|u_{i,1,1} - u_i\| \leq \|u_{i,1,0} - u_i\|.$$

This completes the proof. q.e.d.

In practical application, it is important to choose μ_i as small as possible. Therefore, let us put

$$\begin{aligned} \rho_1 &= 0 \quad \text{and} \\ \rho_i &= (\gamma^2 + \rho_{i-1}(1-\gamma^2))^{\mu_i}. \end{aligned} \tag{82}$$

For fixed γ and fixed $(\mu_k)_{k\in\mathbb{N}}$ it is simple to calculate the limit $\lim_{i\to\infty}\rho_i$ numerically and to decide if the limit is smaller than 1 or not. We did this for several recursion parameters $(\mu_k)_{k\in\mathbb{N}}$. Table The result of this analysis is shown in table 1.
In case of a constant recursion parameter $(\mu_k)_{k\in\mathbb{N}}$ one can find an explicit formula, which indicates if the the limit $\lim_{i\to\infty}\rho_i$ is smaller than 1 or not. Lemma 2 states this formula and Table 2 shows some results of this formula.

Lemma 2. *Assume that* $(\mu_i) = \mu \in \mathbb{N}\setminus\{1\}$ *and*

$$\gamma < \gamma_\mu := \sqrt{1-\frac{1}{\mu}}.$$

Then, the equation $\left(\gamma^2 + \rho(1-\gamma^2)\right)^\mu = \rho$ *has a solution* $\rho \in [0,1[$. *The elements of the sequence (82) are contained in the interval* $[0,\rho]$.

Proof. Let us first prove that the equation $\left(\gamma^2 + \rho(1-\gamma^2)\right)^\mu = \rho$ has one solution $0 \le \rho < 1$. A short calculation shows

$$\begin{aligned} \left(\gamma^2 + \rho(1-\gamma^2)\right)^\mu - \rho &= \left(1 + (\rho-1)(1-\gamma^2)\right)^\mu - \rho = \\ &= (1-\rho)\,p(\rho,\gamma), \quad \text{where} \\ p(\rho,\gamma) &= 1 - \sum_{k=1}^{\mu}\binom{\mu}{k}(\rho-1)^{k-1}(1-\gamma^2)^k. \end{aligned}$$

Since $0 \le \gamma < \gamma_\mu$, the polynomial $p(\rho,\gamma)$ has the properties

$$\begin{aligned} p(1,\gamma) &= 1 - \mu(1-\gamma^2) \ < \ 0 \quad \text{and} \\ p(0,\gamma) &= 1 + \sum_{k=1}^{\mu}\binom{\mu}{k}(-1)^k(1-\gamma^2)^k = (1-(1-\gamma^2))^\mu \ \ge \ 0. \end{aligned}$$

Thus, by the continuity of the function $p(\rho,\gamma)$, there is a $0 \le \rho < 1$ such that $p(\rho,\gamma) = 0$. This implies $\left(\gamma^2 + \rho(1-\gamma^2)\right)^\mu = \rho$. By induction we get that $\rho_i \in [0,\rho]$. q.e.d.

(μ_i)	$1+\frac{1}{10}$	$1+\frac{1}{8}$	$1+\frac{1}{6}$	$1+\frac{1}{4}$	$1+\frac{1}{2}$
$\gamma_{(\mu_i)}$	0.2587	0.2880	0.33030	0.39887	0.54119

Table 1: If $\gamma < \gamma_{(\mu_i)}$, then there is a $\rho < 1$ such that $\rho_i < \rho$ for every i.

μ	2	3	4	5	6	7	8
$\gamma_\mu = \sqrt{\frac{\mu-1}{\mu}}$	0.70710	0.81650	0.86602	0.89443	0.91287	0.92582	0.93541

Table 2: If $\gamma < \gamma_\mu$, then there is a $\rho < 1$ such that $\rho_i < \rho$ for every i.

Example 11 (W-cycle $\mu = 2$.)**.** *The equation* $\left(\gamma^2 + \rho(1-\gamma^2)\right)^2 = \rho$ *has the solution*

$$\rho = \frac{\gamma^4}{(\gamma^2-1)^2}. \tag{83}$$

For the hierarchical construction it is $\gamma = \sqrt{\frac{3}{8}}$ *(see Table **??**). This leads to* $\rho_i \leq \rho = \frac{9}{25} = 0.36$*.*

Theorem 3 can be extended to the case of approximate subspace corrections that satisfy assumption **A** (see Algorithm **??**). This gives Theorem 4 the proof of which can be found in [12]

Theorem 4 (Convergence of the multilevel cycle with an approximate subspace correction)**.** *Let*

$$\rho_i = \sup_{u_i - u_{i,1,0} \in \mathcal{V}_i} \frac{\|u_i - u_{i,\mu_i,3}\|}{\|u_i - u_{i,1,0}\|}$$

be the convergence rate of the general cycle with ν *smoothing operations. Then, for every* $\epsilon > 0$*, there exists a number* ν_ϵ *which depends only on* ϵ*,* γ*,* μ_i*,* C_{sm}*, and* ρ_{sm} *such that:*

If $\rho_{i-1} \leq 1$*, then the following recursion formula holds*

$$\begin{aligned} \rho_i &\leq (\gamma^2 + \rho_{i-1}(1-\gamma^2))^{\mu_i} + \epsilon, \\ \rho_1 &= 0, \end{aligned}$$

for every $\nu \geq \nu_\epsilon$*.*

Furthermore, the following inequality holds:

$$\rho_i \leq \theta_{\mu_i,1} + \theta_{\mu_i,2} + \theta_{\mu_i,3}$$

where

$$\begin{aligned}\theta_{\mu,1} &= (\gamma^2 + \rho_{i-1}(1-\gamma^2))\,(\theta_{\mu-1,1} + \theta_{\mu-1,2}),\\ \theta_{\mu,2} &= (\gamma + \rho_{i-1}(1-\gamma))\,\theta_{\mu-1,3}, \quad \textit{and}\\ \theta_{\mu,3} &= C_{sm}\rho_{sm}^{\nu}\sqrt{1+\rho_{i-1}^2}(\theta_{\mu-1,1} + \theta_{\mu-1,2} + \theta_{\mu-1,3})\end{aligned}$$

for $\mu = 2, \dots, \mu_i$ *and*

$$\begin{aligned}\theta_{1,1} &= (\gamma^2 + \rho_{i-1}(1-\gamma^2)),\\ \theta_{1,2} &= C_{sm}\rho_{sm}^{\nu}(\gamma + \rho_{i-1}(1-\gamma)), \quad \textit{and}\\ \theta_{1,3} &= C_{sm}\rho_{sm}^{\nu}\sqrt{(\gamma + C_{sm}\rho_{sm}^{\nu}(1-\gamma))^2 + \rho_{i-1}^2(C_{sm}\rho_{sm}^{\nu}+1)^2}.\end{aligned}$$

4.2 Result for a Non-Symmetric Bilinear Forms

To obtain a robust estimation of the convergence rate of the multilevel algorithm in case of non-symmetric bilinear forms, we have to estimate the convergence rate of the multilevel algorithm in a norm which includes the non-symmetric part of a. A natural norm with this property is the operator norm of a. For the definition of this norm we need a Hilbert space Therefore, let us assume that $\mathcal{V}_n$ is a Hilbert space with scalar product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$. Then, we can define the following semi-norms on $\mathcal{V}_n$

$$\begin{aligned}\|u\|_{\mathcal{V}_i} &:= \sup_{v_i\in\mathcal{V}_i}\frac{a(u,v_i)}{\|v_i\|} \quad \text{for } u\in\mathcal{V}_n \quad \text{and}\\ \|u\|_{\mathcal{W}_i} &:= \sup_{w_i\in\mathcal{W}_i}\frac{a(u,w_i)}{\|w_i\|} \quad \text{for } u\in\mathcal{V}_n.\end{aligned}$$

Obviously, $\|\cdot\|_{\mathcal{V}_i}$ is a norm on $\mathcal{V}_i$ and $\|\cdot\|_{\mathcal{W}_i}$ is a norm on $\mathcal{W}_i$. These norms contain the non-symmetric part of a. The scalar product $\langle\cdot,\cdot\rangle$ on $\mathcal{V}$ can be defined in different ways. Often, one can construct this scalar product with the help of the symmetric part of a. But, we do not have to specify the scalar product $\langle\cdot,\cdot\rangle$ for the general theory in this section.

> **Assumption A:** Assume that there are constants $C_{\text{sm}} > 0$ and $\rho_{\text{sm}} \in [0,1)$, independent of i, such that
>
> $$\|(B_i)^{\nu}(w)\|_{\mathcal{W}_i} \le C_{\text{sm}}\rho_{\text{sm}}^{\nu}\|w\|_{\mathcal{W}_i} \quad \forall w\in\mathcal{W}_i, \quad 2\le i\le n.$$

The second assumption is the strengthened Cauchy-Schwarz inequality in the Hilbert space $\mathcal{V}$:

> **Assumption B:** Assume that there is a constant $0 \le \tilde{\gamma} < 1$ such that
>
> $$\gamma(\mathcal{V}_{i-1},\mathcal{W}_i) = \sup_{v\in\mathcal{V}_{i-1},w\in\mathcal{W}_i}\frac{\langle v,w\rangle}{\|v\|\,\|w\|} \le \tilde{\gamma}, \quad 2\le i\le n.$$

The third assumption must involve the bilinear form a. It is something like a generalization of the strengthened Cauchy-Schwarz inequality for nonsymmetric bilinear forms.

Assumption C: Assume that there is a constant $0 < \gamma < 1$ such that

$$\begin{aligned} \|v\|_{\mathcal{W}_i} &\leq \gamma \|v\|_{\mathcal{V}_{i-1}} \quad \forall\, v \in \mathcal{V}_{i-1}, \quad 2 \leq i \leq n, \\ \|w\|_{\mathcal{V}_{i-1}} &\leq \gamma \|w\|_{\mathcal{W}_i} \quad \forall\, w \in \mathcal{W}_i, \quad 2 \leq i \leq n. \end{aligned}$$

Using these assumptions one can prove (see [13]):

Theorem 5 (Convergence of the multilevel cycle)**.** *Let*

$$\rho_i = \sup_{u_i - u_{i,1,0} \in \mathcal{V}_i} \frac{\|u_i - u_{i,\mu_i,3}\|_{\mathcal{V}_i}}{\|u_i - u_{i,1,0}\|_{\mathcal{V}_i}}, \quad 1 \leq i \leq n,$$

which is the sharp convergence factor bound for the general cycle with ν approximate subspace corrections. Then, for every $\epsilon > 0$, there exists a number ν_ϵ that depends only on ϵ, γ, $\tilde{\gamma}$, C_{sm}, and ρ_{sm} such that the following holds:

If $\rho_{i-1} \leq 1$, then

$$\begin{aligned} \rho_i &\leq \sqrt{\frac{1+\gamma^2}{1-\tilde{\gamma}}}\,((1+\gamma^2)\rho_{i-1} + \gamma^2)^{\mu_i} + \epsilon, \quad 2 \leq i \leq n, \\ \rho_1 &= 0, \end{aligned}$$

for every $\nu \geq \nu_\epsilon$.

It is important to choose (μ_i) as small as possible. Table 3 helps to choose (μ_i) for constant $(\mu_i) = \mu$. If one chooses $\gamma = \tilde{\gamma} < \gamma_\mu$, then the convergence rate of the multilevel cycle is smaller than the value ρ_μ in Table 3.

μ	2	3	4	5	6	7	8
γ_μ	0.395	0.518	0.591	0.642	0.680	0.710	0.734
ρ_μ	0.122	0.104	0.0732	0.0612	0.0527	0.0487	0.0442

Table 3: For a non-symmetric bilinear form a choose the smallest μ such that $\gamma < \gamma_\mu$ and $\tilde{\gamma} < \gamma_\mu$. Then, the convergence rate of the multi-level cycle is smaller than ρ_μ.

