Chapter 6

The Multigrid Method

Remark 6.1 *Motivation.* The two-level method leaves an open question: How to solve the coarse grid equation

$$A^{2h}\mathbf{e}^{2h} = I_h^{2h}\left(\mathbf{r}^h\right) =: \mathbf{r}^{2h} \tag{6.1}$$

efficiently? The answer might be apparent: by a two-level method. The form (6.1) is not much different from the original problem. Thus, if one applies the two-level method to the original equation, its application to (6.1) should be easy. A recursive application of this idea, of using the two-level method for solving the coarse grid equation, leads to the multigrid method. □

6.1 Multigrid Cycles

Remark 6.2 *Notations.* To simplify the notations, the right-hand side vector of the residual equation will be denoted by $\mathbf{f}^{2h}$ instead of $\mathbf{r}^{2h}$ since it is just another right-hand side vector. The solution vector on the finest grid will be denoted by $\mathbf{u}^h$ and the current iterate by $\mathbf{v}^h$. Instead of denoting the solution vector on the coarse grid by $\mathbf{e}^{2h}$, it will be denoted by $\mathbf{v}^{2h}$. These notations can be used in an implementation of the method. □

Example 6.3 *A multigrid method.* Now, the two-level method will be imbedded into itself. It will be assumed that there are $l+1$ grids, $l \geq 0$, where the finest grid has the grid spacing h and the grid spacing increase by the factor 2 for each coarser grid. Let $L = 2^l$.
- Apply the smoother ν_1 times to $A^h\mathbf{u}^h = \mathbf{f}^h$ with the initial guess $\mathbf{v}^h$. The results is denoted by $\mathbf{v}^h$.
- Compute $\mathbf{f}^{2h} = I_h^{2h}\mathbf{r}^h = I_h^{2h}\left(\mathbf{f}^h - A^h\mathbf{v}^h\right)$.
  - ○ Apply the smoother ν_1 times to $A^{2h}\mathbf{u}^{2h} = \mathbf{f}^{2h}$ with the initial guess $\mathbf{v}^{2h} = \mathbf{0}$. Denote the result by $\mathbf{v}^{2h}$.
  - ○ Compute $\mathbf{f}^{4h} = I_{2h}^{4h}\mathbf{r}^{2h} = I_{2h}^{4h}\left(\mathbf{f}^{2h} - A^{2h}\mathbf{v}^{2h}\right)$.

$$\vdots$$

  - − Solve $A^{Lh}\mathbf{u}^{Lh} = \mathbf{f}^{Lh}$.

$$\vdots$$

  - ○ Correct $\mathbf{v}^{2h} := \mathbf{v}^{2h} + I_{4h}^{2h}\mathbf{v}^{4h}$.
  - ○ Apply smoother ν_2 times to $A^{2h}\mathbf{u}^{2h} = \mathbf{f}^{2h}$ with the initial guess $\mathbf{v}^{2h}$.
- Correct $\mathbf{v}^h := \mathbf{v}^h + I_{2h}^h\mathbf{v}^{2h}$.
- Apply the smoother ν_2 times to $A^h\mathbf{u}^h = \mathbf{f}^h$ with the initial guess $\mathbf{v}^h$.

$\square$

Example 6.4 *Multigrid method with γ-cycle.* The multigrid scheme from Example 6.3 is just one possibility to perform a multigrid method. It belongs to a family of multigrid methods, the so-called multigrid methods with γ-cycle that have the following compact recursive definition:

$\mathbf{v}^h \leftarrow M_\gamma^h(\mathbf{v}^h, \mathbf{f}^h)$

1. Pre smoothing: Apply the smoother ν_1 times to $A^h \mathbf{u}^h = \mathbf{f}^h$ with the initial guess $\mathbf{v}^h$.
2. If Ω^h is the coarsest grid
   − solve the problem.
   else
   − Restrict to the next coarser grid: $\mathbf{f}^{2h} \leftarrow I_h^{2h} \left(\mathbf{f}^h - A^h \mathbf{v}^h \right)$.
   − Set initial iterate on the next coarser grid: $\mathbf{v}^{2h} = \mathbf{0}$.
   − If Ω^h is the finest grid, set $\gamma = 1$.
   − Call the γ-cycle scheme γ times for the next coarser grid:

$$\mathbf{v}^{2h} \leftarrow M_\gamma^{2h} \left(\mathbf{v}^{2h}, \mathbf{f}^{2h} \right).$$

3. Correct with the prolongated update: $\mathbf{v}^h \leftarrow \mathbf{v}^h + I_{2h}^h \mathbf{v}^{2h}$.
4. Post smoothing: Apply the smoother ν_2 times to $A^h \mathbf{u}^h = \mathbf{f}^h$ with the initial guess $\mathbf{v}^h$.

In practice, only $\gamma = 1$ (V-cycle) and $\gamma = 2$ (W-cycle) are used. The names become clear if one has a look on how they move through the hierarchy of grids, see Figures 6.1 and 6.2. $\square$
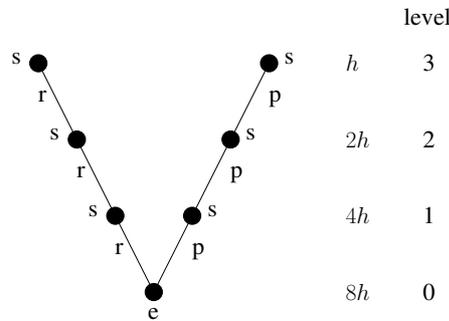

Figure 6.1: Multigrid V-cycle ($\gamma = 1$), s – smoothing, r – restriction, p – prolongated, e – exact solver.

