Multigrid Methods

A brief tutorial for the course
Analysis of Numerical Methods

by

Bernhard Müller
Division of Scientific Computing,
Department of Information Technology,
Uppsala University, Sweden

November 2004

Abstract

This brief tutorial on multigrid methods is a part of the course 'Analysis of Numerical Methods' for fourth year students in the scientific computing specialization of technical physics at Uppsala University. The brief tutorial is based on the compendium 'Analys av numeriska metoder' (in Swedish) by my colleagues Bertil Gustafsson, Björn Sjögreen and Michael Thuné, September 1990. Chapter 12 of the compendium on 'Multigridmetoder' by Michael Thuné is updated here. The course notes on the subject by my colleague Jarmo Rantakokko were of great use as well.

This brief tutorial on multigrid methods gives an introduction to multigrid methods and their convergence analysis. It is based on the books by Wolfgang Hackbusch, "Multi-Grid Methods and Applications", Springer-Verlag, 1985, and by William L. Briggs, Van Emden Henson and Steve F. McCormick, "A Multigrid Tutorial", Second Edition, SIAM, 2000.

Uppsala in November 2004

Bernhard Müller

Contents

Chapter 1

Introduction

Multigrid methods have been extremely successful in many branches of scientific computing. We focus here on the efficient solution of linear systems arising from discretizations of elliptic PDEs.

The multigrid algorithm was discovered and analyzed by R.P. Fedorenko in the 1960ies. In the 1970ies, Achi Brandt recognized the efficiency and potential of the multigrid approach and applied it to engineering problems. Wolfgang Hackbusch pioneered the analysis of multigrid methods in the 1970ies.

Multigrid methods have been the subject of many conferences, journal articles and books, cf. e.g. `http://www.mgnet.org` and are used in many codes in science and engineering.

Chapter 2

Model Problem

We consider the 1D model problem [Hackbusch(1985)]

$$\begin{cases} -u''(x) = f(x) & 0 < x < 1 \\ u(0) = u(1) = 0 \end{cases} \tag{2.1}$$

For the discretization of (2.1), we choose an equidistant distribution of the grid points

$$x_j = jh\,, \qquad j = 0, 1, ..., n, n+1,$$

where $h = \frac{1}{n+1}$, cf. Fig. 2.1. Note that the details of the grid definition are important! The ODE boundary value problem (2.1) is discretized by the standard second order finite difference approximation

$$-D_+D_-u_j = f_j\,, \qquad j = 1, ..., n, \tag{2.2}$$

i.e. $-\frac{1}{h^2}(u_{j+1} - 2u_j + u_{j-1}) = f_j\,, j = 1, ..., n.$ The discretization yields the linear $n \times n$ system of equations

$$\mathbf{Au} = \mathbf{f}\,, \tag{2.3}$$

where

$$\mathbf{A} = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix}, \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{pmatrix}, \mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{pmatrix},$$

with u_j approximating the exact solution of (2.1) $u(x_j)$ at interior grid point x_j and $f_j = f(x_j)$, $j = 1, ..., n$.

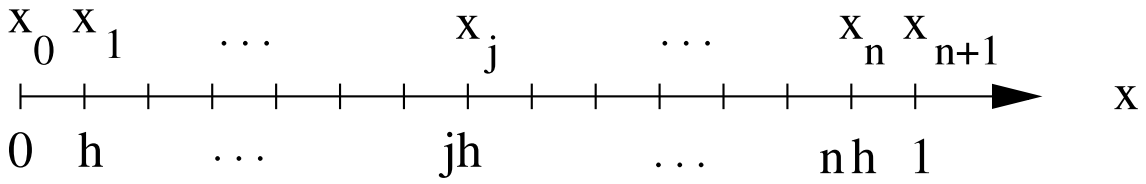

Figure 2.1: *Grid for 1D model problem.*

We want to solve the linear system (2.3) by an iterative method and analyse its convergence. We choose the **damped Jacobi method**:

$$\mathbf{u}^{(m+1)} = \mathbf{u}^{(m)} - \omega \mathbf{D}^{-1}(\mathbf{A}\mathbf{u}^{(m)} - \mathbf{f}), \tag{2.4}$$

where $\mathbf{D} = \frac{2}{h^2}\mathbf{I}$ denotes the diagonal of $\mathbf{A}$. ω is a damping factor, which is $\omega = 1$ for the original Jacobi method. Starting from an initial guess $\mathbf{u}^{(0)}$, we want the iterative method to converge as fast as possible, e.g. until $||\mathbf{u}^{(m+1)} - \mathbf{u}^{(m)}||_2 \leq tol$ with a prescribed tolerance tol. Convergence of an iterative method

$$\mathbf{u}^{(m+1)} = \mathbf{G}\mathbf{u}^{(m)} + \mathbf{d} \tag{2.5}$$

is determined by the iteration matrix $\mathbf{G}$: If the spectral radius of $\mathbf{G}$ is smaller than 1, i.e. $\rho(\mathbf{G}) < 1$, the iterative method (2.5) converges. The lower $\rho(\mathbf{G})$ is, the faster is the convergence.

Lemma 1:

The eigenvalues of the iteration matrix $\mathbf{G}_\omega = \mathbf{I} - \omega \mathbf{D}^{-1}\mathbf{A}$ are

$$\lambda_\mu = 1 - 2\omega sin^2\left(\frac{\mu\pi h}{2}\right), \mu = 1, ..., n \tag{2.6}$$

and the corresponding eigenvectors of $\mathbf{G}_\omega$ are

$$\mathbf{v}_\mu = \sqrt{2h}\begin{pmatrix} sin(\mu\pi h) \\ sin(\mu\pi 2h) \\ \vdots \\ sin(\mu\pi(n-1)h) \\ sin(\mu\pi nh) \end{pmatrix}, \mu = 1, ..., n \tag{2.7}$$

Proof:

Prove first that $\alpha_\mu = \frac{4}{h^2} sin^2\left(\frac{\mu\pi h}{2}\right)$ is an eigenvalue of matrix $\mathbf{A}$ in (2.3) and $\mathbf{v}_\mu$ the corresponding eigenvector of $\mathbf{A}$, where $\mu = 1, ..., n$. Exercise!
Because of $\mathbf{D}^{-1} = \frac{h^2}{2}\mathbf{I}$, we can conclude from

$$\mathbf{A}\mathbf{v}_\mu = \alpha_\mu \mathbf{v}_\mu$$

that

$$\mathbf{G}_\omega \mathbf{v}_\mu = (\mathbf{I} - \omega\frac{h^2}{2}\mathbf{A})\mathbf{v}_\mu = (1 - \omega\frac{h^2}{2}\alpha_\mu)\mathbf{v}_\mu .$$

Thus, $\lambda_\mu = 1 - \omega\frac{h^2}{2}\frac{4}{h^2} sin^2\left(\frac{\mu\pi h}{2}\right) = 1 - 2\omega sin^2\left(\frac{\mu\pi h}{2}\right)$ is an eigenvector of $\mathbf{G}_\omega$ and $\mathbf{v}_\mu$ the corresponding eigenvector.