4.3 Analysis of the Strengthened Cauchy-Schwarz Inequality

4.3.1 Introduction

Consider a finite element space $\mathcal{V}_n$ and a symmetric positive definite bilinear form

$$a : \mathcal{V}_n \times \mathcal{V}_n \to \mathbb{R}.$$

Then, $\mathcal{V}_n$ is a Hilbert space with scalar product a. Assume that $\mathcal{V}_{n-1}$ is a coarse subspace of $\mathcal{V}_n$. We are looking for a complementary space $\mathcal{W}_n$ such that the multilevel cycle 1 is a fast iterative solver. The theory in section 4.1 and 4.2 shows that we need a complementary space $\mathcal{W}_n$ spanned by functions with a small support and such that the constant in the strengthened Cauchy-Schwarz inequality

$$\gamma(\mathcal{V}_{n-1}, \mathcal{W}_n, a) := \sup_{u \in \mathcal{V}_{n-1}, v \in \mathcal{W}_n} \frac{|a(u,v)|}{\sqrt{a(u,u)}\sqrt{a(v,v)}}$$

is small. To solve this problem, we first try to construct suitable complementary spaces in one dimension (see section 4.3.3 and ...). Then, simple tensor product constructions will lead to a construction for the two-dimensional case.
The simplest way to construct a complementary space is to use the hierarchical basis. Let us explain this by a 1-dimensional example. Let $\mathring{V}_n \subset H_0^1(]0,1[)$ be the space of piecewise linear functions of meshsize $h = 2^{-n}$ and v_p^n the corresponding nodal basis functions. Then, we get

$$W_n^{\text{hier}} := \operatorname{span}_{\mathbb{R}}\left\{ v_p^n = v_p \;\middle|\; p \in \Omega_n \backslash \Omega_{n-1} \right\}.$$

Obviously, this construction leads to the direct sum

$$\mathring{V}_n = \mathring{V}_{n-1} \oplus W_n^{\text{hier}}.$$

Let us calculate the constant in the strengthened Cauchy-Schwarz inequality with respect to the H^1- and L^2-bilinear forms

$$\int_0^1 \frac{\partial u}{\partial x}\frac{\partial v}{\partial x}\, dx \quad \text{and} \quad \int_0^1 uv\, dx.$$

By Theorem **??**, v_p, $p \in \Omega_n$ is an orthogonal basis of $\mathring{V}_n$ and v_p, $p \in \Omega_{n-1}$ is an orthogonal basis of $\mathring{V}_{n-1}$. This implies

$$\gamma(\mathring{V}_{n-1}, W_n^{\text{hier}}, H^1) := \sup_{u \in \mathring{V}_{n-1}, v \in W_n^{\text{hier}}} \frac{\int_0^1 \frac{\partial u}{\partial x}\frac{\partial v}{\partial x}\, dx}{|u|_{H^1}|v|_{H^1}} = 0,$$

where $|v|_{H^1}^2 := \int_0^1 \left(\frac{\partial v}{\partial x}\right)^2 dx$. This certainly is the optimal construction of a complementary space with respect to the H^1-bilinear form. Unfortunately, the hierarchical basis is not H^1-orthogonal in two dimensions.
Now, let us study the constant in the strengthened Cauchy-Schwarz inequality with respect to the L^2 bilinear form. Let us do this in several steps.

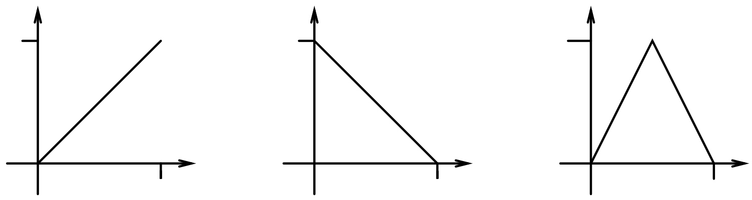


Figure 16: The functions x, $1-x$, and $v_{\frac{1}{2}}$.

Step 1. *Localization.*
Consider the basis functions x, $1-x$, and $v_{\frac{1}{2}}$ in Figure 16. Let us assume that we can prove

$$\left\|(c_L(1-x)+c_R x)\right\|^2_{L^2(]0,1[)}+\left\|b_M v_{\frac{1}{2}}\right\|^2_{L^2(]0,1[)} \leq K\left\|(c_L(1-x)+c_R x)+b_M v_{\frac{1}{2}}\right\|^2_{L^2(]0,1[)} \tag{84}$$

for every parameter $c_L, c_R, b_M \in \mathbb{R}$, where $K > 1$ is a fixed constant. Then, this implies

$$\|u\|^2_{L^2(]ih,(i+1)h[)} + \|v\|^2_{L^2(]ih,(i+1)h[)} \leq K\|u+v\|^2_{L^2(]ih,(i+1)h[)}$$

for every $u \in \mathring{V}_{n-1}$, $v \in W_n^{\text{hier}}$, and $i = 0, \cdots, 2^n - 1$, where $h = 2^{-n}$.
Summing up these inequalities yields

$$\|u\|^2_{L^2(]0,1[)} + \|v\|^2_{L^2(]0,1[)} \leq K\|u+v\|^2_{L^2(]0,1[)}$$

for every $u \in \mathring{V}_{n-1}$, $v \in W_n^{\text{hier}}$. Now, we can apply Lemma 3. This gives

$$\gamma(\mathring{V}_{n-1}, W_n^{\text{hier}}, L^2) \leq 1 - K^{-1}.$$

Therefore, it is enough to prove (84).

Step 2. *Algebraic analysis.*
A short calculation shows that (84) is equivalent to

$$\frac{1}{3}\left(c_L^2 + c_R^2 + c_L c_R + b_M^2\right) \leq K\frac{1}{3}\left(c_L^2 + c_R^2 + c_L c_R + b_M^2\right) + K\frac{1}{2}\left(c_L + c_R\right)b_M.$$

This equation is equivalent to

$$0 \leq c_L^2 + c_R^2 + c_L c_R + b_M^2 + \frac{K}{K-1}\frac{3}{2}(c_L + c_R) b_M.$$

Of course, this equation should hold for the special case $c_L = c_R = c$. This leads to the inequality

$$0 \leq 3c^2 + b_M^2 + 2\frac{K}{K-1}\frac{3}{2} c b_M.$$

This inequality is correct if

$$\sqrt{\frac{3}{4}} = \frac{K-1}{K} = 1 - K^{-1}, \tag{85}$$

since then

$$0 \leq 3c^2 + b_M^2 + 2\frac{K}{K-1}\frac{3}{2} c b_M = \left(\sqrt{3}c + b_M\right)^2.$$

Therefore, we choose K such that (85) holds. Then, we get

$$0 \leq \frac{1}{4}(c_L - c_R)^2 + \left(\sqrt{3}\frac{1}{2}(c_L - c_R) + b_M\right)^2 = c_L^2 + c_R^2 + c_L c_R + b_M^2 + \frac{K}{K-1}\frac{3}{2}(c_L + c_R) b_M$$

and (84) is proved for this choice of K.
Combining Step 1. - 2. shows that

$$\gamma(\mathring{V}_{n-1}, W_n^{\text{hier}}, L^2) := \sup_{u \in \mathring{V}_{n-1}, v \in W_n^{\text{hier}}} \frac{\int_0^1 uv \; dx}{\|u\|_{L^2} \|v\|_{L^2}} \leq \sqrt{\frac{3}{4}} \approx 0.86603.$$

Summarizing the above estimations, we state the following

Proposition 1 (Hierarchical basis in 1D)**.**

$$\begin{aligned} \gamma(\mathring{V}_{n-1}, W_n^{hier}, H^1) &= \sup_{u \in \mathring{V}_{n-1}, v \in W_n^{hier}} \frac{\int_0^1 \frac{\partial u}{\partial x}\frac{\partial v}{\partial x} \; dx}{|u|_{H^1} |v|_{H^1}} = 0, \\ \gamma(\mathring{V}_{n-1}, W_n^{hier}, L^2) &= \sup_{u \in \mathring{V}_{n-1}, v \in W_n^{hier}} \frac{\int_0^1 uv \; dx}{\|u\|_{L^2} \|v\|_{L^2}} \leq \sqrt{\frac{3}{4}} \approx 0.86603. \end{aligned}$$

In several applications, the constant 0.86603 is a too large constant for obtaining a fast multilevel algorithm. In section 4.3.3 - 4.3.5, we show how to construct complementary spaces $\mathcal{W}_n$ spanned by functions with a small support and such that the constant in the strengthened Cauchy-Schwarz inequality is smaller than 0.4 for the H^1 and L^2-bilinear form. In the next section, we explain how to estimate the strengthened Cauchy-Schwarz inequality for the hierarchical basis in 2D.

4.3.2 Hierarchical Decomposition

In this section, we study the constant in the strengthened Cauchy-Schwarz inequality for the hierarchical construction Example 9 in two dimensions. For reasons of simplicity, let Ω be the unit square $]0,1[^2$. But, the results in this sections also hold for a polygonal domain such that the corners of Ω are contained in $\mathbb{Z} \times \mathbb{Z}$ and such that the boundary of Ω is the union of vertical and horizontal lines. Let $V_{n,n}$ be the space of piecewise bilinear functions with meshsize $h = 2^{-n}$ on Ω. Recall that $\mathring{V}_{n,n}$ is the space of piecewise bilinear functions of meshsize $h = 2^{-n}$ in x- and y-direction and that $W_{n,n}^{\text{hier}}$ is the space defined by

$$W_{n,n}^{\text{hier}} := \operatorname{span}_{\mathbb{R}}\left\{ v_{(p,q)}^n \mid (p,q) \in \Omega_n \times \Omega_n \backslash \Omega_{n-1} \times \Omega_{n-1} \right\}.$$

Observe, that the space $W_{n,n}^{\text{hier}}$ can be described in the following way

$$\begin{aligned} W_{n,n}^{\text{hier}} &:= \left\{ u \in \mathring{V}_{n,n} \;\middle|\; u(p) = 0 \quad \text{for } p \in \Omega_{n-1} \times \Omega_{n-1} \right\} \quad \text{if } n > 1 \text{ and} \\ W_{1,1}^{\text{hier}} &:= \mathring{V}_{1,1}. \end{aligned}$$

Let

$$a : \mathring{V}_{n,n} \times \mathring{V}_{n,n} \to \mathbb{R}$$

be a symmetric positive definite bilinear form. The aim of this section is to estimate the constant in the strengthened Cauchy-Schwarz inequality

$$\gamma(\mathring{V}_{n-1,n-1}, W_{n,n}^{\text{hier}}, a) := \max_{v \in \mathring{V}_{n-1,n-1}, w \in W_{n,n}^{\text{hier}}} \frac{a(v,w)}{\sqrt{a(v,v)}\sqrt{a(w,w)}},$$

where we write $\frac{0}{0} := 0$. The result of our analysis is printed in Table 4. First, we explain our analysis in general. Assume that a is one of the bilinear forms

$$\begin{aligned} a(u,v) &= \int_\Omega \langle \nabla u, \nabla v \rangle \; d(x,y), \\ a(u,v) &= \int_\Omega \frac{\partial u}{\partial(\cos\phi, \sin\phi)} \frac{\partial v}{\partial(\cos\phi, \sin\phi)} \; d(x,y), \quad \text{where } 0 \le \phi \le 2\pi, \text{ or} \\ a(u,v) &= \int_\Omega u\, v \; d(x,y). \end{aligned}$$

Here, we abbreviate

$$\frac{\partial w}{\partial(\cos\phi, \sin\phi)} := \frac{\partial w}{\partial x} \cos\phi + \frac{\partial w}{\partial y} \sin\phi.$$

$a(u,v)$	$\gamma(\text{Hier}, n, a) \leq$
$\int_\Omega \langle \nabla u, \nabla v \rangle \; d(x,y)$	$\sqrt{\frac{3}{8}} \approx 0.612372$
$\int_\Omega \frac{\partial u}{\partial(\cos\phi, \sin\phi)} \frac{\partial v}{\partial(\cos\phi, \sin\phi)} \; d(x,y)$	$\sqrt{\frac{3}{4}} \approx 0.866025$
$\int_\Omega u\, v \; d(x,y)$	$\sqrt{\frac{15}{16}} \approx 0.968246$

Table 4: Constant in the strengthened Cauchy-Schwarz inequality for the hierarchical decomposition.

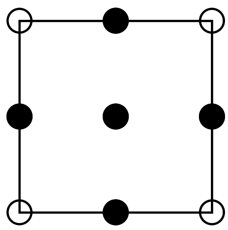


Figure 17: W_{square} is the space of piecewise bilinear functions on a grid of meshsize $\frac{1}{2}$ which are zero at the coarse grid points (above white points).

$$\gamma(\mathring{V}_{n-1,n-1}, W_{n,n}^{\text{hier}}, a) := \max_{v \in \mathring{V}_{n-1,n-1}, w \in W_{n,n}^{\text{hier}}} \frac{a(v,w)}{\sqrt{a(v,v)}\sqrt{a(w,w)}},$$

Step 1. *Localization.*
Define

$$\begin{aligned}
\Omega_{\text{square}} &:= \;]0,1[^2, \\
V_{\text{square}} &:= \left\{ u \;\middle|\; u \text{ is bilinear on } \Omega_{\text{square}} \right\} \\
&= \text{span}\left\{ xy, (1-x)y, x(1-y), (1-x)(1-y) \right\}, \quad \text{and} \\
W_{\text{square}} &:= \Big\{ u \in \mathcal{C}(\bar{\Omega}_{\text{square}}) \;\Big|\; u(0,0) = u(1,0) = u(0,1) = u(1,1) = 0 \text{ and } u \text{ is bilinear} \\
&\qquad \text{on the subdomains }]0,0.5[^2,\ ]0.5,1.0[^2,\ ]0,0.5[\times]0.5,1.0[,\ ]0.5,1.0[\times]0,0.5[\Big\}.
\end{aligned}$$

The grid points corresponding to the space W_{square} are depicted in Figure 17. Let a_{square} be the bilinear a, but replace the integral by the integral over the domain Ω_{square}. Let $\|\cdot\|$ be the semi-norm induced by the bilinear form a or a_{square}. Assume that $K > 1$ is a constant such that

$$\|v\|^2 + \|w\|^2 \le K\|v+w\|^2 \quad \text{for every } v \in V_{\text{square}} \text{ and } w \in W_{\text{square}}. \quad (86)$$