Example 6.5 *Multigrid F-cycle.* In between the V-cycle and the W-cycle is the F-cycle, see Figure 6.3. The F-cycle starts with the restriction to the coarsest grid. In the prolongation process, after having reached each level the first time, again a restriction to the coarsest grid is performed. $\square$

Remark 6.6 *To the multigrid cycles.*
• The system on the coarsest grid is often small or even very small. Then, it can be solved efficiently with a direct method (Gaussian elimination, Cholesky factorization). Otherwise, one can apply a few steps of an iterative scheme to computed a sufficiently good approximate solution.
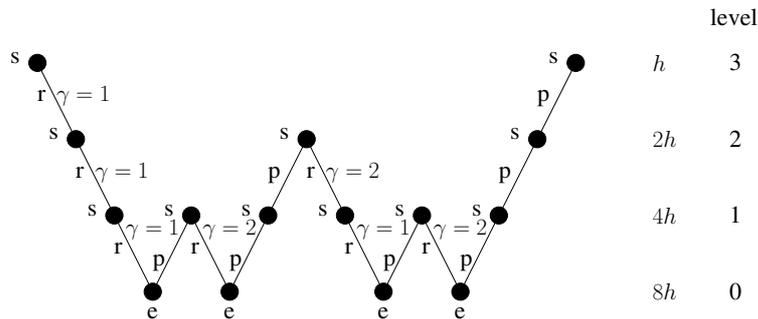
Figure 6.2: Multigrid W-cycle ($\gamma = 2$), s – smoothing, r – restriction, p – prolongated, e – exact solver.
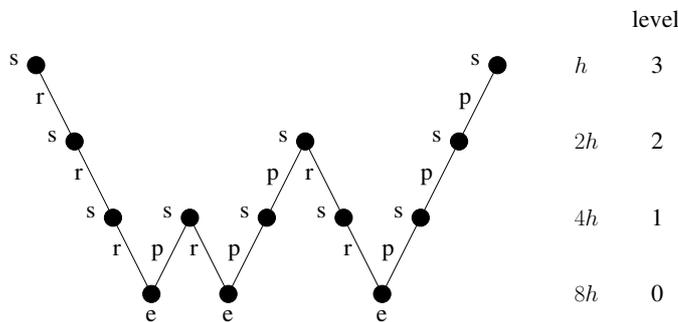

Figure 6.3: Multigrid F-cycle, s – smoothing, r – restriction, p – prolongated, e – exact solver.

- In our experience, it is sometimes (depending on the problem) helpful to damp the correction after having prolongated the update. Let $\beta \in (0, 1]$ be given, then instead of Step 3 of the multigrid γ-cycle, the update has the form

$$\mathbf{v}^h \leftarrow \mathbf{v}^h + \beta I_{2h}^h \mathbf{v}^{2h}.$$

- The initial guess for the first pre smoothing step on the finest grid can be obtained by a nested iteration, see Remark 4.5. In the nested iteration, the system is first solved (or smoothed) on a very coarse gird, then one goes to the next finer grid and smoothes the system on this grid and so on, until the finest grid is reached. This approach is called full multigrid. If one uses on each grid, which is not the finest grid, one multigrid V-cycle for smoothing, the so-called full multigrid V-cycle is performed, see Figure 6.4. The full multigrid V-cycle looks like a F-cycle without restriction and pre smoothing.

    In practice, one solves the systems on the coarser grids up to a certain accuracy before one enters the next finer grid.

    $\square$

6.2 Convergence of the W-cycle

Remark 6.7 *Contents.* It will be proved that the sufficient conditions for the convergence of the two-level method, Theorem 5.11, almost imply the convergence of the multigrid W-cycle. The rate of convergence will be bounded by a number $\rho(\nu) < 1$ which depends on the number of pre smoothing steps and which is independent of the finest step size h and of the number of levels involved in the multigrid scheme.
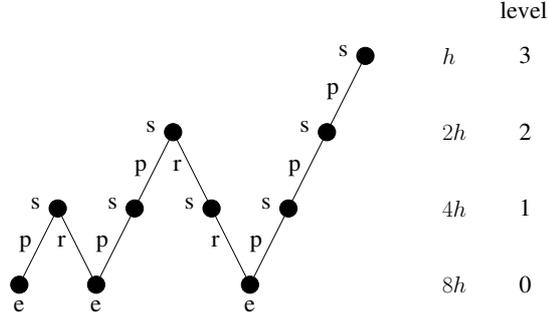
Figure 6.4: Full multigrid V-cycle, s – smoothing, r – restriction, p – prolongated, e – exact solver.

This technique cannot be applied to the multigrid V-cycle. The convergence theory for the V-cycle is more complicated and beyond the scope of this course. □

Remark 6.8 *Preliminarities.* As usual, one has to study the iteration matrix for the investigation of the convergence of an iterative solver. The levels of the multigrid hierarchy are numbered by $0, \ldots, l$, where level 0 is the coarsest grid. The iteration matrix of the two-level method on level l, where the corresponding mesh width should be h, is denoted by S_l and it has the form, see (5.5)

$$S_l(\nu) = \left(I - I_{l-1}^l \left(A_{l-1} \right)^{-1} I_l^{l-1} A_l \right) S_{\mathrm{sm},l}^\nu. \tag{6.2}$$

This iteration matrix is the matrix without post smoothing.