Remarks:

1. Lemma 1 indicates that the damping factor must satisfy $0 < \omega \leq 1$ to get convergence, i.e. $\rho(\mathbf{G}_\omega) = \max_{\mu=1,...,n} |\lambda_\mu| < 1$. We use the fact that the eigenvalues λ_μ lie between $1 - 2\omega sin^2\left(\frac{\pi}{2(n+1)}\right)$ and $1 - 2\omega sin^2\left(\frac{\pi n}{2(n+1)}\right)$, because $\frac{1}{n+1} \leq \mu h \leq \frac{n}{n+1}$. Consequently,

$$\rho(\mathbf{G}_\omega) = 1 - \mathcal{O}(h^2) \tag{2.8}$$

2. Suppose we want to reduce the error of the kth iteration $\mathbf{e}^{(m)} = \mathbf{u}^{(m)} - \mathbf{u}^*$, where $\mathbf{u}^*$ is the exact solution of (2.3), by a factor of 10^{-d} of the initial error $\mathbf{e}^{(0)}$. The error is measured in the 2-norm. We have

$$\mathbf{e}^{(m)} = \mathbf{u}^{(m)} - \mathbf{u}^* = \mathbf{G}_\omega(\mathbf{u}^{(m-1)} - \mathbf{u}^*) = \mathbf{G}_\omega \mathbf{e}^{(m-1)} = ...\mathbf{G}_\omega^m \mathbf{e}^{(0)} .$$

Since $||\mathbf{e}^{(m)}||_2 \leq ||\mathbf{G}_\omega||_2^m ||\mathbf{e}^{(0)}||_2$, we require $||\mathbf{G}_\omega||_2^m \leq 10^{-d}$. Thus, a sufficient condition for the number of iterations of the damped Jacobi method to reduce the initial error by a factor of 10^{-d} is:

$$m \geq \frac{d}{-log_{10}(||\mathbf{G}_\omega||_2)} . \tag{2.9}$$

We have $\rho(\mathbf{G}_\omega) \leq ||\mathbf{G}_\omega||_2$ and approximate $||\mathbf{G}_\omega||_2$ by $\rho(\mathbf{G}_\omega)$. Then, we can conclude from (2.8) that it will take $\mathcal{O}(h^{-2}) = \mathcal{O}(n^2)$ iterations. Since one iteration takes $\mathcal{O}(n)$ operations, $\mathcal{O}(n^3)$ operations are needed for convergence. Thus, the damped Jacobi method and similar iteration methods cannot compete with the direct solution of the tridiagonal linear system (2.3), which takes $\mathcal{O}(n)$ operations.

Multigrid Concept

Multigrid (MG) methods are based on the concept of combining two strategies:

1. **Error smoothing on fine grid.**
   We require the high wave numbers, i.e. $k = \mu h$ with $\frac{1}{2} \leq \mu h < 1$, to be damped. In the model problem, we therefore choose the damping factor ω of the damped Jacobi method (2.4) such that the eigenvalues λ of $\mathbf{G}_\omega$, cf. (2.6), satisfy $|\lambda| \leq \frac{1}{2}$. In Fig. 2.2, we see that $\omega = \frac{1}{2}$ or $\omega = \frac{2}{3}$ are good choices. After a few (say s) iterations, the error modes with high wave numbers are so much reduced that the residual

   $$\mathbf{r} = \mathbf{f} - \mathbf{A}\mathbf{u}^{(s)} \tag{2.10}$$

   is smooth enough to be transfered to the coarse grid to be corrected there. The residual on the fine grid can be transfered to the coarse grid by restricting it to the values of the coarse grid. We assume that the coarse grid size h_0 is twice the fine one h, i.e. $h_0 = 2h$, cf. Fig.s 3.1 and 3.2.

2. **Error correction on coarse grid.**
   The error correction on the coarse grid is based on iterative improvement. Let us denote the restricted residual on the coarse grid by $\mathbf{r}_0$. Then, we can correct the approximate fine grid solution $\mathbf{u}^{(s)}$ by solving the linear system

   $$\mathbf{A}_0 \mathbf{v}_0 = \mathbf{r}_0 , \tag{2.11}$$

   where $\mathbf{A}_0$ is the matrix discretizing the negative second derivative of the ODE (2.1) on the coarse grid, i.e. $\mathbf{A}_0$ is defined by (2.3) with $n_0 = \frac{1}{h_0} - 1$ instead of $n = \frac{1}{h} - 1$.
   Let us first assume that the smaller system (2.11) is solved by a direct method. In fact, that is done for the two-grid multigrid method, where we only have one fine and one coarse grid. The coarse grid solution, i.e. here the exact solution $\mathbf{v}_0$ of (2.11), is then transfered or prolongated to the fine grid, e.g. by interpolation. The prolongated solution denoted by $\mathbf{v}$ is then used to correct the approximate fine grid solution $\mathbf{u}^{(s)}$ by

   $$\mathbf{u} = \mathbf{u}^{(s)} + \mathbf{v} . \tag{2.12}$$

If $\mathbf{v}$ were the exact solution of $\mathbf{Av} = \mathbf{r}$, then $\mathbf{u}$ would be the exact solution to the system (2.3), because $\mathbf{Au} = \mathbf{Au}^{(s)} + \mathbf{Av} = \mathbf{Au}^{(s)} + \mathbf{r} = \mathbf{Au}^{(s)} + (\mathbf{f} - \mathbf{Au}^{(s)}) = \mathbf{f}$. But even if $\mathbf{v}$ is only an approximate correction obtained by prolongation from the coarse grid correction, we can get fast convergence, because the low wave number error components have been eliminated on the coarse grid.

In a multigrid method, we do not solve the coarse grid system (2.11) directly, but iteratively using the same algorithm as to solve (2.3) on the fine grid. The high wave numbers $k = \mu h_0$ with $\frac{1}{2} \leq \mu h_0 < 1$, on the coarse grid with $h_0 = 2h$ are the low wave numbers $k = \mu h$ with $\frac{1}{4} \leq \mu h < \frac{1}{2}$, on the fine grid. Since high wave number error components are effectively damped by a few damped Jacobi iterations, we apply those iterations on the coarse grid. Then, the coarse grid residual is transfered to the next coarser grid, where the high wave numbers corresponding to the low wave numbers $k = \mu h$ with $\frac{1}{8} \leq \mu h < \frac{1}{4}$, on the fine grid are effectively damped by a few damped Jacobi iterations. The process is repeated, until we reach the coarsest grid where the system, e.g. just one scalar equation, is solved directly.