Then, this inequality holds on every cell of the grid $\Omega_n \times \Omega_n$. Summing up these inequalities gives

$$\|v\|^2 + \|w\|^2 \le K\|v+w\|^2 \quad \text{for every } v \in V_{n-1,n-1} \text{ and } w \in W_{n,n}^{\text{hier}}.$$

By Lemma 3, we obtain

$$\gamma(\mathring{V}_{n-1,n-1}, W_{n,n}^{\text{hier}}, a) \le \frac{K-1}{K}.$$

Step 2. *Analysis of the matrix equation.*
Let w_1, w_2, w_3, and w_4 be a basis of V_{square} and let w_5, w_6, w_7, w_8, and w_9 be a basis of W_{square}. Let $\mathcal{A} = (a_{i,j})_{1\le i,j\le 9}$ be the matrix of the bilinear form a_{square} with respect to the basis $\{w_1, w_1, \cdots, w_9\}$. Now, let $\mathcal{B}$ be the block diagonal matrix of $\mathcal{A}$:

$$\mathcal{B} = \left\{ \begin{array}{ccccccccc} a_{11} & a_{12} & a_{13} & a_{14} & 0 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & 0 & 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} & 0 & 0 & 0 & 0 & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{55} & a_{56} & a_{57} & a_{58} & a_{59} \\ 0 & 0 & 0 & 0 & a_{65} & a_{66} & a_{67} & a_{68} & a_{69} \\ 0 & 0 & 0 & 0 & a_{75} & a_{76} & a_{77} & a_{78} & a_{79} \\ 0 & 0 & 0 & 0 & a_{85} & a_{86} & a_{87} & a_{88} & a_{89} \\ 0 & 0 & 0 & 0 & a_{95} & a_{96} & a_{97} & a_{98} & a_{99} \end{array} \right\}.$$

Furthermore, put $\mathcal{C}_K := K\,\mathcal{A} - \mathcal{B}$. Then, inequality (86) is equivalent to
$\mathcal{C}_K$ is a positive semi-definite matrix.
This is equivalent to
$\mathcal{C}_K$ has no negative eigenvalues.
Let $P_K(x) := \det(E\,x\, - \,\mathcal{C}_K)$ be the characteristic polynomial of $\mathcal{C}_K$.
Then, inequality (86) is equivalent to
The roots of $P_K(x)$ are not negative.
Therefore, we have to solve the following algebraic problem:

> Find the minimal real number $K > 1$ such that the roots of $P_K(x)$ are not negative.

Step 3. Analysis of the characteristic polynomial.

The characteristic polynomial $P_K(x)$ has the following form

$$P_K(x) = p_0(K) + p_1(K)x + p_2(K)x^2 + \cdots + p_9(K)x^9$$

where $p_i(K)$ are polynomials. Let $s \in \mathbb{N}_0$ be the maximal number such that $p_0(K) = p_1(K) = \cdots = p_{s-1}(K) = 0$. Then $p_s(K)$ is the product of all roots of $P_K(x)$ which are not zero for every K. We suppose that the optimal value of K is such that $P_K(x)$ has s or more zero roots. This leads to the ansatz

> **Ansatz:**
> Find the roots $K_1, K_2, \ldots, K_l$ of the polynomial $p_s(K)$.
> Find the minimal $K_j > 1$ such that the roots of $P_{K_j}(x)$ are not negative.

For each of our bilinear forms this ansatz succeeded. Therefore, we get

$$\gamma(\mathring{V}_{n-1,n-1}, W_{n,n}^{\text{hier}}, a) \leq \frac{K_j - 1}{K_j}.$$

We calculated the value K_j with the help of a algebra manipulation program. The polynomials $P_K(x)$ are very long expressions. Therefore, we do not want to write them down here. But let us explain the calculation of K_j for two bilinear forms in more detail.
Example 1: $a(u,v) = \int_\Omega \langle \nabla u, \nabla v \rangle \; d(x,y)$.
For this bilinear form we get $s = 1$ and

$$p_1(K) = -\frac{81}{64}(K-1)^4(8 - 16K + 5K^2)^2.$$

This leads to $K_j = \frac{2}{5}(4 + \sqrt{6})$ and

$$\gamma(\mathring{V}_{n-1,n-1}, W_{n,n}^{\text{hier}}, a) \leq \frac{3 + 2\sqrt{6}}{8 + 2\sqrt{6}} = \sqrt{\frac{3}{8}}.$$

Example 2: $a(u,v) = \int_\Omega \frac{\partial u}{\partial(\cos\phi, \sin\phi)} \frac{\partial v}{\partial(\cos\phi, \sin\phi)} \; d(x,y)$. For this bilinear form we get $s = 2$. The problem of this bilinear form is that $p_2(K)$ depends on ϕ. But, $K_j = 2(2 + \sqrt{3})$ is a root of $p_2(K)$ for every ϕ. Therefore, we analyze the polynomial $Q_\phi(x) = P_{K_j}(x)\, x^{-3}$. We have to show that all roots of $Q_\phi(x)$ are non-negative for every $\phi \in \mathbb{R}$. By symmetry, it is enough to study $Q_\phi(x)$ for $0 \leq \phi \leq \frac{\pi}{4}$. One can show that the roots of $Q_\phi(\frac{\pi}{4})$ are greater than 1 and that the roots of $Q_\phi(0)$ are non-negative. Now, we use the continuity of the roots with respect to ϕ. If there is a $0 < \phi_n \leq \frac{\pi}{4}$ such that $Q_{\phi_n}(x)$ has a negative root, then there must be a $\phi_n \leq \phi_0 \leq \frac{\pi}{4}$ such that $Q_{\phi_0}(x)$ has a zero root. This implies $Q_{\phi_0}(0) = 0$. But it is

$$Q_\phi(0) = \frac{1}{4}(1351 + 780\sqrt{3})(4 - \cos(4\phi))\sin^2(2\phi) > 0 \quad \text{for } 0 < \phi \leq \tfrac{\pi}{4}.$$

Therefore, we get

$$\gamma(\mathring{V}_{n-1,n-1}, W_{n,n}^{\text{hier}}, a) \leq \frac{3+2\sqrt{3}}{4+2\sqrt{3}} = \sqrt{\frac{3}{4}}.$$

for every ϕ.

4.3.3 Prewavelets

In several applications, the hierarchical construction leads to a too large constant γ. A smaller canstant can be obtained by prewavelets. Let us explain these functions in one dimension. The prewavelet functions are a basis of the L_2-orthogonal complement space. This space $\mathring{W}_n \subset \mathring{V}_n$ is defined by

$$\mathring{W}_n := \left\{ w \in \mathring{V}_n \;\middle|\; \int_0^1 wv \; dx = 0 \quad \text{for every } v \in \mathring{V}_{n-1} \right\}.$$

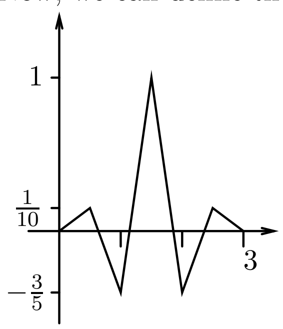
Then, we obtain the following L_2-orthogonal direct sum

$$\mathring{V}_n = \mathring{W}_n \oplus_{L^2} \mathring{V}_{n-1}.$$

A prewavelet is a function which induces by shifting a basis of the space $\mathring{W}_n$. The boundary makes it difficult to define a shifting in the space $\mathring{W}_n$. Therefore, we first define prewavelet functions for a related space. Let $\bar{V}_n \subset \mathcal{C}(\mathbb{R})$ be the space of 2-periodic and piecewise linear functions on a uniform grid of meshsize $h = 2^{-n}$, where $n \in \mathbb{N}$. Figure 18 shows an example of a function in $\bar{V}_n$ restricted on the interval $[-1, 1]$. By 2-periodicity, every function $w \in \bar{V}_n$ is uniquely defined by its values on the interval $[-1, 1]$ and has the property $w(-1) = w(1)$.
$\bar{V}_n$ is an L^2- and H^1-Hilbert space with respect the standard inner products $\int_{-1}^1 w\, v \; dx$ and $\int_{-1}^1 w\, v + \frac{\partial w}{\partial x} \frac{\partial v}{\partial x} \; dx$, respectively. The corresponding respective norms are $\| \cdot \|_{L^2}$ and $\| \cdot \|_{H^1}$. The semi-norm $| \cdot |_{H^1}$ on $\bar{V}_n$ is defined by $|w|_{H^1}^2 := \int_{-1}^1 \left(\frac{\partial w}{\partial x} \right)^2 \; dx$. Now let $\bar{W}_n \subset \bar{V}_n$ be the L_2-orthogonal complement space of $\bar{V}_{n-1}$ for $n \geq 2$. This is the space

$$\bar{W}_n := \left\{ w \in \bar{V}_n \;\middle|\; \int_{-1}^1 wv \; dx = 0 \quad \text{for every } v \in \bar{V}_{n-1} \right\}.$$

Now, we can define the prewavelet functions. Let ψ_M be the function



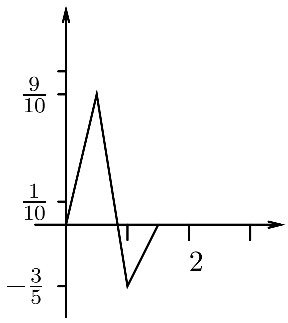
Furthermore, we define the function $\bar{\varphi}_{n,k} \in \bar{W}_n$ by

$$\bar{\varphi}_{n,k}(x) := \begin{cases} \psi_M(x\,2^{n-1} - k + 2) & \text{if } 0 \leq x\,2^{n-1} - k + 2 \leq 3, \\ 0 & \text{elsewhere} \end{cases},$$

where $k = \pm 1, \ldots, \pm 2^{n-1}$. The functions $\bar{\varphi}_{n,k}$, $k = \pm 1, \ldots, \pm 2^{n-1}$ form a basis of $\bar{W}_n$.
Now, we can define the prewavelet functions in $\mathring{W}_n$. Observe that the functions $\bar{\varphi}_{n,k}\big|_{]0,1[}$ are contained in $\mathring{W}_n$ if $k = 2, \ldots, 2^{n-1} - 1$. To construct a prewavelet function near the boundary with Dirichlet boundary condition, we observe that

$$(\bar{\varphi}_{n,1} - \bar{\varphi}_{n,-1})\,\big|_{]0,1[} \in \mathring{W}_n.$$

These considerations lead to the following construction. Define the function ψ_L^D by:



Now, define the functions $\varphi_{n,k}$ by

$$\begin{aligned} \varphi_{n,1}(x) &:= \psi_L^D(x\,2^{n-1}), \\ \varphi_{n,k}(x) &:= \psi_M(x\,2^{n-1} - k + 2) \quad \text{for } k = 2, \ldots 2^{n-1} - 1, \\ \varphi_{n,2^{n-1}}(x) &:= \psi_L^D((1-x)\,2^{n-1}). \end{aligned}$$

Obviously, the constant in the strengthened Cauchy-Schwarz inequality between $\mathring{V}_{n-1}$ and W_n is

$$\gamma(\mathring{V}_{n-1}, W_n, L^2) := \max_{v \in \mathring{V}_{n-1}, w \in W_n} \frac{\int_0^1 vw\,dx}{\|v\|_{L^2}\|w\|_{L^2}} = 0$$

with respect to the L^2-bilinear form. Now, we want to estimate the constant γ with respect to the H^1-bilinear form. This is the constant

$$\gamma(\mathring{V}_{n-1}, W_n, H^1) := \max_{v \in \mathring{V}_{n-1}, w \in W_n} \frac{\int_0^1 \frac{\partial v}{\partial x}\frac{\partial w}{\partial x}\,dx}{|v|_{H^1}|w|_{H^1}},$$

where $|v|_{H^1}^2 := \int_0^1 \left(\frac{\partial v}{\partial x}\right)^2 dx$.

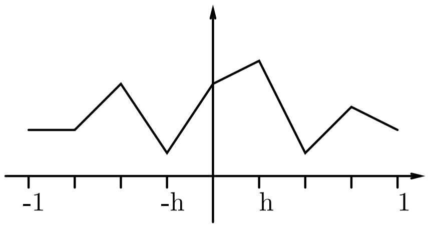


Figure 18: Example of a function in $\bar{V}_n$.

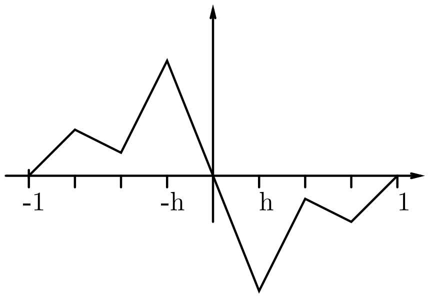


Figure 19: Symmetric extension $\tilde{u}$.