The solution of $A_l \mathbf{u}_l = \mathbf{f}_l$ is a fixed point of the multigrid γ-cycle. This statement follows from the fact that it is a fixed point of the two-level method, see Remark 5.6.
□

Lemma 6.9 Iteration matrix of the multigrid γ-cycle. *The iteration matrix of the multigrid γ-cycle scheme is given by*

$$
\begin{aligned}
S_{\mathrm{mg},l}(\nu) &= S_l(\nu) \quad \text{if } l = 1, \\
S_{\mathrm{mg},l}(\nu) &= S_l(\nu) + I_{l-1}^l \left(S_{\mathrm{mg},l-1}(\nu) \right)^\gamma A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu \quad \text{for } l \geq 2. \tag{6.3}
\end{aligned}
$$

Proof: For $l = 1$, the two-level method and the multigrid γ-cycle scheme are identical and the statement of the lemma follows immediately.

The proof for $l \geq 2$ will be performed by induction. Assume that (6.3) holds for $l - 1$. The iteration matrix $S_{\mathrm{mg},l}(\nu)$ can be written in the form

$$S_{\mathrm{mg},l}(\nu) = C_{\mathrm{mg},l} S_{\mathrm{sm},l}^\nu,$$

where $C_{\mathrm{mg},l}$ represents the iteration matrix of the complete coarse grid correction, i.e., everything which was done on the levels $0, \ldots, l - 1$. This matrix has to be determined. To this end, consider the multigrid method with $\mathbf{f}_l = \mathbf{0}$ and let $\mathbf{u}_l$ being arbitrary. For the restricted residual, it holds

$$\mathbf{f}_{l-1} = I_l^{l-1} \left(\mathbf{f}_l - A_l \mathbf{u}_l \right) = -I_l^{l-1} A_l \mathbf{u}_l.$$

Then, in the multigrid γ-cycle, γ iterates $\mathbf{v}_{l-1}^{(1)}, \ldots, \mathbf{v}_{l-1}^{(\gamma)}$ are computed, starting with the initial iterate $\mathbf{v}_{l-1}^{(0)} = \mathbf{0}$. The multigrid γ-cycle on level $l - 1$, which is applied to

$$A_{l-1} \mathbf{u}_{l-1} = \mathbf{f}_{l-1}, \tag{6.4}$$

can be described with the basic form of a fixed point iteration given in (3.3)

$$\mathbf{v}_{l-1}^{(j+1)} = S_{\mathrm{mg},l-1}(\nu) \mathbf{v}_{l-1}^{(j)} + N_{l-1} \mathbf{f}_{l-1}. \tag{6.5}$$

40

From Remark 6.8 it follows that the solution of (6.4) is the fixed point of (6.5). One obtains

$$
\begin{aligned}
\mathbf{v}_{l-1}^{(1)} &= S_{\mathrm{mg},l-1}(\nu)\mathbf{v}_{l-1}^{(0)} + N_{l-1}\mathbf{f}_{l-1} = N_{l-1}\mathbf{f}_{l-1} \\
\mathbf{v}_{l-1}^{(2)} &= S_{\mathrm{mg},l-1}(\nu)N_{l-1}\mathbf{f}_{l-1} + N_{l-1}\mathbf{f}_{l-1} \\
\mathbf{v}_{l-1}^{(3)} &= S_{\mathrm{mg},l-1}(\nu)\left(S_{\mathrm{mg},l-1}(\nu)N_{l-1}\mathbf{f}_{l-1} + N_{l-1}\mathbf{f}_{l-1}\right) + N_{l-1}\mathbf{f}_{l-1} \\
&= \left(S_{\mathrm{mg},l-1}(\nu)\right)^2 N_{l-1}\mathbf{f}_{l-1} + S_{\mathrm{mg},l-1}(\nu)N_{l-1}\mathbf{f}_{l-1} + N_{l-1}\mathbf{f}_{l-1} \\
&\ \vdots \\
\mathbf{v}_{l-1}^{(\gamma)} &= \sum_{k=0}^{\gamma-1}\left(S_{\mathrm{mg},l-1}(\nu)\right)^k N_{l-1}\mathbf{f}_{l-1} \\
&= \sum_{k=0}^{\gamma-1}\left(S_{\mathrm{mg},l-1}(\nu)\right)^k N_{l-1}\left(-I_l^{l-1}\left(A_l\mathbf{u}_l\right)\right).
\end{aligned}
\tag{6.6}
$$

Let $\mathbf{u}_{l-1}$ be the fixed point of (6.5) and the solution of (6.4), then it is

$$
\begin{aligned}
\mathbf{u}_{l-1} &= S_{\mathrm{mg},l-1}(\nu)\mathbf{u}_{l-1} + N_{l-1}\mathbf{f}_{l-1} = S_{\mathrm{mg},l-1}(\nu)\mathbf{u}_{l-1} + N_{l-1}A_{l-1}\mathbf{u}_{l-1} \\
&= \left(S_{\mathrm{mg},l-1}(\nu) + N_{l-1}A_{l-1}\right)\mathbf{u}_{l-1}.
\end{aligned}
$$