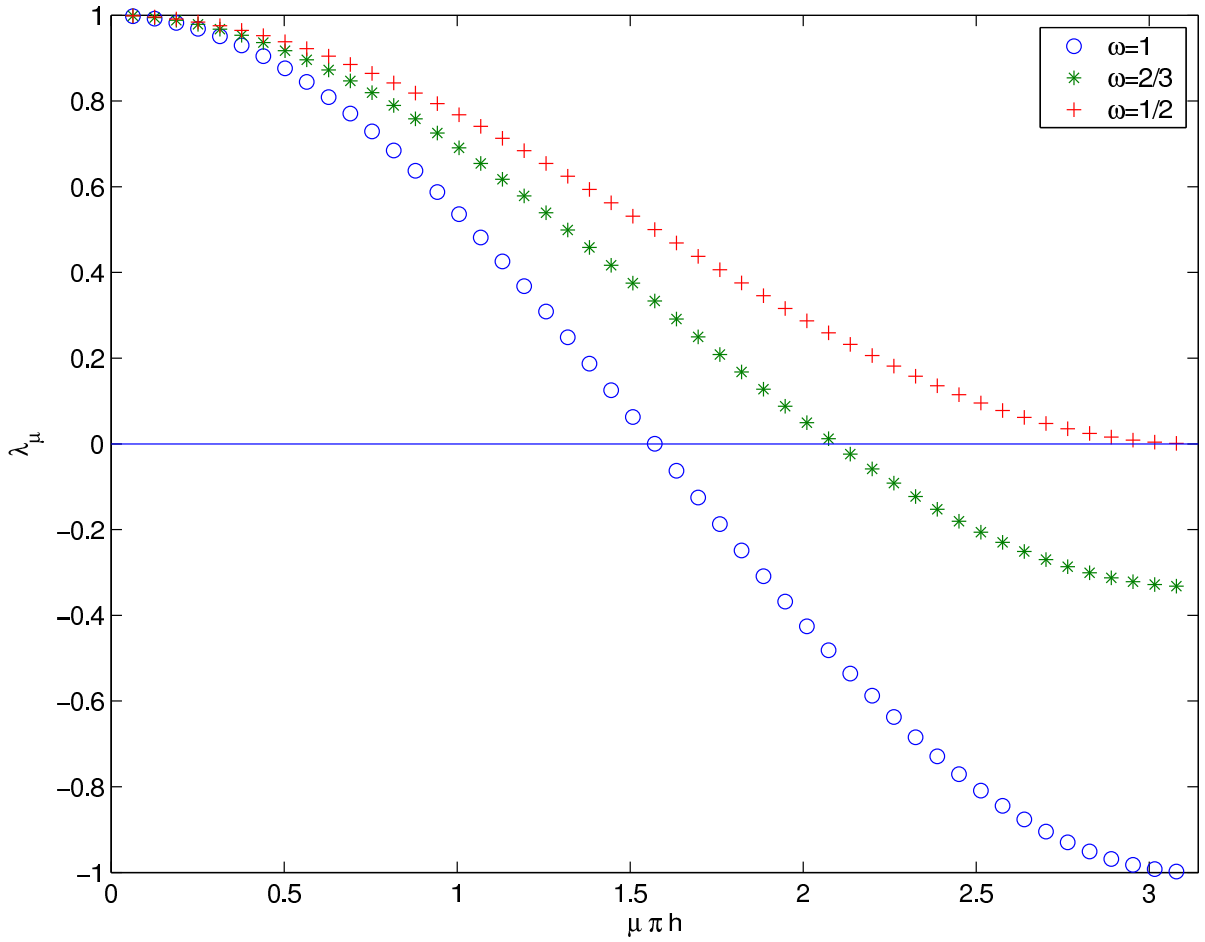

Figure 2.2: *Eigenvalues λ of $\mathbf{G}_\omega$ for damped Jacobi method, cf. (2.6).*

Chapter 3

Algorithm

As we discussed in the previous chapter, the multigrid algorithm is based on two strategies:

1. error smoothing on fine grid and
2. error correction on coarse grid.

In the simplest case, these strategies are combined for one fine grid and one coarse grid. The resulting algorithm of the two-grid multigrid (TGM) method is outlined in the following section and analysed in chapter 'Convergence'. If the correction equation on the coarse grid is not solved directly but iteratively by the multigrid strategy, we get the basic multigrid (MG) method. Then, we work with a number of grids: from the finest to the coarsest. The MG algorithm is presented in the second section.

3.1 Two-Grid Multigrid Method

Suppose $\mathbf{u}$ is the initial guess or the previously obtained approximation of the solution of (2.3) on the fine grid and $\mathbf{f}$ the corresponding right hand side. We perform the following steps:

1. run ν iterations with the damped Jacobi method (2.4) to smooth the residual on the fine grid $l = 1$,
2. restrict the residual $\mathbf{f} - \mathbf{A}_1\mathbf{u}$ to the coarse grid $l = 0$,
3. solve the linear system $\mathbf{A}_0\mathbf{v} = \mathbf{r}$ on the coarse grid $l = 0$ exactly by LU decomposition,
4. prolongate the correction $\mathbf{v}$ to the fine grid $l = 1$ and compute the corrected solution.

The two-grid multigrid (TGM) algorithm can be summarized by the function template

$$\begin{aligned} &\text{function} \quad \mathbf{u} = \mathrm{TGM}(\mathbf{u}, \mathbf{f}) \\ &\\ &\mathbf{u} := \mathbf{S}_1^{\nu}(\mathbf{u}, \mathbf{f}) \\ &\mathbf{r}_0 := \mathbf{R}(\mathbf{f} - \mathbf{A}_1\mathbf{u}) \\ &\mathbf{v}_0 := \mathbf{A}_0^{-1}\mathbf{r}_0 \\ &\mathbf{u} := \mathbf{u} + \mathbf{P}\mathbf{v}_0 \end{aligned} \tag{3.1}$$

$\mathbf{a} := \mathbf{b}$ denotes that array $\mathbf{a}$ is overwritten by array $\mathbf{b}$. The subscripts refer to the grid levels. Remember that $h = \frac{1}{n+1}$, cf. Fig. 2.1. The step sizes h are related by $h_1 = 2^{-1}h_0$ and the number of interior grid points n by $n_1 + 1 = 2(n_0 + 1)$.