Theorem 6. *The following estimate holds:*

$$\gamma(\mathring{V}_{n-1}, W_n, H^1) \le \sqrt{\frac{3}{19}} \approx 0.39736.$$

Proof. Step 1. *Remove boundary conditions.*
Define the constant

$$\gamma(\bar{V}_{n-1}, \bar{W}_n, H^1) := \max_{v\in\bar{V}_{n-1}, w\in\bar{W}_n} \frac{\int_{-1}^{1} \frac{\partial v}{\partial x}\frac{\partial w}{\partial x}\, dx}{|v|_{H^1(]-1,1[)}|w|_{H^1(]-1,1[)}}$$

where $|v|^2_{H^1(]-1,1[)} := \int_{-1}^{1} \left(\frac{\partial v}{\partial x}\right)^2 dx$. Furthermore, define the symmetric extension operator $\tilde{} : V_n \to \bar{V}_n$ by (see Figure 19)

$$\tilde{u}(x) = -\tilde{u}(-x) = u(x) \quad \forall\ 0 \le x \le 1, \quad u \in V_n.$$

Then, it is enough to prove

$$\gamma(\bar{V}_{n-1}, \bar{W}_n, H^1) \le \sqrt{\frac{3}{19}}, \tag{87}$$

since

$$\begin{aligned} \gamma(\mathring{V}_{n-1}, W_n, H^1) &= \max_{v\in\mathring{V}_{n-1}, w\in W_n} \frac{\int_{0}^{1} \frac{\partial v}{\partial x}\frac{\partial w}{\partial x}\, dx}{|v|_{H^1}|w|_{H^1}} \\ &= \max_{v\in\mathring{V}_{n-1}, w\in W_n} \frac{\int_{-1}^{1} \frac{\partial \tilde{v}}{\partial x}\frac{\partial \tilde{w}}{\partial x}\, dx}{|\tilde{v}|_{H^1(]-1,1[)}|\tilde{w}|_{H^1(]-1,1[)}} \\ &\le \gamma(\bar{V}_{n-1}, \bar{W}_n, H^1). \end{aligned}$$

Step 2. *Localization.*
The set of nodal points of the space $\bar{V}_n$ is

$$\bar{\Omega}_n := \left\{ p \,=\, k\, 2^{-n} \mid k = -2^n + 1, \cdots, 2^n \right\}$$

and the nodal basis function $v_p^n \in \bar{V}_n$ at the nodal point $p \in \bar{\Omega}_n$ is defined by

$$v_p^n(k\,2^{-n}) = \begin{cases} 0 & \text{if } p \neq k\,2^{-n} \\ 1 & \text{if } p = k\,2^{-n} \end{cases},$$

where $k = -2^n + 1, \cdots, 2^n$. The set of functions v_p^{n-1}, $p \in \bar{\Omega}_{n-1}$ is a basis of $\bar{V}_{n-1}$. Recall that the functions $\bar{\varphi}_{n,k}$, $k = {}^+_- 1, \ldots, {}^+_- 2^{n-1}$ form a basis of $\bar{W}_n$.

Assume that $v \in \bar{V}_{n-1}$ and $w \in \bar{W}_n$. There are real values b_k and c_k such that

$$v = \sum_{k=-2^{n-1}+1}^{2^{n-1}} c_k v_{k2^{-(n-1)}}^{n-1} \quad \text{and} \quad w = \sum_{k=1}^{2^{n-1}} (b_k \bar{\varphi}_{n,k} + b_{-k} \bar{\varphi}_{n,-k}).$$

After some calculation, one obtains

$$\begin{aligned} & K|v+w|_{H^1}^2 - \left(|v|_{H^1}^2 + |w|_{H^1}^2\right) = \\ & = \frac{1}{50\,h} \sum_{k=-2^{n-1}+1}^{2^{n-1}} 256\, b_k^2 (K-1) + 25\,(b_{k-1}^2 + b_{k+1}^2)(K-1) + \\ & \quad 25\,(c_{k+1} - c_k)^2 (K-1) + 128\, b_k (b_{k-1} + b_{k+1})(K-1) + \\ & \quad 30\,(c_k - c_{k+1})(b_{k+1} - b_{k-1}) K + 14\, b_{k-1} b_{k+1} (K-1). \end{aligned}$$

Now, we see that contrary to the hiererchical basis, prewavelets do not naturally lead to pure local equations. Therefore, we introduce a real parameter β, to obtain local problems. Then, we get

$$\begin{aligned} & K|v+w|_{H^1}^2 - \left(|v|_{H^1}^2 + |w|_{H^1}^2\right) = \\ & = \frac{1}{50\,h} \sum_{k=-2^{n-1}+1}^{2^{n-1}} (256 - 2\beta)\, b_k^2 (K-1) + (25+\beta)\,(b_{k-1}^2 + b_{k+1}^2)(K-1) + \\ & \quad 25\,(c_{k+1} - c_k)^2 (K-1) + 128\, b_k (b_{k-1} + b_{k+1})(K-1) + \\ & \quad 30\,(c_k - c_{k+1})(b_{k+1} - b_{k-1}) K + 14\, b_{k-1} b_{k+1} (K-1). \end{aligned}$$

Assume that we can prove for suitable fixed parameters β and $K > 1$:

$$0 \leq \Psi_{\beta,K}(b_-, b, b_+, c, c_+), \tag{88}$$

for every $b_-, b, b_+, c, c_+ \in \mathbb{R}$, where

$$\begin{aligned} \Psi_{\beta,K}(b_-, b, b_+, c, c_+) \;:=\; & (256 - 2\beta)\, b^2 (K-1) + (25+\beta)\,(b_-^2 + b_+^2)(K-1) + \\ & 25\,(c_+ - c)^2 (K-1) + 128\, b (b_- + b_+)(K-1) + \\ & 30\,(c - c_+)(b_+ - b_-) K + 14\, b_- b_+ (K-1). \end{aligned}$$

Then, this implies

$$0 \leq K|v+w|_{H^1}^2 - \left(|v|_{H^1}^2 + |w|_{H^1}^2\right).$$

Now, we can apply Lemma 3. This gives

$$\gamma(\bar{V}_{n-1}, \bar{W}_n, H^1) \leq 1 - K^{-1}.$$

Therefore, it is enough to study (88).
STEP 3. *Algebraic analysis.*
To find suitable parameters β and K, one has to study the case $c = -c_+$ and $b_+ = -b_-$ in more detail. Then, by using an algebra manipulation program, we found that the choice

$$K = \frac{1}{16}(19+\sqrt{57}) \quad \text{and} \quad \beta = \frac{12(45+7\sqrt{57})}{11+\sqrt{57}}$$

leads to

$$\begin{aligned} 0 &\leq \left(\sqrt{\alpha}b + \frac{128}{2}\frac{K-1}{\sqrt{\alpha}}(b_- + b_+)\right)^2 + \\ &\quad \left((c-c_+)\sqrt{25(K-1)} + \frac{30K}{2}\frac{1}{\sqrt{25(K-1)}}(b_+ - b_-)\right)^2 \\ &= \Psi_{\beta,K}(b_-, b, b_+, c, c_+), \end{aligned}$$

where $\alpha = \frac{2(3+\sqrt{57})^3}{3(11+\sqrt{57})}$.
Combining STEP 1. - 3. shows that

$$\gamma(\mathring{V}_{n-1}, W_n, H^1) \leq \gamma(\bar{V}_{n-1}, \bar{W}_n, H^1) \leq \frac{3+\sqrt{57}}{19+\sqrt{57}} = \sqrt{\frac{3}{19}} \approx 0.39736.$$

q.e.d.

4.3.4 Generalized Prewavelets

The prewavelets in section 4.3.3 lead to L^2-orthogonal spaces $\mathring{W}_n$ and $\mathring{V}_n$. We showed that the constant in the strengthened Cauchy-Schwarz inequality between these subspaces of the Hilbert space $H^1(]0,1[)$ is smaller than

$$\sqrt{\frac{3}{19}} \approx 0.39736.$$

In several applications (e.g. anisotropic PDE's) it is not necessary that the space $\mathring{W}_n$ is L^2-orthogonal to the space $\mathring{V}_{n-1}$. Furthermore, the large support of the prewavelet functions increases the computational amount of a corresponding multilevel algorithm. Therefore, we construct a space $\mathring{W}'_n$ which is spanned by functions with a smaller support and which exactly has the properties we need. These properties are

- $\mathring{W}'_n$ is a subspace of $\mathring{V}_n$ such that

  $$\mathring{V}_n = \mathring{W}'_n \oplus \mathring{V}_{n-1}.$$

- The constant in the strengthened Cauchy-Schwarz inequality is as small as possible in the Hilbert space $L^2(]0,1[)$ and in the Hilbert space $H^1(]0,1[)$. This means that

  $$\max\left(\gamma(\mathring{V}_{n-1}, \mathring{W}'_n, H^1)\ ,\ \gamma(\mathring{V}_{n-1}, \mathring{W}'_n, L^2)\right)$$

  is as small as possible, where

  $$\begin{aligned}\gamma(\mathring{V}_{n-1}, \mathring{W}'_n, H^1) &:= \max_{v\in\mathring{V}_{n-1}, w\in\mathring{W}'_n} \frac{\int_0^1 \frac{\partial v}{\partial x}\frac{\partial w}{\partial x}\,dx}{|v|_{H^1}|w|_{H^1}} \quad \text{and} \\ \gamma(\mathring{V}_{n-1}, \mathring{W}'_n, L^2) &:= \max_{v\in\mathring{V}_{n-1}, w\in\mathring{W}'_n} \frac{\int_0^1 vw\,dx}{\|v\|_{L^2}\|w\|_{L^2}}.\end{aligned}$$

- $\mathring{W}'_n$ is spanned by functions with a small support.

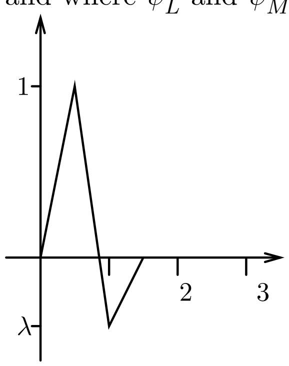
These considerations lead to the following construction of the space $\mathring{W}'_n$. Let $\mathring{W}'_n$ be the space spanned by the functions $\varphi'_{n,1}, \ldots, \varphi'_{n,2^{n-1}}$:

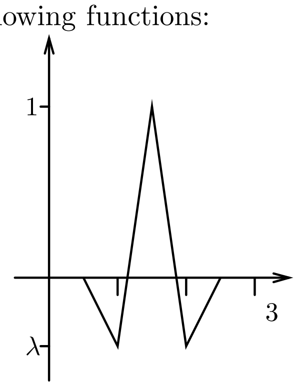
$$\mathring{W}'_n := \mathrm{span}_{\mathbb{R}}\left\{\varphi'_{n,k} \mid k = 1, \cdots, 2^{n-1}\right\}$$

where

$$\begin{aligned}\varphi'_{n,1}(x) &:= \psi'_L(x\,2^{n-1}), \\ \varphi'_{n,k}(x) &:= \psi'_M(x\,2^{n-1} - k + 2) \quad \text{for } k = 2, \ldots 2^{n-1} - 1, \text{ and} \\ \varphi'_{n,2^{n-1}}(x) &:= \psi'_L((1-x)\,2^{n-1}).\end{aligned}$$

and where ψ'_L and ψ'_M are the following functions:





By the construction of the space $\mathring{W}_n'$, we get the direct sum

$$\mathring{V}_n = \mathring{W}_n' \oplus \mathring{V}_{n-1}.$$

Let us call the functions $\varphi_{n,k}'$ generalized prewavelets. The support of these functions is smaller than the support of the prewavelets. Therefore, the evaluation of the generalized prewavelets cost less computational time than the evaluation of the prewavelets. Generalized prewavelets lead to a smaller constant in the strengthened Cauchy-Schwarz inequality than prewavelets, if the parameter λ is chosen in an optimal way. Figure 20 depicts an nearly sharp estimation of the constants $\gamma(\mathring{V}_{n-1}, \mathring{W}_n', H^1)$ and $\gamma(\mathring{V}_{n-1}, \mathring{W}_n', L^2)$. The optimal parameter λ is at the intersections in two graphs in Figure 20. A detailed analysis of the constants $\gamma(\mathring{V}_{n-1}, \mathring{W}_n', H^1)$ and $\gamma(\mathring{V}_{n-1}, \mathring{W}_n', L^2)$ leads two the following theorem the proof of which can be found in ... .