It follows that

$$
I = S_{\mathrm{mg},l-1}(\nu) + N_{l-1}A_{l-1}
$$

and

$$
N_{l-1} = \left(I - S_{\mathrm{mg},l-1}(\nu)\right)A_{l-1}^{-1}.
\tag{6.7}
$$

Using (telescopic sum)

$$
\sum_{k=0}^{\gamma-1} x^k(1-x) = \sum_{k=0}^{\gamma-1} x^k - \sum_{k=0}^{\gamma-1} x^{k+1} = 1 - x^\gamma,
$$

one obtains from (6.6) and (6.7)

$$
\begin{aligned}
\mathbf{v}_{l-1}^{(\gamma)} &= \left(\sum_{k=0}^{\gamma-1}\left(S_{\mathrm{mg},l-1}(\nu)\right)^k\left(I - S_{\mathrm{mg},l-1}(\nu)\right)\right) A_{l-1}^{-1}\left(-I_l^{l-1}\left(A_l\mathbf{u}_l\right)\right) \\
&= \left(I - \left(S_{\mathrm{mg},l-1}(\nu)\right)^\gamma\right) A_{l-1}^{-1}\left(-I_l^{l-1}\left(A_l\mathbf{u}_l\right)\right).
\end{aligned}
\tag{6.8}
$$

From the coarse grid correction, step 3 auf the multigrid γ-cycle scheme, see Example 6.4, it follows for the result of the multigrid γ-cycle that

$$
\mathbf{u}_l^{\mathrm{new}} := C_{\mathrm{mg},l}\mathbf{u}_l = \mathbf{u}_l + I_{l-1}^l\mathbf{v}_{l-1}^{(\gamma)}.
$$

Inserting (6.8), one obtains for the iteration matrix of the coarse grid correction

$$
\begin{aligned}
C_{\mathrm{mg},l} &= I + I_{l-1}^l\left(I - \left(S_{\mathrm{mg},l-1}(\nu)\right)^\gamma\right)A_{l-1}^{-1}\left(-I_l^{l-1}A_l\right) \\
&= I - I_{l-1}^l A_{l-1}^{-1}I_l^{l-1}A_l + I_{l-1}^l\left(S_{\mathrm{mg},l-1}(\nu)\right)^\gamma A_{l-1}^{-1}I_l^{l-1}A_l.
\end{aligned}
$$

Hence, the iteration matrix of the multigrid γ-cycle scheme is given by

$$
\begin{aligned}
S_{\mathrm{mg},l}(\nu) &= C_{\mathrm{mg},l}S_{\mathrm{sm},l}^\nu \\
&= \left(I - I_{l-1}^l A_{l-1}^{-1}I_l^{l-1}A_l\right)S_{\mathrm{sm},l}^\nu + I_{l-1}^l\left(S_{\mathrm{mg},l-1}(\nu)\right)^\gamma A_{l-1}^{-1}I_l^{l-1}A_l S_{\mathrm{sm},l}^\nu.
\end{aligned}
$$

The first term is equal to $S_l(\nu)$, see (6.2). Thus, (6.3) is proved for level l under the assumption that it holds for level $l-1$.

One can write the iteration matrix for $l = 1$ also in form (6.3), using the definition $S_{\mathrm{mg},0}(\nu) := 0$. Then, (6.3) holds for $l = 1$ and hence it holds for all $l \geq 1$ by induction. $\blacksquare$

Remark 6.10 *Estimate of the spectral norm of the iteration matrix.* The iteration matrix $S_{\mathrm{mg},l}(\nu)$ of the multigrid γ-cycle scheme is the sum of the iteration matrix of the two-level method and a perturbation. It will be shown that this perturbation is, under certain assumptions, small.

The spectral norm of $S_{\mathrm{mg},l}(\nu)$ will be estimated in a first step by the triangle inequality and the rule for estimating the norm of products of matrices

$$
\begin{aligned}
\left\| S_{\mathrm{mg},l}(\nu) \right\|_2 &\leq \left\| S_l(\nu) \right\|_2 + \left\| I_{l-1}^l \left(S_{\mathrm{mg},l-1}(\nu) \right)^\gamma A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu \right\|_2 \\
&\leq \left\| S_l(\nu) \right\|_2 + \left\| I_{l-1}^l \right\|_2 \left\| S_{\mathrm{mg},l-1}(\nu) \right\|_2^\gamma \left\| A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu \right\|_2. \quad (6.9)
\end{aligned}
$$

Now, bounds for all factors on the right-hand side of (6.9) are needed. □

Remark 6.11 *Assumptions on the prolongation operator.* It will be assumed that the prolongation is a bounded linear operator with a bound independent of l, i.e., there is a constant $\bar{c}_p$ such that

$$
\left\| I_{l-1}^l \right\|_2 \leq \bar{c}_p \quad \forall\, l \geq 1. \tag{6.10}
$$

In addition, a bound of $\left\| I_{l-1}^l \right\|_2$ from below will be needed. Thus, it will be assumed that there is a constant $\underline{c}_p > 0$ such that for all $\mathbf{u}_{l-1}$ defined on level $l-1$ it is

$$
\underline{c}_p^{-1} \left\| \mathbf{u}_{l-1} \right\|_2 \leq \left\| I_{l-1}^l \mathbf{u}_{l-1} \right\|_2 \quad \forall\, l \geq 1. \tag{6.11}
$$

The assumptions (6.10) and (6.11) are satisfied for the prolongation operator defined in Section 4.2. These properties can be deduced, e.g., by using the definition of the operator norm, *exercise.* □