The damped Jacobi method discussed in the previous chapter can be expressed as:

$$\mathbf{u}^{(m+1)} = \mathbf{S}_1\mathbf{u}^{(m)} + \mathbf{d}\,, \tag{3.2}$$

where $\mathbf{S}_1 = \mathbf{G}_\omega = \mathbf{I}_1 - \omega\mathbf{D}_1^{-1}\mathbf{A}_1$ and $\mathbf{d} = \omega\mathbf{D}_1^{-1}\mathbf{f}$. The index 1 refers again to the fine grid. $\mathbf{S}_1^\nu$ in (3.1) means that we apply the damped Jacobi method ν times.

The residual $\mathbf{r} = \mathbf{f} - \mathbf{A}_1\mathbf{u}$ is transfered or restricted to the coarse grid by averaging, cf. Fig. 3.1:

$$r_i^{(0)} = \frac{1}{4}\left(r_{2i-1}^{(1)} + 2r_{2i}^{(1)} + r_{2i+1}^{(1)}\right)\,, \quad i = 1, ..., n_0\,. \tag{3.3}$$

The superscripts of the components refer to the grid level. The restriction defined by (3.3) can also be expressed by

$$\mathbf{r}_0 = \mathbf{R}\mathbf{r}_1\,, \tag{3.4}$$

where

$$\mathbf{R} = \frac{1}{4}\begin{pmatrix} 1 & 2 & 1 & & & & & & \\ & & 1 & 2 & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & 2 & 1 \end{pmatrix}$$

is a $n_0 \times n_1$ matrix.

The coarse grid correction $\mathbf{v}_0 = \mathbf{A}_0^{-1}\mathbf{r}_0$ is transfered or prolongated to the fine grid by injection and averaging at joint and intermediate grid points, respectively:

$$\begin{array}{lcll} v_1^{(1)} & = & \frac{1}{2}(v_0^{(0)} + v_1^{(0)}) = \frac{1}{2}v_1^{(0)} & , \\ v_{2i}^{(1)} & = & v_i^{(0)} & , i = 1, ..., n_0\,, \\ v_{2i+1}^{(1)} & = & \frac{1}{2}(v_i^{(0)} + v_{i+1}^{(0)}) & , i = 1, ..., n_0 - 1\,, \\ v_{n_1}^{(1)} & = & \frac{1}{2}(v_{n_0}^0 + v_{n_0+1}^0) = \frac{1}{2}v_{n_0}^0 & , \end{array} \tag{3.5}$$

where the boundary conditions $v_0^{(0)} = v_{n_0+1}^0 = 0$ are used. The prolongation defined by (3.5) and illustrated in Fig.3.2 can also be expressed by

$$\mathbf{v}_1 = \mathbf{P}\mathbf{v}_0\,, \tag{3.6}$$

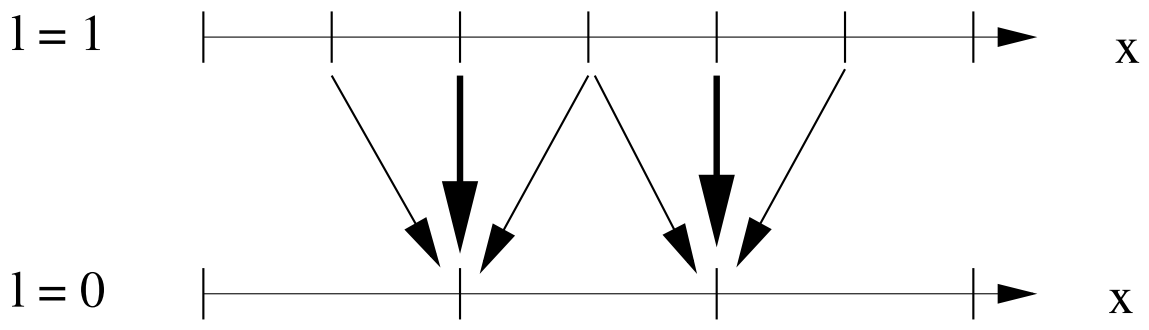

Figure 3.1: *Restriction from fine grid ($l = 1$) to coarse grid ($l = 0$).*

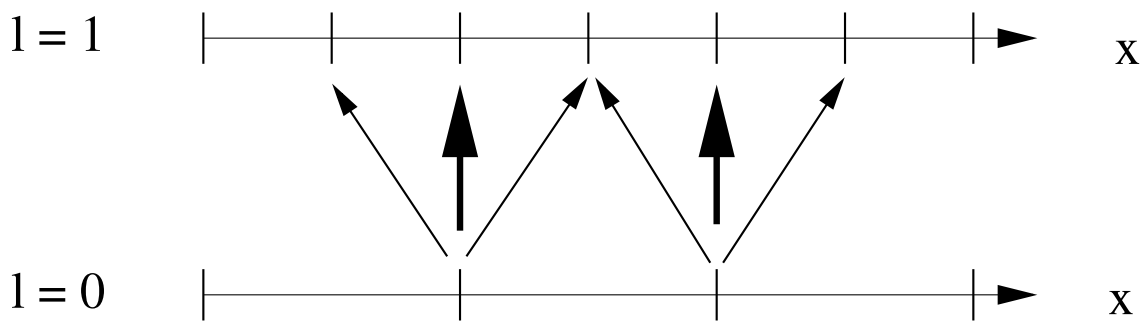

Figure 3.2: *Prolongation from coarse grid ($l = 0$) to fine grid ($l = 1$).*

where

$$\mathbf{P} = \frac{1}{2}\begin{pmatrix} 1 & & & & & & \\ 2 & & & & & & \\ 1 & 1 & & & & & \\ & 2 & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & & 1 & & \\ & & & & 2 & & \\ & & & & 1 & 1 & \\ & & & & & 2 & \\ & & & & & 1 & \end{pmatrix}$$

is a $n_1 \times n_0$ matrix.