Theorem 7. *Put the parameter*

$$\lambda_{opt} = -0.442736$$

for the generalized prewavelets. Then, the following estimates hold

$$\gamma(\mathring{V}_{n-1}, \mathring{W}_n', H^1) \leq \frac{\lambda_{opt}}{\lambda_{opt} - 1} \leq 0.30688 \quad \textit{and} \quad \gamma(\mathring{V}_{n-1}, \mathring{W}_n', L^2) \leq \frac{\lambda_{opt}}{\lambda_{opt} - 1} \leq 0.30688.$$

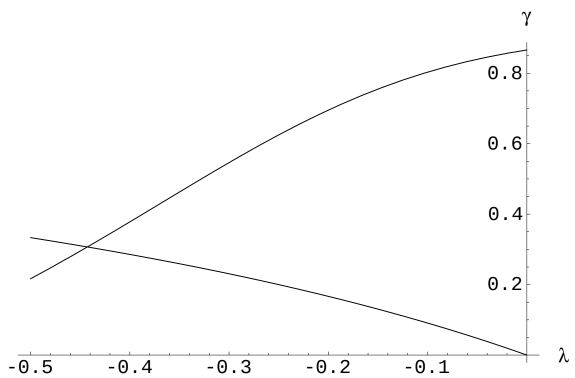


Figure 20: Estimation of the constants in strengthened Cauchy-Schwarz inequality for spaces spanned by generalized prewavelets

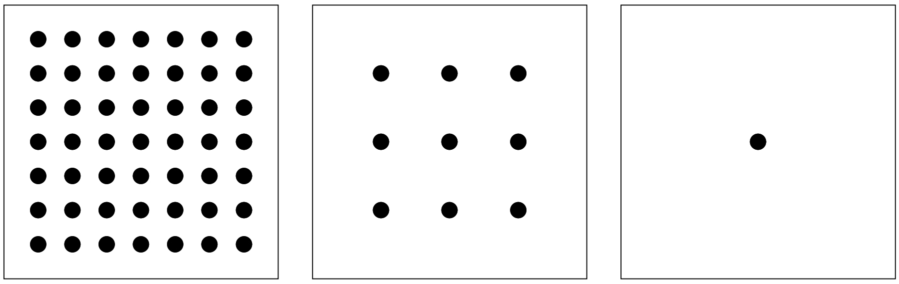


Figure 21: Semi-coarsening of a grid.

4.3.5 2D-Splittings by Prewavelets and Generalized Prewavelets

Let us recall the tensor product construction of two subspaces $V, W \subset H^1(]0,1[)$

$$V \otimes W := \operatorname{span}_{\mathbb{R}}\{v(x) \cdot w(y) \mid v \in V \quad \text{and} \quad w \in W\} \subset H^1(]0,1[^2).$$

Now, consider the fine-grid space

$$\mathcal{V}_n := \mathring{V}_n \otimes \mathring{V}_m,$$

where $m, n > 1$ are two parameters.
There are two ways to construct a coarse-grid space $\mathcal{V}_{n-1}$ of $\mathcal{V}_n$:

- *Semi-coarsening.* (See Figure 21). Now, m is a fixed parameter and n is the parameter indicating the level. Thus, define $\mathcal{V}_{n-1}^{\text{semi}} := \mathring{V}_{n-1} \otimes \mathring{V}_m$.
- *2-directional coarsening.* (See Figure 2.1.1). Now, let $n = m$ be parameter indicating the level. Define $\mathcal{V}_{n-1} := \mathring{V}_{n-1} \otimes \mathring{V}_{n-1}$.

Let us first study the case of semi-coarsening. The complementary space $\mathcal{W}_n$ can be constructed by prewavelets or generalized prewavelets:

- *Prewavelets.* Define $\mathcal{W}_n^{\text{semi}} := \mathring{W}_n \otimes \mathring{V}_m$.
- *Generalized prewavelets.* Define $\mathcal{W}_n'^{\text{semi}} := \mathring{W}_n' \otimes \mathring{V}_m$, where we choose $\lambda = \lambda_{opt} = -0.442736$.

Let $b(y) \geq 0$ be a non-negative function such that the following integrals exist. The constants in the strengthened Cauchy-Schwarz inequality are printed in Table 5 and Table 6 for some bilinear forms. Let us prove the estimate

$$\gamma(\mathcal{V}_{n-1}^{\text{semi}}, \mathcal{W}_n^{\text{semi}}, a) \leq \sqrt{\frac{3}{19}}$$

$a(u,v)$	$\gamma(\mathcal{V}_{n-1}^{\text{semi}}, \mathcal{W}_n^{\text{semi}}, a) \leq$
$\int_\Omega \langle \nabla u, \nabla v \rangle \; d(x,y)$	$\sqrt{\frac{3}{19}} \approx 0.39736$
$\int_\Omega b(y) \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} \; d(x,y)$	$\sqrt{\frac{3}{19}} \approx 0.39736$
$\int_\Omega b(y) \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} \; d(x,y)$	0
$\int_\Omega b(y) \, u \, v \; d(x,y)$	0

Table 5: Constant in the strengthened Cauchy-Schwarz inequality for semi-coarsening and prewavelets.

$a(u,v)$	$\gamma(\mathcal{V}_{n-1}^{\text{semi}}, \mathcal{W}_n'^{\text{semi}}, a) \leq$
$\int_\Omega \langle \nabla u, \nabla v \rangle \; d(x,y)$	0.30688
$\int_\Omega b(y) \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} \; d(x,y)$	0.30688
$\int_\Omega b(y) \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} \; d(x,y)$	0.30688
$\int_\Omega b(y) \, u \, v \; d(x,y)$	0.30688

Table 6: Constant in the strengthened Cauchy-Schwarz inequality for semi-coarsening and generalized prewavelets.

for $a(u,v) = \int_\Omega b(y) \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} \; d(x,y)$. Let $u \in \mathcal{V}_n$ and $v \in \mathcal{W}_n^{\text{semi}}$. Observe, that by the tensor product construction, the functions

$$x \to u(x,y) \quad \text{and} \quad x \to v(x,y)$$

are contained in $\mathring{V}_{n-1}$ and $\mathring{W}_n$, respectively for every fixed $y \in [0,1]$. By Theorem 6, we get

$$\left| \int_0^1 \frac{\partial u}{\partial x}(x,y) \frac{\partial v}{\partial x}(x,y) \; dx \right| \leq \sqrt{\frac{3}{19}} \sqrt{\int_0^1 \left(\frac{\partial u}{\partial x}\right)^2 (x,y) \; dx} \sqrt{\int_0^1 \left(\frac{\partial v}{\partial x}\right)^2 (x,y) \; dx}$$

for every fixed $y \in [0,1]$. Thus, we obtain

$$\begin{aligned} \left| \int_\Omega b(y) \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} \; d(x,y) \right| &\leq \int_0^1 b(y) \left| \int_0^1 \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} \; dx \right| \; dy \\ &\leq \int_0^1 b(y) \sqrt{\frac{3}{19}} \sqrt{\int_0^1 \left(\frac{\partial u}{\partial x}\right)^2 \; dx} \sqrt{\int_0^1 \left(\frac{\partial v}{\partial x}\right)^2 \; dx} \; dy \\ &\leq \sqrt{\frac{3}{19}} \sqrt{\int_\Omega b(y) \left(\frac{\partial u}{\partial x}\right)^2 \; d(x,y)} \sqrt{\int_\Omega b(y) \left(\frac{\partial v}{\partial x}\right)^2 \; d(x,y)}. \end{aligned}$$

The other estimates in Table 5 and Table 6 can be proved in the same way. Now, let us study the case of a 2-directional coarsening. Assume that $n = m$. Then, it is $\mathcal{V}_n = \mathring{V}_n \otimes \mathring{V}_n$ and $\mathcal{V}_{n-1} = \mathring{V}_{n-1} \otimes \mathring{V}_{n-1}$. The complementary space $\mathcal{W}_n$ can be constructed by prewavelets or generalized prewavelets:

- *Prewavelets.* Define $\mathcal{W}_n := (\mathring{W}_n \otimes \mathring{V}_n) + (\mathring{V}_n \otimes \mathring{W}_n)$. We can write $\mathcal{W}_n$ in two possible ways as a direct sum:

$$\mathcal{W}_n := (\mathring{W}_n \otimes \mathring{V}_n) \oplus (\mathring{V}_{n-1} \otimes \mathring{W}_n) = (\mathring{W}_n \otimes \mathring{V}_{n-1}) \oplus (\mathring{V}_n \otimes \mathring{W}_n). \quad (89)$$

- *Generalized prewavelets.* Define $\mathcal{W}'_n := (\mathring{W}'_n \otimes \mathring{V}_n) + (\mathring{V}'_n \otimes \mathring{W}_n)$, where we choose $\lambda = \lambda_{opt} = -0.442736$. We can write $\mathcal{W}'_n$ in two possible ways as a direct sum:

$$\mathcal{W}'_n = (\mathring{W}'_n \otimes \mathring{V}_n) \oplus (\mathring{V}_{n-1} \otimes \mathring{W}'_n) = (\mathring{W}'_n \otimes \mathring{V}_{n-1}) \oplus (\mathring{V}_n \otimes \mathring{W}'_n). \quad (90)$$

Let us prove the first equation in (89). The other equations in (90) and (89) follow by the same arguments. By $\mathring{V}_n = \mathring{V}_{n-1} \oplus \mathring{W}_n$, we get

$$\mathcal{W}_n = (\mathring{W}_n \otimes \mathring{V}_n) + (\mathring{V}_n \otimes \mathring{W}_n) = (\mathring{W}_n \otimes \mathring{V}_{n-1}) + (\mathring{W}_n \otimes \mathring{W}_n) + (\mathring{V}_{n-1} \otimes \mathring{W}_n).$$

By a dimension argument, the last sum in this equation must be a direct sum. Therefore, we get

$$\mathcal{W}_n = (\mathring{W}_n \otimes \mathring{V}_{n-1}) \oplus (\mathring{W}_n \otimes \mathring{W}_n) \oplus (\mathring{V}_{n-1} \otimes \mathring{W}_n) = (\mathring{W}_n \otimes \mathring{V}_n) \oplus (\mathring{V}_{n-1} \otimes \mathring{W}_n).$$

This shows the first equation in (89).

$a(u,v)$	$\gamma(\mathcal{V}_{n-1}, \mathcal{W}_n, a) \leq$
$\int_\Omega \langle \nabla u, \nabla v \rangle \; d(x,y)$	$\sqrt{\frac{3}{19}} \approx 0.39736$
$\int_\Omega \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} \; d(x,y)$	$\sqrt{\frac{3}{19}} \approx 0.39736$
$\int_\Omega \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} \; d(x,y)$	$\sqrt{\frac{3}{19}} \approx 0.39736$
$\int_\Omega u\, v \; d(x,y)$	0

Table 7: Constant in the strengthened Cauchy-Schwarz inequality for prewavelets.

$a(u,v)$	$\gamma(\mathcal{V}_{n-1}, \mathcal{W}'_n, a) \leq$
$\int_\Omega \langle \nabla u, \nabla v \rangle \; d(x,y)$	0.31[a]
$\int_\Omega \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} \; d(x,y)$	0.38 [a]
$\int_\Omega \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} \; d(x,y)$	0.38[a]
$\int_\Omega u\, v \; d(x,y)$	0.30[a]

Table 8: Constant in the strengthened Cauchy-Schwarz inequality for generalized prewavelets.

[a]numerical result

The constants in the strengthened Cauchy-Schwarz inequality of the above splittings are printed in Table 7 and Table 8 for some bilinear forms. Let us prove the estimate

$$\gamma(\mathcal{V}_{n-1}, \mathcal{W}_n, a) \leq \sqrt{\frac{3}{19}}$$

for $a(u,v) = \int_\Omega \frac{\partial u}{\partial x}\frac{\partial v}{\partial x}\, d(x,y)$. Assume that $u \in \mathcal{V}_{n-1}$ and $v = v_c + v_f \in \mathcal{W}_n$, where $v_c \in \mathring{W}_n \otimes \mathring{V}_{n-1}$ and $v_f \in \mathring{V}_n \otimes \mathring{W}_n$. By Theorem 6, we get

$$\begin{aligned}
\left|\int_\Omega \frac{\partial u}{\partial x}\frac{\partial v}{\partial x}\, d(x,y)\right| &= \left|\int_\Omega \frac{\partial u}{\partial x}\frac{\partial v_c}{\partial x}\, d(x,y)\right| \\
&\leq \sqrt{\frac{3}{19}}\sqrt{\int_\Omega \left(\frac{\partial u}{\partial x}\right)^2 d(x,y)}\sqrt{\int_\Omega \left(\frac{\partial v_c}{\partial x}\right)^2 d(x,y)} \\
&\leq \sqrt{\frac{3}{19}}\sqrt{\int_\Omega \left(\frac{\partial u}{\partial x}\right)^2 d(x,y)}\sqrt{\int_\Omega \left(\frac{\partial v}{\partial x}\right)^2 d(x,y)}.
\end{aligned}$$

The other estimates of Table 7 follow in the same way.
For the implementation a multilevel algorithm based on the space $\mathcal{W}'_n$ it is necessary to find basis functions of $\mathcal{W}'_n$ which support is as small as possible. To this end, we define the space

$$\mathring{W}_n^{\text{even}} = \operatorname{span}_{\mathbb{R}}\left\{ v_p^n \;\middle|\; p \in \Omega_{n-1} \right\}.$$

A short calculation shows the following idendity

$$\mathring{V}_n = \mathring{W}'_n \oplus \mathring{W}_n^{\text{even}}.$$

This leads to the splitting

$$\mathcal{W}'_n = \left(\mathring{W}'_n \otimes \mathring{W}_n^{\text{even}}\right) \oplus (\mathring{W}'_n \otimes \mathring{W}'_n) \oplus \left(\mathring{W}_n^{\text{even}} \otimes \mathring{W}'_n\right).$$

Using the natural basis of the spaces $\mathring{W}'_n$ and $\mathring{W}_n^{\text{even}}$ gives a natural tensor product construction for a basis of $\mathcal{W}_n$. This basis consists of functions with a small support.