Remark 6.12 *Assumptions on the smoother.* It will be assumed that there is a constant c_s such that

$$
\left\| S_{\mathrm{sm},l}^\nu \right\|_2 \leq c_s \quad \forall\, l \geq 1, 0 < \nu < \infty. \tag{6.12}
$$

This assumption is satisfied, e.g., for the damped Jacobi iteration, $S_{\mathrm{sm},l} = S_{\mathrm{jac},\omega}$ applied to the model problem, with $c_s = 1$. It was shown in the proof of Lemma 3.10 that $\rho(S_{\mathrm{jac},\omega}) < 1$. Since $S_{\mathrm{jac},\omega}$ is a symmetric matrix, it is $\left\| S_{\mathrm{jac},\omega} \right\|_2 = \rho(S_{\mathrm{jac},\omega})$. It follows that

$$
\left\| S_{\mathrm{sm},l}^\nu \right\|_2 = \left\| S_{\mathrm{sm},l} \right\|_2^\nu = \rho(S_{\mathrm{jac},\omega})^\nu < 1.
$$

□

Lemma 6.13 Estimate of last term in (6.9) with the iteration matrix of the two-level method. *Suppose* (6.11) *and* (6.12), *then*

$$
\left\| A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu \right\|_2 \leq \underline{c}_p \left(c_s + \left\| S_l(\nu) \right\|_2 \right). \tag{6.13}
$$

Proof: One gets with (6.11)

$$
\left\| A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu \mathbf{u}_l \right\|_2 \leq \underline{c}_p \left\| I_{l-1}^l A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu \mathbf{u}_l \right\|_2
$$

for all $\mathbf{u}_l$, where it is noted that $A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu \mathbf{u}_l$ is a vector on level $l-1$. Using the definition of an operator norm gives

$$
\left\| A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu \right\|_2 \leq \underline{c}_p \left\| I_{l-1}^l A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu \right\|_2. \tag{6.14}
$$

The right-hand side of this estimate can be rewritten as follows

$$
\begin{aligned}
I_{l-1}^l A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu &= S_{\mathrm{sm},l}^\nu - A_l^{-1} A_l S_{\mathrm{sm},l}^\nu + I_{l-1}^l A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu \\
&= S_{\mathrm{sm},l}^\nu - \left(A_l^{-1} - I_{l-1}^l A_{l-1}^{-1} I_l^{l-1} \right) A_l S_{\mathrm{sm},l}^\nu \\
&= S_{\mathrm{sm},l}^\nu - S_l(\nu).
\end{aligned}
$$

Using this identity in (6.14), applying the triangle inequality, and assumption (6.12) gives

$$
\begin{aligned}
\left\| A_{l-1}^{-1} I_l^{l-1} A_l S_{\mathrm{sm},l}^\nu \right\|_2 & \leq \underline{c}_p \left(\left\| S_{\mathrm{sm},l}^\nu \right\|_2 + \left\| S_l(\nu) \right\|_2 \right) \\
& \leq \underline{c}_p \left(c_s + \left\| S_l(\nu) \right\|_2 \right).
\end{aligned}
$$

$\blacksquare$

Remark 6.14 *Impact on estimate* (6.9). Only the case will be considered that the number ν of smoothing steps is sufficiently large such that the two-level method converges, i.e., it is

$$
\left\| S_l(\nu) \right\|_2 < 1.
$$

Inserting (6.13) into (6.9) and using the assumption on the number of smoothing steps yields, together with (6.10),

$$
\begin{aligned}
\left\| S_{\mathrm{mg},l}(\nu) \right\|_2 & \leq \left\| S_l(\nu) \right\|_2 + \underline{c}_p \left\| I_{l-1}^l \right\|_2 \left\| S_{\mathrm{mg},l-1}(\nu) \right\|_2^\gamma \left(c_s + \left\| S_l(\nu) \right\|_2 \right) \\
& \leq \left\| S_l(\nu) \right\|_2 + \underline{c}_p \bar{c}_p \left(c_s + 1 \right) \left\| S_{\mathrm{mg},l-1}(\nu) \right\|_2^\gamma \\
& = \left\| S_l(\nu) \right\|_2 + c^* \left\| S_{\mathrm{mg},l-1}(\nu) \right\|_2^\gamma.
\end{aligned}
\tag{6.15}
$$

This inequality is of the recursive form

$$
x_1 = x, \quad x_l \leq x + c^* x_{l-1}^\gamma, \quad l \geq 2,
\tag{6.16}
$$

with $x = \left\| S_l(\nu) \right\|_2 < 1$ and for $l = 1$ the multigrid and the two-level method coincide. $\square$

Lemma 6.15 Bound for the iterates of inequality (6.16). *Assume that $c^* \gamma > 1$. If $\gamma \geq 2$ and*

$$
x \leq x_{\max} := \frac{\gamma - 1}{\gamma} \left(c^* \gamma \right)^{-\frac{1}{\gamma - 1}},
$$

then every iterate of (6.16) *is bounded by*

$$
x_l \leq \frac{\gamma}{\gamma - 1} x < 1.
$$

Proof: The proof of the bound is performed by induction. For $l = 2$, one has

$$
\begin{aligned}
x_2 & \leq x + c^* x_1^\gamma \leq x + c^* x^\gamma = x \left(1 + c^* x^{\gamma - 1} \right) \\
& \leq x \left(1 + c^* x_{\max}^{\gamma - 1} \right) = x \left(1 + c^* \left(\frac{\gamma - 1}{\gamma} \right)^{\gamma - 1} \frac{1}{c^* \gamma} \right) \\
& = x \left(1 + \left(\frac{(\gamma - 1)^{\gamma - 1}}{\gamma^\gamma} \right) \right) = x \left(1 + \frac{1}{\gamma - 1} \left(1 - \frac{1}{\gamma} \right)^\gamma \right) \\
& \leq x \frac{\gamma - 1 + 1}{\gamma - 1} = x \frac{\gamma}{\gamma - 1},
\end{aligned}
$$

since $\left(1 - \frac{1}{\gamma} \right)^\gamma < 1$ (positive power of a real number in $(0,1)$).