Note:

$$\mathbf{P} = 2\,\mathbf{R}^T. \tag{3.7}$$

3.2 Multigrid Method

If the two-grid multigrid method (3.1) is used recursively starting from the finest grid level l with grid size $h_l = 2^{-l}h_0$ down to the coarsest grid level 0 with grid size h_0, we get the basic multigrid (MG) method. One iteration of it is called a V-cycle, cf. Fig. 3.3. Suppose $\mathbf{u}$ is the initial guess or the previously obtained approximation of the solution of (2.3) on the finest grid and $\mathbf{f}$ the corresponding right hand side. We perform the following steps:

1. if the coarsest grid level 0 is reached, solve the linear system $\mathbf{A}_0\mathbf{u} = \mathbf{f}$ exactly by LU decomposition (note that $\mathbf{A}_0\mathbf{v}_0 = \mathbf{r}_0$ is solved exactly, if $\mathbf{v}_0 := \mathrm{MG}(0, \mathbf{v}_0, \mathbf{r}_0)$ is called),

2. otherwise run ν iterations with the damped Jacobi method (2.4) to smooth the residual on the fine grid level l,

3. restrict the residual $\mathbf{f} - \mathbf{A}_l\mathbf{u}$ to the coarse grid $l-1$ and call the restricted residual $\mathbf{r}_{l-1}$,

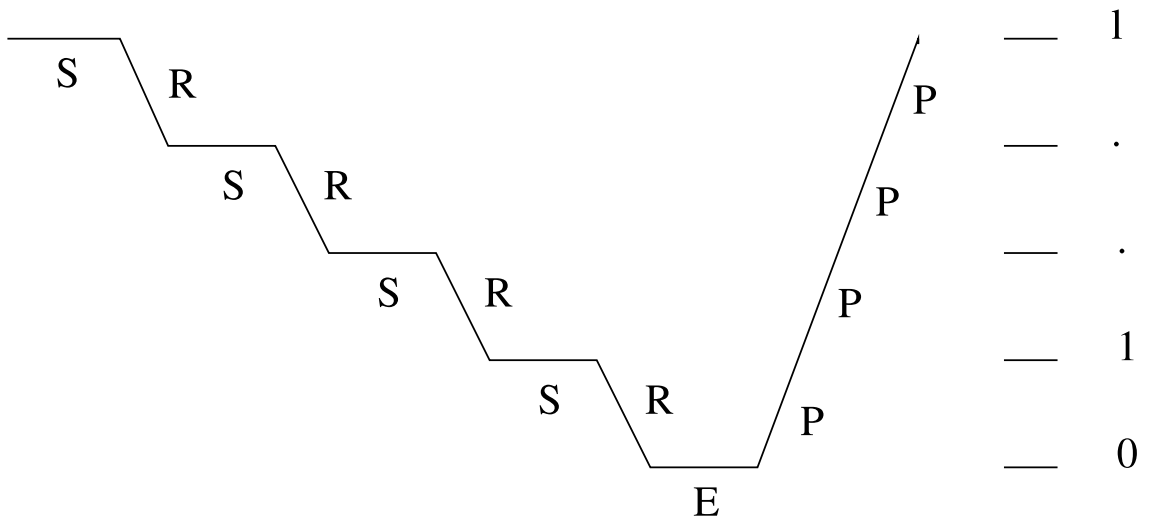

Figure 3.3: *One multigrid iteration called a V-cycle with smoothing (S), restriction (R), exact solution on coarsest grid level* 0 *(E), prolongation and correction (P).* $l = 4$ *in the example.*

4. solve the linear system $\mathbf{A}_{l-1}\mathbf{v}_{l-1} = \mathbf{r}_{l-1}$ on the coarse grid $l-1$ iteratively by using the multigrid algorithm with the starting guess $\mathbf{v}_{l-1} = 0$,

5. prolongate the correction $\mathbf{v}_{l-1}$ to the fine grid and compute the corrected solution.

The multigrid (MG) algorithm can be summarized by the function template

$$
\begin{array}{lll}
\text{function} & \mathbf{u} = \text{MG} & (l, \mathbf{u}, \mathbf{f}) \\
 & & \\
 & \text{if} & l = 0 \text{ then} \\
 & & \mathbf{u} := \mathbf{A}_0^{-1}\mathbf{f} \\
 & \text{else} & \\
 & & \mathbf{u} := \mathbf{S}_l^{\nu}(\mathbf{u}, \mathbf{f}) \\
 & & \mathbf{r}_{l-1} := \mathbf{R}(\mathbf{f} - \mathbf{A}_l\mathbf{u}) \\
 & & \mathbf{v}_{l-1} := 0 \\
 & & \mathbf{v}_{l-1} := \text{MG}(l-1, \mathbf{v}_{l-1}, \mathbf{r}_{l-1}) \\
 & & \mathbf{u} := \mathbf{u} + \mathbf{P}\mathbf{v}_{l-1} \\
 & \text{end if} &
\end{array}
\tag{3.8}
$$

Extensions of the MG algorithm, e.g. W-cycle, are discussed in the literature, cf. [Hackbusch(1985)] [Briggs et al. (2000)].

Chapter 4

Convergence

The convergence analysis of the two grid multigrid method (3.1) follows [Hackbusch(1985)]. The TGM method (3.1) can be written as an iteration method

$$\mathbf{u}^{(m+1)} = \mathbf{G}(\nu)\mathbf{u}^{(m)} + \mathbf{d}\,, \tag{4.1}$$

where $\mathbf{G}(\nu) = (\mathbf{I} - \mathbf{P}\mathbf{A}_0^{-1}\mathbf{R}\mathbf{A}_1)\mathbf{S}_1^{\nu}$. From the convergence theory of iterative methods, we shall use

$$\text{TGM} \quad \text{converges} \Longleftrightarrow \rho(\mathbf{G}(\nu)) < 1\,. \tag{4.2}$$

To estimate the error or the number of iterations for convergence we need to determine $||\mathbf{G}(\nu)||_2$, cf. (2.9).

The spectral radius and 2-norm of the TGM iteration matrix $\mathbf{G}(\nu)$ in (4.1) are determined

- by block-diagonalizing $\mathbf{G}(\nu)$ by means of the eigenvectors of $\mathbf{A}_1$ and $\mathbf{A}_0$ and
- by analysing the blocks.

The $n_1 \times n_1$ matrix $\mathbf{A}_1$ with step size h_1 is given by (2.3). The coarse grid $n_0 \times n_0$ matrix $\mathbf{A}_0$ with step size h_0 looks similar. Remember that the matix dimensions and step sizes are related by $n_1 = 2n_0 + 1$ and $h_1 = \frac{1}{2}h_0$. The eigenvalues and eigenvectors of $\mathbf{A}_i$, $i = 0$ or 1, cf. (2.6) and (2.7), are here denoted by (2.6)

$$\lambda_{\mu}^{(i)} = \frac{4}{h_i^2} sin^2\left(\frac{\mu\pi h_i}{2}\right), \quad \mu = 1, ..., n_i, \tag{4.3}$$

and

$$\mathbf{w}_{\mu}^{(i)} = \sqrt{2h_i}\begin{pmatrix} sin(\mu\pi h_i) \\ sin(\mu\pi 2h_i) \\ \vdots \\ sin(\mu\pi(n_i-1)h_i) \\ sin(\mu\pi n_i h_i) \end{pmatrix}, \quad \mu = 1, ..., n_i. \tag{4.4}$$