4.4 Anisotropic Elliptic Differential Equation

Consider the anisotropic differential equation in section 1.1. Assume that the coefficients in the bilinear form

$$L(u) \;:=\; -\operatorname{div} A \operatorname{grad} u + cu = f \quad \text{on} \quad \Omega \subset \mathbb{R}^2, \quad \text{where}$$

are constant and are of the following form:

$$A = \begin{pmatrix} a_{1,1} & 0 \\ 0 & a_{2,2} \end{pmatrix} \in (L^1_{\text{loc}}(\Omega))^{2\times 2}, \qquad c \in L^1_{\text{loc}}(\Omega).$$

Then, the local mode analysis and the analysis of multigrid algorithms on a complementary space show, that a fast convergence rate can be obtained by the following choice of the coarse grid:

- $\frac{a_{1,1}}{h_x} << \frac{a_{2,2}}{h_y}$ semi-coarsening in y-direction,
- $\frac{a_{1,1}}{h_x} \approx \frac{a_{2,2}}{h_y}$ coarsening in xy-direction,
- $\frac{a_{1,1}}{h_x} >> \frac{a_{2,2}}{h_y}$ semi-coarsening in x-direction.

4.5 PDE's with Jumping Coefficients

Assume that $a(u, v)$ is a bilinear form with jumping coefficients. In general, the entries of the stiffness matrix $a(v_p^h, v_q^h)$ cannot be computed exactly. Then, numerical integration formulas have to be applied. The question is how to obtain the the coarse grid stiffness matrix. One choice is to calculate the stiffness matrix on the coarse grid by numerical integration formulas. This can lead to a very poor approximation of the stiffness matrix on the coarse grid, since the coarse grid integration formulas are of very low order. To understand this, consider a variable coefficient with a large jump on a very small domain. In such situations, a numerical integration of the stiffness matrix on the coarse grid can lead to a failure of the multigrid algorithm. A stable approach is to compute the coarse grid stiffness matrix A_{2h} by a restriction of the the fine grid stiffness matrix A_h according

$$A_{2h} = I_h^{2h} A_{2h} I_{2h}^h. \tag{91}$$

In case of a symmetric positive definite bilinear form $a(u, v)$, the stiffness matrix A_h is symmetric positive definite, too, if the integration formulas are accurate enough (This is simple to obtain in general). Then, a multiplicative subspace correction method will converge, since it is a minimizing algorithm. Nevertheless the convergence rate my be very poor. To understand this, consider the following 1D example

$$a_\epsilon(u, v) = \int_0^{0.25} \epsilon u'v'dx + \int_{0.25}^{0.5} u'v'dx + \int_{0.5}^{1} \epsilon u'v'dx.$$

Furthermore, let

- $v_{0.25}$ be the nodal basis function for linear elements of mesh size 0.25 and
- $v_{0.5}$ be the nodal basis function for linear elements of mesh size 0.5.

Now, consider

$$\gamma_\epsilon := \frac{|a_\epsilon(v_{0.25}, v_{0.5})|}{\|v_{0.25}\|\|v_{0.5}\|}.$$

Observe that

$$\gamma_1 = 0 \quad \text{and} \quad \lim_{\epsilon \to 0} \gamma_\epsilon = 1.$$

This implies, that a subspace correction method leads to a very poor convergence in case of the above bilinear form $a_\epsilon(u, v)$. The same hold for the classical multigrid algorithm. To improve the convergence one can improve the restriction prolongation operator or one can improve the smoothing. The smoothing can be improved by a block smoothing of neighbor points near the discontinuity.

4.6 PDE's with a Convection Term

Consider the convection-diffusion problem

$$-\triangle u + \vec{b}\nabla u + cu = f$$

with suitable boundary conditions and a stable discretization. This can be an upwind discretization of the convection term $\vec{b}\nabla u$ or a streamline diffusion discretization in case of finite elements. Unfortunately, the restriction of the stiffness matrix as in (91) leads to non-stable discretization of the convection term in case of a standard coarsening. There are three ways to avoid this problem:

- Let A_{2h} be the coarse grid discretization matrix (not very good convergence).
- Coarse orthogonal to the streamlines.
- Construct suitable restriction and prolongation operators.

In Section 5, we explain a suitable multigrid algorithm for convection diffusion problems.

4.7 Consequence for PDE's with a Kernel

Let $a : V \times V \to \mathbb{R}$ be a symmetric bilinear form on a Hilbert space V and $f \in V'$.

Definition 8. *The kernel of a is defined by*

$$\text{kern}(a) := \{u \in V \mid a(u, v) = 0 \ \forall v \in V\}.$$

Now, let us consider the problem

Problem 8. *Find $u \in V$ such that*

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

If $\text{kern}(a) \neq \{0\}$, then Problem 8 has no unique solution. Furthermore, Problem 8 has no solution, if $f\Big|_{\text{kern}(a)} \neq 0$.

To obtain a solution, we have to assume

$$f\Big|_{\mathrm{kern}(a)} = 0.$$

Furthermore, let us define the quotient space

$$V_a := V/\mathrm{kern}(a)$$

and let is define

$$\begin{aligned} \bar{a} : V_a \times V_a &\to \mathbb{R} \\ \bar{a}([u],[v]) &= a(u,v) \end{aligned}$$

Here, $u \in V$ is a representant of $[u] \in V_a$. Then, the following problem has a unique solution, if V_a is a finite dimensional vector space:

Problem 9. *Find $u \in V_a$ such that*

$$\bar{a}(u,v) = f(v) \qquad \forall v \in V.$$

Example 12. *Consider the Poisson's equation withe pure Neumann boundary conditions:*

$$\begin{aligned} -\triangle u &= f \quad \text{on } \Omega \\ \frac{\partial u}{\partial \vec{n}}\Big|_{\partial\Omega} &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

The bilinear form of the corresponding weak formulation is

$$a(u,v) = \int_\Omega \nabla u \nabla v \; d.$$

The local kernel of a is

$$\mathrm{kern}(a) = \mathbb{R}$$

the subspace of constant functions. Thus, there is a unique solution, if and only if $\int_\Omega \; d = 0$.

Example 13. *Let $E > 0$ and $0 < \nu < \frac{1}{2}$. Define the symmetric derivative*

$$\begin{aligned} \epsilon_{ij} &:= \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \\ Du &:= \begin{pmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{23} \end{pmatrix} \end{aligned}$$

and the matrix

$$\mathcal{C}^{-1} = \frac{1}{E}\begin{pmatrix} 1 & -\nu & -\nu & & & \\ -\nu & 1 & -\nu & & 0 & \\ -\nu & -\nu & 1 & & & \\ & & & 1+\nu & & \\ & 0 & & & 1+\nu & \\ & & & & & 1+\nu \end{pmatrix},$$

where E and ν are physical constants. The bilinear form corresponding to the problem of linear elasticity is

$$\begin{aligned} a : (H^1(\Omega))^3 \times (H^1(\Omega))^3 &\rightarrow \mathbb{R} \\ (u, v) &\mapsto \int_\Omega (Du)^T \mathcal{C} Dv \; d(x, y, z) \end{aligned}$$

The local kernel of this bilinear form is a 6-dimensional space of the ridged body modes:

$$\operatorname{kern}(a) = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} y \\ -x \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ z \\ -y \end{pmatrix}, \begin{pmatrix} -z \\ 0 \\ x \end{pmatrix} \right\}.$$

Let us assume, we want to construct a subspace correction method with a complementary space. This method is based on a decomposition

$$V_{a,i} = V_{a,i-1} + W_{a,i}.$$

Here, the quotient spaces $V_{a,i}$, $V_{a,i-1}$, and $W_{a,i}$ are based on a decomposition

$$V_i = V_{i-1} \oplus W_i$$

such that

$$V_{a,i} = \big\{[u] \mid u \in V_i\big\}, \quad V_{a,i-1} = \big\{[u] \mid u \in V_{i-1}\big\}, \quad W_i = \big\{[u] \mid u \in W_i\big\}.$$

Theorem 8. *If the constant*

$$\gamma = \sup_{v \in V_{a,i-1}, w \in W_{a,i}} \frac{\bar{a}(v, w)}{\|v\| \, \|w\|}$$

of the strengthened Cauchy-Schwarz inequality between $V_{a,i-1}$ and $W_{a,i}$ is smaller than 1, then

$$\operatorname{kern}(a) \subset W_i \;\; \text{or} \;\; \operatorname{kern}(a) \subset V_{i-1}\ .$$

Proof. Let us assume that $\text{kern}(a) \not\subset W_i$ and $\text{kern}(a) \not\subset V_{i-1}$. This means, there are vectors $0 \neq w \in W_i$ and $0 \neq v \in V_{i-1}$ such that

$$v + w \in \text{kern}(a).$$

This implies

$$[v] = -[w]$$

and

$$\frac{\bar{a}([v],[w])}{\|[v]\| \, \|[w]\|} = 1$$

Thus, we get $\gamma = 1$. □

In case of the Examples 12 and 13, it is very difficult to obtain that $\text{kern}(a) \subset W_i$ for a certain i and such that the corresponding iterative solver is an efficient solver. Thus, one has to construct the coarse grid spaces $V_1 \subset V_2 \subset ... \subset V_{l_{\max}}$ such that

$$\text{kern}(a) \subset V_1.$$

Thus, in case of Example 12, the constant functions must be conained in V_1 and in case of Example 13 the space of ridged body modes.

5 Algebraic Multigrid

5.1 General Description of AMG

Let A be a $n \times n$ matrix and let $b \in \mathbb{R}^n$. We want to solve the following problem:

Find $x_1 \in \mathbb{R}^{|\Omega_0|}$ such that

$$Ax = b.$$

Define

$$A_1 := A \quad \text{and} \quad b_1 = b.$$

Furthermore, let us denote

$$\Omega_1 = \{1, 2, ..., n\}$$

to be the finest grid. An algebraic multigrid constructs a sequence of coarser grids

$$\Omega_m \subset \Omega_{m-1} \subset ... \subset \Omega_1$$

and a restriction and prolongation operator

$$\begin{aligned} I_{k+1}^{k} &: \mathbb{R}^{|\Omega_k|} \to \mathbb{R}^{|\Omega_{k+1}|}, \\ I_{k}^{k+1} &: \mathbb{R}^{|\Omega_{k+1}|} \to \mathbb{R}^{|\Omega_k|} \\ I_{k+1}^{k} &= (I_{k}^{k+1})^T \end{aligned}$$

Then, define

$$A_{k+1} = I_k^{k+1} A_{k+1} I_{k+1}^k.$$

Using a relaxation method $\mathcal{S}_{l,b_l}$ (like the Gauss-Seidel relaxation) leads to the following AMG (algebraic multigrid):

$\underline{AGM(x_l^k, b_l, l)}$

If $l = m$, then $AMG(x_m^k, b_m, m) = A_m^{-1} b_m$

If $l < m$, then

Step 1 (v_1-pre–smoothing)

$x_l^{k,1} = \mathcal{S}_{l,b_l}^{\ v_1}(x_l^k)$

Step 2 (Coarse grid correction)

Residual : $r_l = b_l - A_l x_l^{k,1}$

Restriction : $r_{l+1} = I_l^{l+1} r_l$

Recursive call:

$e_{l+1}^0 = 0$

for $i = 1 \dots \mu$

$e_{l+1}^i = AGM(e_{l+1}^{i+1}, r_{l+1}, l+1)$

$e_{l+1} = e_{l+1}^\mu$

Prolongation : $e_l = I_{l+1}^l e_{l+1}$

Correction : $x_l^{k,2} = x_l^{k,1} + e_l$

Step 3 (v_2-post–smoothing)

$$AGM(x_l^k, b_l, l) = \mathcal{S}_{l,b_l}^{\ v_2}(x_l^{k,2})$$

5.2 Coarse Grid Construction of AMG

The original AMG by [16] is based on matrices which are weak diagonal dominant. These matrices have property

$$\sum_{i \not= j} |a_{ij}| \leq a_{ii}.$$

In case of a FD discretization of a PDE the entries a_{ij} are often negative. This motivates the following definition:

Definition 9 (Strong Connections)**.** *Let* $0 < \alpha < 1$ *be a small value (usually* $\alpha = 0.25$*).*
For each $i \in \Omega^k$ *define the set of strong connections by:*

$$S_i = \{j \in \Omega^k \mid -a_{ij} \geq \alpha \max_{k \neq i} -a_{ik}\}$$

Now, we construct the set of coarse grid points by the following algorithm:

<u>Coloring Sweep for Constructing Coarse Grid</u>

1. Assume that the set of fine grid points Ω_k is defined.
   Now construct the set of coarse grid points C and
   the set of fine grid points F as follows:
2. For each $i \in \Omega^k$, let $\lambda_i = |S_i|$
   (This is the number of strong connections).
3. Pick i with maximal λ_i, such that $i \notin C$ and $i \notin F$. Put i in C.
4. For each $j \in S_i \wedge j \notin C \wedge j \notin F$, put j in F.
   Increment λ_k for each $k \in S_j$.
5. If $\Omega^k \neq C \cup F$ go to 3. .
6. If $\Omega^k = C \cup F$ stop and let $\Omega^{k+1} = C$.