Let the statement be already proved for $l - 1$, then one obtains with the assumption of the induction

$$
\begin{aligned}
x_l & \leq x + c^* x_{l-1}^\gamma \leq x + c^* \left(\frac{\gamma}{\gamma - 1} \right)^\gamma x^\gamma \\
& = x \left(1 + c^* \left(\frac{\gamma}{\gamma - 1} \right)^\gamma x^{\gamma - 1} \right) \leq x \left(1 + c^* \left(\frac{\gamma}{\gamma - 1} \right)^\gamma x_{\max}^{\gamma - 1} \right) \\
& = x \left(1 + \left(\frac{\gamma}{\gamma - 1} \right)^\gamma \left(\frac{\gamma - 1}{\gamma} \right)^{\gamma - 1} \frac{1}{\gamma} \right) = x \left(1 + \frac{1}{\gamma - 1} \right) = x \frac{\gamma}{\gamma - 1}.
\end{aligned}
$$

Using now the assumption on x and the assumption $c^* \gamma > 1$, one gets

$$\frac{\gamma}{\gamma - 1} x \leq \frac{\gamma}{\gamma - 1} x_{\max} = (c^* \gamma)^{-\frac{1}{\gamma - 1}} < 1.$$

$\blacksquare$

Remark 6.16 *To Lemma 6.15.* The condition $\gamma \geq 2$ is used in the definition of $x_{\max}$. Note that $x_{\max} < 1$ since both factors are lower than 1.

In the case of the W-cylce, i.e., $\gamma = 2$, the statement of the Lemma 6.15 implies

$$x \leq x_{\max} = \frac{1}{4c^*}, \quad x_l \leq 2x = 2 \|S_l(\nu)\|_2.$$

$\square$

Theorem 6.17 Convergence of the multigrid γ-cycle for $\gamma \geq 2$. *Suppose $\gamma \geq 2$, (6.10), (6.11), (6.12) with $\nu(h) = \infty$, and the assumptions of Theorem 5.11. Let $\rho \in (0, 1)$ be a fixed number. Then there is a number $\underline{\nu}$ such that*

$$\|S_{\mathrm{mg},l}(\nu)\|_2 \leq \rho < 1, \quad \|S_{\mathrm{mg},l}(\nu)\|_2 \leq \frac{\gamma}{\gamma - 1} C_a \eta(\nu), \quad (6.17)$$

whenever the number of smoothing iterations ν is larger or equal than $\underline{\nu}$. The estimates (6.17) are independent of the level l and the number of levels. The function $\eta(\nu)$ is defined in the smoothing property (5.7) and the constant C_a is defined in the approximation property.

Proof: Starting point of the proof is inequality (6.15). Lemma 6.15 will be applied with
$$x = \|S_l(\nu)\|_2, \quad x_l = \|S_{\mathrm{mg},l}(\nu)\|_2.$$
Without loss of generality, one can choose

$$c^* > \frac{1}{\gamma} \quad \Longleftrightarrow \quad c^* \gamma > 1.$$

In particular, c^* can be chosen so large that

$$x \leq \frac{\gamma - 1}{\gamma} (c^* \gamma)^{-\frac{1}{\gamma - 1}} \leq \frac{\gamma - 1}{\gamma} \rho < 1.$$

Note that large values of c^* imply small values of x, which can be always obtained by applying sufficiently many smoothing steps. Thus, the assumptions of Lemma 6.15 are satisfied and one obtains

$$\|S_{\mathrm{mg},l}(\nu)\|_2 \leq \frac{\gamma}{\gamma - 1} \|S_l(\nu)\|_2 = \frac{\gamma}{\gamma - 1} x \leq \rho.$$

The second estimate is obtained recursively. Using formally the same computations as in the proof of Lemma 6.15, one gets

$$\|S_{\mathrm{mg},2}(\nu)\|_2 \leq \frac{\gamma}{\gamma - 1} \|S_2(\nu)\|_2 \leq \frac{\gamma}{\gamma - 1} C_a \eta(\nu),$$

and by induction
$$\|S_{\mathrm{mg},l}(\nu)\|_2 \leq \frac{\gamma}{\gamma - 1} \|S_l(\nu)\|_2 \leq \frac{\gamma}{\gamma - 1} C_a \eta(\nu),$$
The details of this proof are an *exercise.* $\blacksquare$

Remark 6.18 *To Theorem 6.17.*
- The theorem states the convergence of the multigrid γ-cycle with a rate of convergence that is independent of the level. The estimate of this rate, i.e., $\frac{\gamma}{\gamma-1} C_a \eta(\nu)$, is in general somewhat pessimistic.