Grouping the eigenvectors of $\mathbf{A}_1$ according to the pattern:

first, last, second, second last, ..., middle,

we define the matrix

$$\mathbf{Q}_1 = [\mathbf{w}_1^{(1)}\, \mathbf{w}_{n_1}^{(1)}\, \mathbf{w}_2^{(1)}\, \mathbf{w}_{n_1-1}^{(1)}\, ...\, \mathbf{w}_{n_0}^{(1)}\, \mathbf{w}_{n_0+2}^{(1)}\, \mathbf{w}_{n_0+1}^{(1)}]\,. \tag{4.5}$$

The high wave number modes $\mathbf{w}_{n_0+1}^{(1)}, \mathbf{w}_{n_0+2}^{(1)}, ..., \mathbf{w}_{n_1}^{(1)}$ on the fine grid cannot be represented on the coarse grid, while the low wave number modes $\mathbf{w}_1^{(1)}, \mathbf{w}_2^{(1)}, ..., \mathbf{w}_{n_0}^{(1)}$ can. We define

$$\mathbf{Q}_0 = [\mathbf{w}_1^{(0)}\, \mathbf{w}_2^{(0)}\, ...\, \mathbf{w}_{n_0}^{(0)}]\,. \tag{4.6}$$

Since the eigenvectors (4.4) are orthonormal, i.e. $\mathbf{w}_j^{(i)T}\mathbf{w}_l^{(i)} = \delta_{jl}$ (Exercise!), the matrices $\mathbf{Q}_1$ and $\mathbf{Q}_0$ are orthogonal, i.e. $\mathbf{Q}_i^{-1} = \mathbf{Q}_i^T$.

We formulate three lemmas before stating theorem 1.

Lemma 1:

a) A similarity transformation with $\mathbf{Q}_0$ block-diagonalizes $\mathbf{A}_0$, i.e.

$$\mathbf{Q}_0^{-1}\mathbf{A}_0\mathbf{Q}_0 = \begin{pmatrix} \lambda_1^{(0)} & & & & 0 \\ & \lambda_2^{(0)} & & & \\ & & \ddots & & \\ & & & \lambda_{n_0-1}^{(0)} & \\ 0 & & & & \lambda_{n_0}^{(0)} \end{pmatrix} = \hat{\mathbf{A}}_0\,, \tag{4.7}$$

where $\lambda_\mu^{(0)} = \frac{4}{h_0^2} sin^2\left(\frac{\mu\pi h_0}{2}\right) = \frac{4}{h_1^2} s_\mu^2 c_\mu^2\,, \mu = 1, ..., n$ are the eigenvalues of $\mathbf{A}_0$.
$s_\mu^2 = sin^2\left(\frac{\mu\pi h_1}{2}\right)$ and $c_\mu^2 = cos^2\left(\frac{\mu\pi h_1}{2}\right)$.

b) A similarity transformation with $\mathbf{Q}_1$ block-diagonalizes $\mathbf{A}_1$ and $\mathbf{S}_1$, namely

$$\mathbf{Q}_1^{-1}\mathbf{A}_1\mathbf{Q}_1 = \begin{pmatrix} \mathbf{A}_1^{(1)} & & & & 0 \\ & \mathbf{A}_1^{(2)} & & & \\ & & \ddots & & \\ & & & \mathbf{A}_1^{(n_0)} & \\ 0 & & & & \mathbf{A}_1^{(n_0+1)} \end{pmatrix} = \hat{\mathbf{A}}_1\,, \tag{4.8}$$

where the diagonal 2×2 blocks $\mathbf{A}_1^{(\mu)}$ contain the two eigenvalues $\lambda_\mu^{(1)}$ and $\lambda_{\mu'}^{(1)}$ of $\mathbf{A}_1$, i.e.

$$\mathbf{A}_1^{(\mu)} = \frac{4}{h_1^2}\begin{pmatrix} s_\mu^2 & 0 \\ 0 & c_\mu^2 \end{pmatrix}, \quad \mu = 1, ..., n_0\,. \tag{4.9}$$

The lower right diagonal block is the scalar

$$\mathbf{A}_1^{(n_0+1)} = \lambda_{n_0+1}^{(1)} = \frac{2}{h_1^2}\,.$$

The block-diagonalization of $\mathbf{S}_1$ reads:

$$\mathbf{Q}_1^{-1}\mathbf{S}_1\mathbf{Q}_1 = \begin{pmatrix} \mathbf{S}_1^{(1)} & & & & 0 \\ & \mathbf{S}_1^{(2)} & & & \\ & & \ddots & & \\ & & & \mathbf{S}_1^{(n_0)} & \\ 0 & & & & \mathbf{S}_1^{(n_0+1)} \end{pmatrix} = \hat{\mathbf{S}}_1\,, \tag{4.10}$$

where the diagonal 2×2 blocks $\mathbf{S}_1^{(\mu)}$ contain the two eigenvalues $1 - \frac{h_1^2}{4}\lambda_\mu^{(1)}$ and $1 - \frac{h_1^2}{4}\lambda_{\mu'}^{(1)}$ of $\mathbf{S}_1 = \mathbf{I} - \frac{h_1^2}{4}\mathbf{A}_1$, i.e.

$$\mathbf{S}_1^{(\mu)} = \begin{pmatrix} c_\mu^2 & 0 \\ 0 & s_\mu^2 \end{pmatrix}, \quad \mu = 1, ..., n_0 \,. \tag{4.11}$$

The lower right diagonal block is the scalar

$$\mathbf{S}_1^{(n_0+1)} = 1 - \frac{h_1^2}{4}\lambda_{n_0+1}^{(1)} = \frac{1}{2} \,.$$

c) Similarity transformations with $\mathbf{Q}_0$ and $\mathbf{Q}_1$ block-diagonalize $\mathbf{R}$ and $\mathbf{P}$:

$$\mathbf{Q}_0^{-1}\mathbf{R}\mathbf{Q}_1 = \begin{pmatrix} \mathbf{R}_1^{(1)} & & & & 0 & 0 \\ & \mathbf{R}_1^{(2)} & & & & 0 \\ & & \ddots & & & \vdots \\ & & & \mathbf{R}_1^{(n_0-1)} & & 0 \\ 0 & & & & \mathbf{R}_1^{(n_0)} & 0 \end{pmatrix} = \hat{\mathbf{R}}, \tag{4.12}$$

$$\mathbf{Q}_1^{-1}\mathbf{P}_1\mathbf{Q}_0 = \begin{pmatrix} \mathbf{P}_1^{(1)} & & & & 0 \\ & \mathbf{P}_1^{(2)} & & & \\ & & \ddots & & \\ & & & \mathbf{P}_1^{(n_0-1)} & \\ 0 & & & & \mathbf{P}_1^{(n_0)} \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} = \hat{\mathbf{P}}, \tag{4.13}$$

where

$$\mathbf{R}^{(\mu)} = \frac{1}{\sqrt{2}}\left(c_\mu^2, -s_\mu^2\right), \qquad \mathbf{P}^{(\mu)} = \sqrt{2}\begin{pmatrix} c_\mu^2 \\ -s_\mu^2 \end{pmatrix}, \quad \mu = 1, ..., n_0 \,. \tag{4.14}$$