5.3 Interpolation of AMG

To construct an interpolation operator, let us use the notation:

$$\begin{aligned} N_i &= \{j \neq i \mid a_{ij} \neq 0\} \quad \text{(neighborhood of } i\text{)} \\ C_i &= S_i \cap C \\ D_i^s &= S_i \cap F \\ D_i^w &= \text{„weak“ connections such that:} \\ N_i &= C_i \cup D_i^s \cup D_i^w. \end{aligned}$$

Let us define a general interpolation operator as follows

$$(I_{k+1}^k)_i = \begin{cases} v_i^{k+1} & \text{if } \quad i \in C \\ \sum_{j \in C_i} w_{ij} v_j^{k+1} & \text{if } \quad i \in F. \end{cases}$$

Assume that (e_i) is the algebraic error. To derive an interpolation formula, we assume the following property of the algebraic error:

$$a_{ii} e_i \approx -\sum_{j \in N_i} a_{ij} e_j$$

for every $i \in F$. This property is motivated by two facts. First relaxation leads to na small residuum. Second, an exact correction on a complementary spanned by the fine grid points leads to the equation

$$a_{ii}\tilde{u}_i = -\sum_{j \in N_i} a_{ij}\tilde{u}_j + b_i$$

for every $i \in F$. Since,

$$a_{ii}u_i = -\sum_{j \in N_i} a_{ij}u_j + b_i.$$

the algebraic error $e_i = u_i - \tilde{u}_i$ satisfies

$$a_{ii}e_i = -\sum_{j \in N_i} a_{ij}e_j$$

for every $i \in F$. Thus, we get

$$\begin{array}{rll} a_{ii}e_i \approx & -\sum_{j \in C_i} a_{ij}e_j & \text{coarse strong points} \\ & -\sum_{l \in D_i^s} a_{il}e_l & \text{fine strong points} \\ & -\sum_{m \in D_i^w} a_{im}e_m & \text{weak points.} \end{array}$$

In this equation, we replace

- e_m by e_i and
- $e_l = \left(\sum_{j \in C_i} a_{lj}e_j\right) / \sum_{j \in C_i} a_{lj}$.

Thus, we construct the interpolation operator such that:

$$\begin{array}{rrl} a_{ii}e_i = & -\sum_{j \in C_i} a_{ij}e_j & \text{coarse strong points} \\ & -\sum_{l \in D_i^s} a_{il}\left(\sum_{j \in C_i} a_{lj}e_j\right) / \sum_{j \in C_i} a_{lj} & \text{fine strong points} \\ & -\sum_{m \in D_i^w} a_{im}e_i & \text{weak points.} \end{array}$$

This implies

$$e_i = \frac{-\sum_{j \in C_i}\left(a_{ij} + \sum_{l \in D_i^s} a_{il}a_{lj} / \sum_{k \in C_i} a_{lk}\right) e_j}{a_{ii} + \sum_{m \in D_i^w} a_{im}}$$

6 Appendix A: Hilbert spaces

The basic tool in the analysis of partial differential equations is the Hilbert space.

Definition 10 (Hilbert space)**.** *Let $\mathcal{H}$ be a real vector space. A bilinear form $\langle\cdot,\cdot\rangle : \mathcal{H}\times\mathcal{H}\to\mathbb{R}$ is called scalar product if the following rules hold:*

- $\langle v,w\rangle = \langle w,v\rangle$ *if $w\in\mathcal{H}$, $v\in\mathcal{H}$ (symmetric) .*
- $\langle v,v\rangle \geq 0$ *for all $v\in\mathcal{H}$ (positive).*
- $\langle v,v\rangle = 0$ *only if $v=0$ (definite).*

$\mathcal{H}$ is called a Hilbert space if $\mathcal{H}$ is complete with respect to the norm $\|\cdot\|$ defined by $\|v\| = \sqrt{\langle v,v\rangle}$.

The simplest example of a Hilbert space is the finite dimensional Euclidean vector space $\mathbb{R}^n$, $n\in\mathbb{N}$ with the Euclidean scalar product

$$\big\langle (v_i)_{1\leq i\leq n}, (w_i)_{1\leq i\leq n}\big\rangle := \sum_{i=1}^{n} v_i w_i.$$

Another example is the space $L^2(\Omega)$ of square integrable functions on a bounded open domain $\Omega\subset\mathbb{R}^d$, $d\in\mathbb{N}$ (see [15]). The scalar product on this space is

$$\big\langle v,w\big\rangle_{L^2} := \int_\Omega vw\,dz \quad \text{for every } v,w\in\mathcal{H}.$$

An elementary property of a scalar product is stated in the following theorem.

Theorem 9 (Cauchy Schwarz inequality)**.** *Let $\langle\cdot,\cdot\rangle$ be the scalar product of a Hilbert space $\mathcal{H}$ with norm $\|\cdot\|$. Then, the following inequality holds*

$$\langle v,w\rangle \leq \|v\|\,\|w\|, \quad v,w\in\mathcal{H}.$$

Proof. By the binomial formula $\langle a-b,a-b\rangle = \langle a,a\rangle + \langle b,b\rangle - 2\langle a,b\rangle$, we get

$$\begin{aligned} 0 &\leq \left\|\frac{v}{\|v\|}-\frac{w}{\|w\|}\right\|^2 = \left\|\frac{v}{\|v\|}\right\|^2 + \left\|\frac{w}{\|w\|}\right\|^2 - 2\left\langle\frac{v}{\|v\|},\frac{w}{\|w\|}\right\rangle = \\ &= 2-2\left\langle\frac{v}{\|v\|},\frac{w}{\|w\|}\right\rangle. \end{aligned}$$

This implies

$$\langle v,w\rangle \leq \|v\|\,\|w\|.$$

q.e.d.

By the Cauchy Schwarz inequality, the angle between vectors $v, w \neq 0$ is well-defined by

$$\angle(v, w) := \arccos \frac{|\langle v, w\rangle|}{\|v\| \, \|w\|},$$

since the fraction in this definition is ≤ 1. The *angle* between two subspaces $\mathcal{V}$ and $\mathcal{W}$ of a Hilbert space $\mathcal{H}$ is defined by

$$\angle(\mathcal{V}, \mathcal{W}) := \arccos \sup_{v\in\mathcal{V}, w\in\mathcal{W}} \frac{\langle v, w\rangle}{\|v\| \, \|w\|}.$$

Let us define the constant

$$\gamma(\mathcal{V}, \mathcal{W}) := \sup_{v\in\mathcal{V}, w\in\mathcal{W}} \frac{\langle v, w\rangle}{\|v\| \, \|w\|},$$

where we write $\frac{0}{0} := 0$, for simplicity. By the Cauchy Schwarz inequality, it is $\gamma(\mathcal{V}, \mathcal{W}) \leq 1$. But for special subspaces the constant $\gamma(\mathcal{V}, \mathcal{W})$ may be smaller than 1. Then, we call $\gamma(\mathcal{V}, \mathcal{W})$ the *constant in the strengthened Cauchy Schwarz inequality* between $\mathcal{V}$ and $\mathcal{W}$, since the following inequality holds

$$\langle v, w\rangle \ \leq \ \gamma(\mathcal{V}, \mathcal{W}) \, \|v\| \, \|w\|, \quad v \in \mathcal{V}, w \in \mathcal{W}.$$

A simple calculation shows that the strengthened Cauchy Schwarz inequality and the angle between subspaces satisfy the equation

$$\angle(\mathcal{V}, \mathcal{W}) = \arccos(\gamma(\mathcal{V}, \mathcal{W})). \tag{92}$$

The subspaces $\mathcal{V}$ and $\mathcal{W}$ of the Hilbert space $\mathcal{H}$ are called *orthogonal* iff $\gamma(\mathcal{V}, \mathcal{W}) = 0$.

Example 14. *Consider the Euclidean Hilbert space $\mathcal{H} = \mathbb{R}^3$ with the subspaces $\mathcal{V} := (1, 0, 1)\mathbb{R}$ and $\mathcal{W} := (1, 0, 0)\mathbb{R} + (0, 1, 0)\mathbb{R}$. Now, a simple calculation shows*

$$\gamma(\mathcal{V}, \mathcal{W}) = \frac{1}{\sqrt{2}} \quad \textit{and} \quad \angle(\mathcal{V}, \mathcal{W}) = 45^\circ.$$

In the above example, the calculation of the constant in the strengthened Cauchy Schwarz inequality is straight forward. But in case of higher dimensional spaces such a calculation can be more complicated. Then, the following lemma is very helpful.

Lemma 3. *Let $\mathcal{H}$ be a finite dimensional Hilbert space with scalar product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$. Furthermore, let $\mathcal{V}$ and $\mathcal{W}$ be subspaces of $\mathcal{H}$. Then, the following inequalities are equivalent for constants $K > 1$*

$$\begin{array}{rcl} \|v\|^2 + \|w\|^2 & \leq & K\|v+w\|^2, \quad v \in \mathcal{V}, w \in \mathcal{W} \\ & \Updownarrow & \\ \gamma(\mathcal{V},\mathcal{W}) & \leq & 1 - K^{-1}. \end{array}$$

Proof. "$\Downarrow$" Choose $v \in \mathcal{V}$ and $w \in \mathcal{W}$ such that $\|v\| = \|w\| = 1$. Then, we get

$$\begin{array}{rcl} \|v\|^2 + \|w\|^2 & \leq & K\|v+(-w)\|^2 \\ & \Downarrow & \\ 2\langle v,w\rangle K & \leq & (K-1)\,\|v\|^2 + (K-1)\,\|w\|^2 \\ & \Downarrow & \\ \langle v,w\rangle & \leq & \dfrac{K-1}{K}\,\|v\|\,\|w\|. \end{array}$$

This shows

$$\gamma(V,W) \leq 1 - K^{-1}.$$

"$\Uparrow$" Choose $v \in \mathcal{V}$ and $w \in \mathcal{W}$. $\gamma(V,W) \leq 1 - K^{-1}$ implies that

$$-\langle v,w\rangle \leq \frac{K-1}{K}\,\|v\|\,\|w\|.$$

Therefore, we get

$$\begin{array}{rcl} 0 & \leq & (\|v\| - \|w\|)^2 \\ & \Downarrow & \\ 2\,\|v\|\,\|w\| & \leq & \|v\|^2 + \|w\|^2 \\ & \Downarrow & \\ -2\,\langle v,w\rangle\,K & \leq & (K-1)\,\|v\|^2 + (K-1)\,\|w\|^2 \\ & \Downarrow & \\ \|v\|^2 + \|w\|^2 & \leq & K\|v+w\|^2. \end{array}$$

q.e.d.

In chapter **??**, we will apply this lemma to finite element spaces.
If different scalar products are given on a vector space, then the strengthened Cauchy Schwarz with respect to a scalar product a is defined by

$$\gamma(\mathcal{V},\mathcal{W},a) := \sup_{v\in\mathcal{V},w\in\mathcal{W}} \frac{a(v,w)}{\sqrt{a(v,v)}\sqrt{a(w,w)}}.$$

A basic lemma describing the constant $\gamma(\mathcal{V},\mathcal{W},a)$ for different bilinear forms is the following:

Lemma 4. *Consider two scalar products a and b on the vector space $\mathcal{H}$. Let λ_a and λ_b be positive constants and $\mathcal{V}$ and $\mathcal{W}$ two subspaces of $\mathcal{H}$. Then the following inequality holds*

$$\gamma(\mathcal{V},\mathcal{W},\lambda_a a+\lambda_b b)\leq \max\left(\gamma(\mathcal{V},\mathcal{W},a),\gamma(\mathcal{V},\mathcal{W},b)\right).$$

Beweis!! und Lemma spaeter zitieren!!

For the weak formulation of partial differential equations, we the the *dual space* $\mathcal{H}'$ of an Hilbert space $\mathcal{H}$. This space is the set of continuous and linear functions $f:\mathcal{H}\to\mathbb{R}$. The dual space is a vector space with norm

$$\|f'\|:=\sup_{v\in\mathcal{H}}\frac{f(v)}{\|v\|}.$$

The mapping $v\mapsto\langle w,v\rangle$ is an element of the dual space $\mathcal{H}'$ for every fixed $w\in\mathcal{H}$. Furthermore, every function $f\in\mathcal{H}'$ can be written as $v\mapsto\langle w,v\rangle$. This is stated in the following theorem.

Theorem 10 (Riesz representation theorem)**.** *Let $\mathcal{H}$ be a Hilbert space with scalar product $\langle\cdot,\cdot\rangle$. Then, for every $f\in\mathcal{H}'$ exists a unique $w\in\mathcal{H}$ such that*

$$\langle w,v\rangle=f(v)\quad\textit{for every}\quad v\in\mathcal{H}.$$

The proof of this theorem can be found in [15].
Now, let

$$\begin{aligned} a:\mathcal{H}\times\mathcal{H} &\to \mathbb{R}\\ (v,w) &\mapsto a(v,w)\end{aligned}$$

be a bounded and positive definite bilinear form. This means that a is a bilinear form and that there are constants $c,m>0$ such that

$$|a(v,w)|\leq c\|v\|\,\|w\|,\quad a(w,w)\geq m\|w\|^2\quad\text{for every } v,w\in\mathcal{H}.$$

The following theorem can be treated as a generalization of Theorem 10.