- A similar result can be proved if only post smoothing and no pre smoothing is applied, as well as in the case that both pre and post smoothing are used.
- The convergence proof for the V-cycle, i.e., $\gamma = 1$, does not rely on the convergence of the two-level method. In this proof, the multigrid iteration matrix is analyzed directly, e.g., see (Hackbusch, 1985, pp. 164).
- For problems without symmetric positive definite system matrix, multigrid works often quite well. But only very little is proved on the convergence of multigrid methods for such problems. Results on the multigrid convergence for problems without symmetric positive definite matrix are in general for problems which are only a slight perturbation of a s.p.d. problem. But many interesting problems are not small perturbations of a s.p.d. problem, like convection-dominated convection-diffusion equations or the Navier–Stokes equations. In these fields, many questions concerning the theory of multigrid methods are open. Some results for convection-diffusion problems can be found in Reusken (2002); Olshanskii and Reusken (2004).

$\square$

6.3 Computational Work of the Multigrid γ-Cycle

Remark 6.19 *Goal.* So far it is proved that the rate of convergence for the multigrid γ-cycle is bounded by a number $\rho < 1$ independently of the level. That means, the number of iterations for solving the equation up to a certain accuracy is bounded from above by a constant which is independent of the level, i.e., one needs on each grid level essentially the same number of iterations to solve the equation. This behavior is in contrast to the classical iteration schemes or the PCG method, where the number of iterations increases by the factor of 4 or 2, respectively, if the grid is refined once, cf. Table 2.1.

Let N_l the number of degrees of freedom on level l, $1 \leq l \leq L$. To obtain an optimal algorithm, one needs to show that the number of operations (flops) per multigrid cycle behaves like $\mathcal{O}(N_l)$. Since the number of multigrid cycles for the solution of the linear system is bounded uniformly, i.e., independently of l, it follows that then also the solution of the linear system requires $\mathcal{O}(N_l)$ operations. $\square$

Remark 6.20 *Assumptions on the computational costs of the components of the multigrid method.* The following bounds for the number of operations are assumed for the basic components of the multigrid method:
- one smoothing step $\mathbf{u}_l := S_l(\mathbf{u}_l)$

$$\text{flops} \ \leq c_s N_l, \quad l \geq 1,$$

- restriction $\mathbf{f}_{l-1} = I_l^{l-1}(\mathbf{f}_l - A_l \mathbf{v}_l)$

$$\text{flops} \ \leq c_r N_l, \quad l \geq 1,$$

- prolongation and correction $\mathbf{u}_l := \mathbf{u}_l + I_{l-1}^l \mathbf{v}_{l-1}$

$$\text{flops} \ \leq c_p N_l, \quad l \geq 1,$$

- coarsest grid problem $\mathbf{u}_0 = A_0^{-1} \mathbf{f}_0$

$$\text{flops} \ \leq c_0.$$

For sparse matrices and the prolongation and restriction which were introduced in Chapter 4, the bounds are true. The system on the coarsest grid can be solved, e.g.,

by Gaussian elimination. Then, c_0 depends on the number of degrees of freedom on the coarsest grid, but not on N_l.

Let
$$c_h = \sup_{l \geq 1} \frac{N_{l-1}}{N_l}.$$

For uniformly refined grids, i.e., $h_{l-1} = 2h_l$, this constant has the value $c_h = 2^{-d}$, where d is the dimension of the domain. $\qquad\square$

Theorem 6.21 Number of operations for the multigrid γ-cycle. *Set $\theta = c_h\gamma$, let $\theta < 1$ and let the assumptions from Remark 6.20 be satisfied. Then, one cycle of the multigrid γ-cycle with ν smoothing steps on each level requires $c_l N_l$ operations, where*
$$c_l < \frac{\nu c_s + c_r + c_p}{1 - \theta} + \theta^{l-1} \frac{c_0}{N_1}. \tag{6.18}$$

Proof: One iteration at level l involves γ^{l-k} iterations at level k, $1 \leq k \leq l$, since there are
- γ iterations on level $l - 1$,
- at each of these iterations, γ iterations on level $l - 2$, i.e., γ^2 iterations on level $l - 2$,
- and so on.

On level 0, γ^{l-1} coarsest grid systems have to be solved, since in each of the γ^{l-1} situations where one is on level 1, level 0 is called. Using the assumptions on the costs of the basic components of the multigrid method, on obtains the following costs

$$
\begin{aligned}
&(\nu c_s + c_r + c_p)\, N_l + \gamma\, (\nu c_s + c_r + c_p)\, N_{l-1} + \gamma^2\, (\nu c_s + c_r + c_p)\, N_{l-2} \\
&\qquad + \ldots + \gamma^{l-1}\, (\nu c_s + c_r + c_p)\, N_1 + \gamma^{l-1} c_0 \\
={} & (\nu c_s + c_r + c_p)\left(N_l + \gamma N_{l-1} + \ldots + \gamma^{l-1} N_1\right) + \gamma^{l-1} c_0 \\
={} & (\nu c_s + c_r + c_p)\, N_l \left(1 + \gamma \frac{N_{l-1}}{N_l} + \ldots + \gamma^{l-1} \frac{N_1}{N_l}\right) + \gamma^{l-1} c_0 \\
={} & (\nu c_s + c_r + c_p)\, N_l \left(1 + \gamma \frac{N_{l-1}}{N_l} + \gamma^2 \frac{N_{l-2}}{N_{l-1}} \frac{N_{l-1}}{N_l} + \ldots\right) + \gamma^{l-1} c_0 \\
\leq{} & (\nu c_s + c_r + c_p)\, N_l \left(1 + \gamma c_h + \gamma^2 c_h^2 + \ldots + \gamma^{l-1} c_h^{l-1}\right) + \gamma^{l-1} c_0 \\
={} & (\nu c_s + c_r + c_p)\, N_l \left(1 + \theta + \theta^2 + \ldots + \theta^{l-1}\right) + \gamma^{l-1} c_0 \tag{6.19} \\
\leq{} & (\nu c_s + c_r + c_p)\, \frac{N_l}{1 - \theta} + \theta^{l-1} \frac{c_0}{c_h^{l-1}} \frac{N_l}{N_l} \\
\leq{} & \left(\frac{\nu c_s + c_r + c_p}{1 - \theta} + \theta^{l-1} \frac{c_0}{N_1}\right) N_l,
\end{aligned}
$$