Proof:

$\mathbf{A}_i\mathbf{Q}_i = \mathbf{Q}_i\hat{\mathbf{A}}_i$, because $\mathbf{A}_i\mathbf{w}_\mu^{(i)} = \mathbf{w}_\mu^{(i)}\lambda_\mu^{(i)}$. Using trigonometric formulae, the relations between h_1 and h_0 and between n_1 and n_0 as well as the definitions, e.g. $\mathbf{S}_1 = \mathbf{I}\frac{h_1^2}{4}\mathbf{A}_1$, the lemma can be proved.
Exercise!

Lemma 2:

The iteration matrix of the TGM method (4.1) $\mathbf{G} = (\mathbf{I} - \mathbf{P}\mathbf{A}_0^{-1}\mathbf{R}\mathbf{A}_1)\mathbf{S}_1^\nu$ can be block-diagonalized by a similarity transformation with $\mathbf{Q}_1$:

$$\mathbf{Q}_1^{-1}\mathbf{G}\mathbf{Q}_1 = \begin{pmatrix} \mathbf{G}^{(1)} & & & & 0 \\ & \mathbf{G}^{(2)} & & & \\ & & \ddots & & \\ & & & \mathbf{G}^{(n_0)} & \\ 0 & & & & \mathbf{G}^{(n_0+1)} \end{pmatrix} = \hat{\mathbf{G}}, \tag{4.15}$$

where

$$\mathbf{G}^{(\mu)} = \begin{pmatrix} s_\mu^2 & c_\mu^2 \\ s_\mu^2 & c_\mu^2 \end{pmatrix}\begin{pmatrix} c_\mu^2 & 0 \\ 0 & s_\mu^2 \end{pmatrix}^\nu, \quad \mu = 1, ..., n_0, \tag{4.16}$$

and

$$\mathbf{G}^{(n_0+1)} = 2^{-\nu}$$

with $s_\mu^2 = sin^2\left(\frac{\mu\pi h_1}{2}\right)$ and $c_\mu^2 = cos^2\left(\frac{\mu\pi h_1}{2}\right)$.

Proof:
Using Lemma 1 and computing each block yields Lemma 2.
Exercise!

Lemma 3:

Let $\mathbf{G}^{(\mu)}$ denote the block matrices of the block-diagonalization $\hat{\mathbf{G}}$ of the the iteration matrix $\mathbf{G}$ of the TGM method (4.1), cf. Lemma 2.

The spectral radius and 2-norm of $\mathbf{G}$ are determined by

a)

$$\rho(\mathbf{G}) = max_{1\le\mu\le n_0+1}\{\rho(\mathbf{G}^{(\mu)})\}\,. \tag{4.17}$$

b)

$$||\mathbf{G}||_2 = max_{1\le\mu\le n_0+1}\{||\mathbf{G}^{(\mu)}||_2)\}\,. \tag{4.18}$$

Proof:
Since $\mathbf{G}$ and $\hat{\mathbf{G}}$ have the same eigenvalues, Lemma 3 follows from Lemma 2.
Exercise!

Theorem 1:

a) The TGM method (3.1) is **convergent**, because the spectral radius of the iteration matrix $\mathbf{G}(\nu)$ is bounded by

$$\rho(\mathbf{G}(\nu)) \le max_{0\le\xi\le 1/2}\{\xi(1-\xi)^\nu + (1-\xi)\xi^\nu\} =: \rho_\nu < 1\,. \tag{4.19}$$

b) The 2-norm of $\mathbf{G}(\nu)$ is bounded by

$$||\mathbf{G}(\nu)||_2 \le \sqrt{2}\, max_{0\le\xi\le 1/2}\{\sqrt{\xi^2(1-\xi)^{2\nu} + (1-\xi)^2\xi^{2\nu}}\} =: \zeta_\nu\,. \tag{4.20}$$

c) As the number of damped Jacobi iterations $\nu \longrightarrow \infty$, the asymptotic behaviour is

$$\rho_\nu \approx \frac{1}{e\nu} \quad \text{and} \quad \zeta_\nu \approx \frac{\sqrt{2}}{e\nu}\,, \tag{4.21}$$

where $e = exp(1)$.

Proof:
Exercise!

Bounds of spectral radius and 2-norm from Theorem 1 are listed in Table 1 [Hackbusch(1985)].

ν	1	2	3	4	5	10
ρ_ν	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	0.0832	0.0671	0.0350
ζ_ν	$\frac{1}{2}$	$\frac{1}{4}$	0.1501	0.1159	0.0947	0.0496

Remarks:

1. **Low and high wave number modes**

The eigenvectors $\mathbf{w}_\mu^{(1)}$, $\mu = 1, ..., n_0$, represent low wave number modes. Since due to $h_0 = 2h_1$

$$(\mathbf{w}_\mu^{(1)})_{2l} = \sqrt{2h_1} sin(2l\mu\pi h_1) = \frac{1}{\sqrt{2}}\sqrt{2h_0} sin(l\mu\pi h_0) = \frac{1}{\sqrt{2}}(\mathbf{w}_\mu^{(0)})_l\,,$$

i.e. $\mathbf{w}_\mu^{(1)}$ and $\frac{1}{\sqrt{2}}\mathbf{w}_\mu^{(0)}$ coincide at common points, the low wave number modes on the fine grid can be represented on the coarse grid.

The eigenvectors $\mathbf{w}_{\mu'}^{(1)}$, $\mu' = n_1 + 1 - \mu$, $\mu = 1, ..., n_0$, represent high wave number modes, which cannot be 'seen' on the coarse grid.

In $\mathbf{Q}_1$, a low and a high wave number mode are grouped together, cf. (4.5).

2. **Damping of low and high wave number modes**

The iteration matrix describes how error modes are damped. Lemma 2 shows how the low and high wave number error modes are damped in the TGM method. The block matrices

$$\mathbf{G}^{(\mu)} = \begin{pmatrix} s_\mu^2 & c_\mu^2 \\ s_\mu^2 & c_\mu^2 \end{pmatrix} \begin{pmatrix} c_\mu^2 & 0 \\ 0 & s_\mu^2 \end{pmatrix}^\nu \qquad \begin{matrix} \textbf{low mode} \\ \textbf{high mode} \end{matrix} \tag{4.22}$$

coarse grid correction **fine grid smoothing**

indicate that the error modes are damped by two processes: smoothing on fine grid by ν damped Jacobi iterations (right matrix in (4.22)) and coarse grid correction (left matrix in (4.22)).