Theorem 11 (Lax-Milgram)**.** *Let $a(\cdot,\cdot)$ be a bounded and positive definite bilinear form on the Hilbert space $\mathcal{H}$. Then, for every $f\in\mathcal{H}'$ exists a unique $w\in\mathcal{H}$ such that*

$$a(w,v)=f(v)\quad\textit{for every}\quad v\in\mathcal{H}.$$

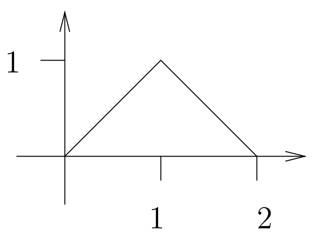
The proof of this theorem can be found in [5].

7 Appendix B: Sobolev spaces

Consider the piecewise linear function (see Figure 22)

$$w(x) = \begin{cases} 0 & \text{for } 0 \leq x \text{ and } x \geq 1, \\ x & \text{for } 0 \leq x \leq 1, \\ 2 - x & \text{for } 1 \leq x \leq 2. \end{cases}$$

The classical derivative $\partial w / \partial x$ of this function is well defined at every point $x \in \mathbb{R}$ with the exception of the points $x = 0$, $x = 1$ and $x = 2$. Therefore, we have to generalize the derivative of a function. Here, it is helpful that, in our applications, we are only interested in the derivative of w in the Hilbert space $L^2(\Omega)$. Therefore, we do not have to define the derivative of w at every point $x \in \mathbb{R}$. The wrong generalization of the derivative would be to define the derivative just by the classical derivative with the exception of a finite set of points. If we use such a concept, then the formula of partial integration would not hold. But in case of a continuous and piecewise differentiable function, this is the right generalization of the derivative of a function in the Hilbert space $L^2(\Omega)$. Therefore, the *generalized derivative* or *weak derivative* of the function in Figure 22 is the function in Figure 23.



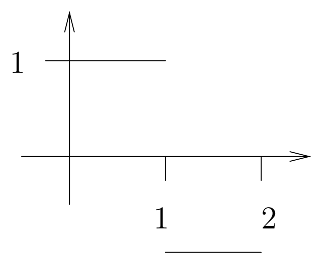


Figure 22: A piecewise linear function w. Figure 23: Weak derivative $\partial w / \partial x$ of w.

On general, we have to define the weak derivative of a function with the help of the partial integration. For reasons of completeness of our presentation, we briefly describe this concept.
To this end, let $\Omega \subset \mathbb{R}^d$ be an open and bounded domain and $\mathcal{C}_0^\infty(\Omega)$ the space of function f such that

- the classical partial derivative

  $$\frac{\partial^{\alpha_1 + \cdots + \alpha_d} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}$$

  exists and is continuous and

- there is a closed subset $\bar{\Omega}_0 \subset \Omega$ such that $f(z) = 0$ for every $z \in \Omega \backslash \bar{\Omega}_0$.

For the definition of the weak derivative, let us introduce the following abbreviation.
If $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_d) \in \mathbb{N}^d$ is a multiindex, then we write

$$|\boldsymbol{\alpha}| := \sum_{i=1}^{d} \alpha_i.$$

Now, we can define the weak derivative.

Definition 11 (Weak derivative)**.** *The function $g \in L^2(\Omega)$ is called the weak derivative of $f \in L^2(\Omega)$, if the following equation holds*

$$\int_\Omega g\,\varphi\,dz \;=\; (-1)^{|\boldsymbol{\alpha}|} \int_\Omega f \,\frac{\partial^{\alpha_1+\cdots+\alpha_d}\varphi}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}\,dz \quad \text{for every } \varphi \in \mathcal{C}_0^\infty(\Omega).$$

Then, we write

$$g =: \frac{\partial^{\alpha_1+\cdots+\alpha_d} f}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}} \in L^2(\Omega).$$

The most important tools in the analysis of partial differential equations are spaces with include the derivative of a function. The Sobolev space is such a space with an additional Hilbert space structure.

Definition 12 (Sobolev space)**.** *Let $\Omega \subset \mathbb{R}^d$ be an open and bounded domain. The Sobolev space $H^m(\Omega)$, $m \in \mathbb{N}$ is defined by*

$$H^m(\Omega) := \left\{ f \in L^2(\Omega) \;\middle|\; \frac{\partial^{\alpha_1+\cdots+\alpha_d} f}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}} \in L^2(\Omega) \text{ for every } \boldsymbol{\alpha} \in \mathbb{N}^d \text{ with } |\boldsymbol{\alpha}| \le m \right\}.$$

The scalar product on this space is defined by

$$\langle f, g\rangle_{H^m} := \int_\Omega \sum_{\boldsymbol{\alpha}\in\mathbb{N}^d,\,|\boldsymbol{\alpha}|\le m} \frac{\partial^{\alpha_1+\cdots+\alpha_d} f}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}\,\frac{\partial^{\alpha_1+\cdots+\alpha_d} g}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}\;dz.$$

This scalar product induces the norm

$$\|f\|_{H^m} := \sqrt{\sum_{\boldsymbol{\alpha}\in\mathbb{N}^d,\,|\boldsymbol{\alpha}|\le m} \left\| \frac{\partial^{\alpha_1+\cdots+\alpha_d} f}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}} \right\|_{L^2(\Omega)}^2 \; dz}.$$

Example 15. *In one dimension, the Sobolev space $H^1(]0,1[)$ is*

$$H^1(]0,1[) := \left\{ f \in L^2(]0,1[) \;\middle|\; \frac{\partial f}{\partial x} \in L^2(\Omega) \right\}.$$

In case of two dimensions, we obtain

$$H^1(]0,1[^2) := \left\{ f \in L^2(]0,1[^2) \;\middle|\; \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \in L^2(\Omega) \right\}.$$

In addition to the Sobolev space $H^1(\Omega)$, we have to define a Sobolev space with Dirichlet boundary conditions, which means that the functions in this space are zero at the boundary of Ω. This is the Sobolev space $H_0^1(\Omega) \subset H^1(\Omega)$. Usually this space is defined by the closure of the space $\mathcal{C}_0^\infty(\Omega)$. Here, we give an equivalent definition of the space $H_0^1(\Omega)$. For our purpose this definition is more convenient, but it is restricted to polygon domains.

Definition 13 (Sobolev space with Dirichlet boundary conditions)**.**
1. Let $\Omega \subset \mathbb{R}^1$ be an interval. Define the space

$$\begin{aligned}\tilde{\mathcal{C}}_0^1(\Omega) \quad := \quad & \Big\{ v \in \mathcal{C}(\bar{\Omega}) \ \big| \ \textit{there are intervals } I_1, \cdots, I_k \textit{ such that } \cup_{i=1}^k I_i = \Omega, \\ & v|_{I_i} \in \mathcal{C}^1(\bar{I}_i) \textit{ for every } i = 1, \cdots, k, \textit{ and } v|_{\partial\Omega} = 0 \Big\}.\end{aligned}$$

The Sobolev space $H_0^1(\Omega)$, is defined by

$$H_0^1(\Omega) := \overline{\tilde{\mathcal{C}}_0^1(\Omega)}^{H^1}.$$

2. Let $\Omega \subset \mathbb{R}^2$ be an open, bounded and polygon domain. Define the space

$$\begin{aligned}\tilde{\mathcal{C}}_0^1(\Omega) \quad := \quad & \Big\{ v \in \mathcal{C}(\bar{\Omega}) \ \big| \ \textit{there are triangles } T_1, \cdots, T_k \textit{ such that } \cup_{i=1}^k T_i = \Omega, \\ & v|_{T_i} \in \mathcal{C}^1(\bar{T}_i) \textit{ for every } i = 1, \cdots, k, \textit{ and } v|_{\partial\Omega} = 0 \Big\}.\end{aligned}$$

The Sobolev space $H_0^1(\Omega)$, is defined by

$$H_0^1(\Omega) := \overline{\tilde{\mathcal{C}}_0^1(\Omega)}^{H^1}.$$

Observe that the space $H_0^1(\Omega)$ is well defined, since the space $\tilde{\mathcal{C}}_0^1(\Omega)$ is a subspace of $H^1(\Omega)$.
For the analysis of Poisson's equations with homogeneous Dirichlet conditions (see section **??**), we need the Sobolev space $H_0^1(\Omega)$. This space can be equipped with another Hilbert space structure. The corresponding scalar product and norm are

$$\langle f, g \rangle_{H_0^1} \quad := \quad \int_\Omega \sum_{i=1}^d \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_i} \, dz \ = \ \int_\Omega \nabla f \nabla g \, dz, \tag{93}$$

$$|f|_{H^m} \quad := \quad \sqrt{\int_\Omega \sum_{i=1}^d \left\| \frac{\partial f}{\partial x_i} \right\|_{L^2(\Omega)}^2 dz} \quad = \sqrt{\|\nabla f\|_{L^2(\Omega)}^2 \, dz}. \tag{94}$$

Here, we use the abbreviation

$$\nabla w := \left(\frac{\partial w}{\partial x_1}, \cdots, \frac{\partial w}{\partial x_d} \right) \quad \text{and} \quad \nabla w \nabla v := \sum_{i=1}^d \frac{\partial w}{\partial x_i} \frac{\partial v}{\partial x_i}.$$

Nevertheless, the two Hilbert space structures on $H_0^1(\Omega)$ lead to equivalent norms.

Lemma 5 (Poincaré's inequality)**.** *There is a constant $c > 0$ which only depends on the bounded domain Ω such that*

$$c^{-1}|f|_{H^1} \leq \|f\|_{H^1} \leq c|f|_{H^1} \quad \text{for every } f \in H_0^1(\Omega).$$

The proof of this lemma can be found in [19].
Observe that the bilinear form $\langle f, g\rangle_{H_0^1}$ is not a scalar product on $H^1(\Omega)$, since this bilinear form is not definite for constant functions.

References

[1] O. Axelsson and P.S. Vassilevski. Algebraic multilevel preconditioning methods. I. *Numer. Math.*, 56:157–177, 1989.

[2] O. Axelsson and P.S. Vassilevski. Algebraic multilevel preconditioning methods. II. *SIAM J. Numer. Anal.*, 27(6):1569–1590, 1990.

[3] D. Braess. The contraction number of a multigrid method for solving the Poisson equation. *Numer. Math.*, 37:387–404, 1981.

[4] D. Braess. The convergence rate of a multigrid method with Gauss-Seidel relaxation for the Poisson equation. *Math. Comp.*, 42(166):505–519, 1984.

[5] D. Braess. *Finite Elemente.* Springer, New York, Berlin, Heidelberg, 1997. English translation available.

[6] A. Brandt. Guide to multigrid development. In A.Dold and E.Eckmann, editors, *Multigrid Methods, Proceedings, Köln-Porz, 1981*, number 960 in Lecture Notes in Mathematics, pages 220–312. Springer, New York, Berlin, Heidelberg, 1982.

[7] William L. Briggs, Van Emden Henson, and Steve F. McCormick. *A Multigrid Tutorial, Second Edition.* Frontiers in Applied Mathematics. SIAM, Philadelphia, 2000.

[8] M. Dobrowolski. Multigrid methods for elliptic systems with constraints. In *Proceedings of the XIth summer school Software and Algorithms of Numerical Mathematics*, pages 79–103, 1995.

[9] W. Hackbusch. *Multi-Grid Methods and Applications.* Springer, 1985.

[10] W. Hackbusch. *Iterative Lösung großer schwachbesetzter Gleichungssysteme.* Teubner, Stuttgart, 1991.

[11] J.F. Maitre and F. Musy. The contraction number of a class of two-level methods; an exact evaluation for some finite element subspaces and model problems. In *Multigrid methods*, number 960 in Lecture notes in mathematics, pages 535–544, 1981.

[12] C. Pflaum. Fast and robust multilevel algorithms. Habilitationsschrift, Universität Würzburg, 1999.

[13] C. Pflaum. Robust convergence of multilevel algorithms for convection-diffusion equations. *SIAM J. Numer. Anal.*, 37(2):443–469, 2000.

[14] M. Renardy and R. C. Rogers. *An Introduction to Partial Differential Equations*. Springer, Berlin, Heidelberg, New York, Tokyo, 1993.

[15] W. Rudin. *Real and Complex Analysis*. McGraw-Hill Book Company, 1984. Deutsche Übersetzung bei Oldenbourg.

[16] J.W. Ruge and K. Stüben. Algebraic Multigrid. In S. F. McCormick, editor, *Multigrid Methods*, volume 3 of *Frontiers in Applied Mathematics*, pages 73–130. SIAM, Philadelphia, 1987.

[17] R. Siegel and J. Howell. *Thermal Radiation Heat Transfer*. Hemisphere Publishing Corporation, 1992.

[18] P. Wesseling. *An Introduction to Multigrid Methods*. John Wiley & Sons, 1992.

[19] J. Wloka. *Partielle Differentialgleichungen*. Teubner, Stuttgart, 1982.