since $c_h^{l-1} \geq N_1/N_l$ for $l \geq 1$. $\qquad\blacksquare$

Remark 6.22 *On the bound* (6.18). The bound (6.18) depends formally on l. One can remove this dependence by using that $\theta^{l-1} < \theta$. However, in the form (6.18) it becomes clearer that the importance of the flops of the coarsest grid solver decreases with increasing level. $\qquad\square$

Example 6.23 *Computational costs for different cycles.* Consider a standard uniform refinement, i.e., it is $c_h = 2^{-d}$, where d is the dimension of the domain.

For one dimension, the theory applies for the V-cycle because $\gamma c_h = 1/2$, but not for the W-cycle since $\gamma c_h = 1$.

In two dimensions, one has for the V-cycle $\gamma c_h = 1/4$ and for the W-cycle $\gamma c_h = 1/2$. Then, one obtains from (6.18) the following estimates for the computational costs:

- V-cycle

$$c_l < \frac{4}{3}\left(\nu c_s + c_r + c_p\right) + \left(\frac{1}{4}\right)^{l-1}\frac{c_0}{N_1},$$

- W-cycle

$$c_l < 2\left(\nu c_s + c_r + c_p\right) + \left(\frac{1}{2}\right)^{l-1}\frac{c_0}{N_1}.$$

Neglecting the flops for the coarsest grid solver, a W-cycle for a two-dimensional problem requires roughly 1.5 times the number of flops of a V-cycle.

In three dimensions, one finds for the V-cycle that $\gamma c_h = 1/8$ and for the W-cycle that $\gamma c_h = 1/4$. Then, the number of flops per cycle is bounded by

- V-cycle

$$c_l < \frac{8}{7}\left(\nu c_s + c_r + c_p\right) + \left(\frac{1}{8}\right)^{l-1}\frac{c_0}{N_1},$$

- W-cycle

$$c_l < \frac{4}{3}\left(\nu c_s + c_r + c_p\right) + \left(\frac{1}{4}\right)^{l-1}\frac{c_0}{N_1}.$$

Hence, the W-cycle is only 1.167 times as expensive as the V-cycle.

These results to think about using different strategies for different dimensions. The V-cycle is always more efficient whereas the W-cycle is generally more stable. Since the efficiency gain of the V-cycle in three dimensions is only small, one should apply there the W-cycle. In two dimensions, one should first try if the V-cycle works. As alternative, one can use in both cases the F-cycle. The computation of the numerical costs of the F-cycle is an *exercise*.   □

Corollary 6.24 Number of flops for $\theta = 1$. *Let the notations be as in Theorem 6.21 and let $\theta = 1$. Then, the number of operations on level l is bounded by*

$$\left((\nu c_s + c_r + c_p)\, l + \frac{c_0}{N_1}\right) N_l.$$

Proof: The proof starts like the proof of Theorem 6.21 until (6.19). Then, one sets $\theta = 1$ in (6.19) to obtain the statement of the corollary.   ■

Example 6.25 *W-cycle in one dimension.* The corollary states that the number of flops for the W-cycle in one dimension is not proportional to N_l. Hence, the W-cycle is not optimal in one dimension.   □

Remark 6.26 *Memory requirements of the multigrid method.* The sparse matrix on level l requires the storage of $c_m N_l$ numbers, where c_m is independent of l. In addition, one has to store the arrays $\mathbf{v}_l$ and $\mathbf{f}_l$, which are $2N_l$ numbers. It follows that the total storage requirements are

$$
\begin{aligned}
(2 + c_m)\sum_{k=0}^{l} N_k &= (2 + c_m)\left(N_l + N_l\frac{N_{l-1}}{N_l} + N_l\frac{N_{l-1}}{N_l}\frac{N_{l-2}}{N_{l-1}} + \dots\right) \\
&\leq (2 + c_m)\, N_l \sum_{k=0}^{l-1} c_h^k \\
&\leq \frac{(2 + c_m)\, N_l}{1 - c_h},
\end{aligned}
$$

if $c_h < 1$. A method that works only on the finest grid requires at least the storage of $(2 + N_l)$ numbers. Thus, for uniform standard refinement, i.e., $c_h = 2^{-d}$, one has for

- $d = 1$: that the multigrid method needs 100 %,
- $d = 2$: that the multigrid method needs 33.3 %,
- $d = 3$: that the multigrid method needs 14.3 %,

more memory than a single grid algorithm on the finest grid. $\qquad\qquad\square$