Since $0 \leq s_\mu^2 \leq \frac{1}{2}$ and $\frac{1}{2} \leq c_\mu^2 \leq 1$, the high wave number modes are effectively diminished by a factor of $s_\mu^{2\nu}$ using ν damped Jacobi iterations on the fine grid, while the low wave number modes are hardly diminished by a factor of $c_\mu^{2\nu}$.

Coarse grid correction has the opposite damping effect: the low and high wave number modes are diminished by a factor of $s_\mu^{2\nu}$ and $c_\mu^{2\nu}$, respectively, on the coarse grid.

Thus, fine grid smoothing and coarse grid correction have a complementary effect on damping error modes.

3. **Convergence rate of TGM method independent of h**

The convergence rate of the TGM method is independent of the step size h, because the spectral radius $\rho(\mathbf{G})$ and the so-called contraction number $||\mathbf{G}||_2$ are independent of h. Thus, the amount of work to obtain convergence is proportional to the number of unknowns n, i.e. $\mathcal{O}(n)$. Thus, the TGM metod is much more efficient than the Jacobi and Gauss-Seidel iteration methods, which require $\mathcal{O}(n^3)$ operations, and the SOR method with optimal relaxation parameter, which requires $\mathcal{O}(n^2)$ operations.

Chapter 5

Computational Work

The computational work per multigrid iteration called one work unit is studied, cf. [Hackbusch(1985)]. The fine grid smoothing (S), residual restriction (R), prolongation and coarse grid correction (C) and the exact solution on the coarsest grid level (0), which is assumed to be fixed, are assumed to have the following operation counts:

statement	number of arithmetic operations	
$\mathbf{u}_l := \mathbf{S}_l(\mathbf{u}_l, \mathbf{f}_l)$	$w_S \leq c_S n_l$	for $l \geq 1$
$\mathbf{r}_{l-1} := \mathbf{R}(\mathbf{f}_l - \mathbf{A}_l \mathbf{u}_l)$	$w_R \leq c_R n_l$	for $l \geq 1$
$\mathbf{u}_l := \mathbf{u}_l + \mathbf{P}\mathbf{v}_{l-1}$	$w_C \leq c_C n_l$	for $l \geq 1$
$\mathbf{u}_0 := \mathbf{A}_0^{-1}\mathbf{f}_0$	$w_0 \leq c_0$	

where n_l is the number of unknowns at grid level l. Since $\mathbf{S}_l$, $\mathbf{R}$ and $\mathbf{P}$ are sparse matrices, the operation counts of these statements are proportional to n_l.

We assume that the ratio of the unknowns is

$$\frac{n_{j-1}}{n_j} \leq 2^{-d}, \quad j = 1, ..., l, \tag{5.1}$$

where d is the dimension of the problem, i.e. $d = 1$ for our 1D model problem. We have $\frac{n_{j-1}}{n_j} < \frac{n_{j-1}}{n_j - 1} = \frac{1}{2}$ in 1D.

Proposition: (4.3.1, p.87, [Hackbusch(1985)])
Suppose $\theta = \gamma 2^{-d} < 1$, where $\gamma = 1$ for the V-cycle (basic MG method (3.8)) and $\gamma = 2$ for the W-cycle [Hackbusch(1985)] [Briggs et al.(2000)]. Then, one iteration of the MG method, i.e. one work unit, requires $c_l n_l$ operations, where

$$c_l \leq \frac{\nu c_S + c_R + c_C}{1 - \theta} + \theta^l \left[\frac{c_0}{\gamma 2^{-dl} n_l} - \frac{\nu c_S + c_R + c_C}{1 - \theta} \right] \tag{5.2}$$

$$< \frac{\nu c_S + c_R + c_C}{1 - \theta} + \theta^{l-1} \frac{c_0}{n_1}. \tag{5.3}$$

Proof:
Let w_l denote the number of aritmetic operations per MG iteration at grid level l with n_l unknowns. Then,

$$w_j \leq (\nu c_S + c_R + c_C) n_j + \gamma w_{j-1}, \quad j = 2, ..., l, \tag{5.4}$$

$$w_1 \leq (\nu c_S + c_R + c_C) n_1 + w_0, \tag{5.5}$$

$$w_0 = c_0 \tag{5.6}$$

$$\begin{aligned} \Longrightarrow w_l &\leq (\nu c_S + c_R + c_C)(n_l + \gamma n_{l-1} + \gamma^2 n_{l-2} + ... + \gamma^{l-1} n_1) + \gamma^{l-1} c_0 \quad (5.7) \\ &\leq (\nu c_S + c_R + c_C) n_l (1 + \theta + \theta^2 + ... + \theta^{l-1}) + \theta^{l-1} 2^{d(l-1)} c_0 \quad (5.8) \end{aligned}$$

Using the formula for the partial sum of a geometric series and $\frac{2^{d(l-1)}}{n_l} \leq \frac{1}{n_1}$, the proposition can be proved.

Chapter 6

Exercises

1. Derive the vector $\mathbf{d}$ in the TGM method (4.1)

$$\mathbf{u}^{(m+1)} = \mathbf{G}(\nu)\mathbf{u}^{(m)} + \mathbf{d}\,. \tag{6.1}$$

2. Show that the eigenvectors of $\mathbf{A}$ and $\mathbf{G}_\omega$ (2.7) are orthogonal.

3. Prove the lemmas and theorem in chapter 4.

4. The efficiency

$$E = -\frac{\text{operations per MG iteration}}{log(\text{ contraction number})} = -\frac{w_l}{log(\zeta_\nu)} \tag{6.2}$$

   is the work required to reduce the error by a factor of e^{-1}. Assume that $w_l \approx (\nu+1)c_S n_l$. What is the optimal number ν of damped Jacobi smoothing iterations? *Hint*: You may assume the asymptotic expression for ζ_ν.

5. Estimate the operation count of the MG method in three dimensions.

6. a) Determine the number of arithmetic operations for the MG method to solve the 1D model problem, cf. chapter 2. Consider $\nu = 3$ smoothing iterations with the damped Jacobi method and assume restriction and prolongation as discussed above.

   b) Compare the operation count for the TGM method for the 1D model problem ($\nu = 1, 2, 3$) with the operation count for the damped Jacobi method for $h = 2^{-m}, m = 2, ..., 10$.