A new lemma in multigrid convergence theory

EINDHOVEN UNIVERSITY OF TECHNOLOGY
Department of Mathematics and Computing Science

RANA 91-07
May 1991

A NEW LEMMA IN MULTIGRID CONVERGENCE THEORY

by
A. Reusken

ISSN: 0926-4507
Reports on Applied and Numerical Analysis
Department of Mathematics and Computing Science
Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven
The Netherlands

A NEW LEMMA IN MULTIGRID CONVERGENCE THEORY

by

Arnold Reusken

A new lemma in multigrid convergence theory

1. Introduction

Consider a sequence of discretized boundary value problems:

$$(1.1) \qquad \mathbf{L}_l\,\mathbf{u}_l = \mathbf{f}_l \qquad l = 0,1,2,\ldots\,.$$

The index l corresponds to a measure h_l of the meshwidth used in the discretization, and $h_0 > h_1 > \ldots > h_{l-1} > h_l > \ldots$ with $\lim_{l\to\infty} h_l = 0$. Let $\mathbf{L}_l : X_l \to X_l$ with $X_l = \mathbb{R}^{n_l}$, and let there be given a *prolongation* $\mathbf{p} : X_{l-1} \to X_l$ and a *restriction* $\mathbf{r} : X_l \to X_{l-1}$.
Furthermore we have a *smoothing method* (e.g. damped Jacobi, Gauss-Seidel) denoted by $\mathcal{S}_l(\mathbf{u}_l, \mathbf{f}_l)$. The basic *two-grid method* can be represented as follows:

```
procedure  Φ_l^TGM(u_l, f_l) ;
   begin   for i := 1 to ν do u_l := S_l(u_l, f_l) ;
           d_{l-1} := r(L_l u_l − f_l) ;
           e_{l-1} := L_{l-1}^{-1} d_{l-1} ;                 (*)
           u_l := u_l − p e_{l-1} ;
           Φ^TGM := u_l
   end ;
```

If $\mathbf{S}_l$ denotes the iteration matrix of $\mathcal{S}_l$ then the iteration matrix of the two-grid method is given by $\mathbf{T}_l(\nu) = (\mathbf{I} - \mathbf{p}\,\mathbf{L}_{l-1}^{-1}\,\mathbf{r}\,\mathbf{L}_l)\,\mathbf{S}_l^{\nu} = (\mathbf{L}_l^{-1} - \mathbf{p}\,\mathbf{L}_{l-1}^{-1}\,\mathbf{r})\,\mathbf{L}_l\,\mathbf{S}_l^{\nu}$. In the *multigrid method* the coarse grid system in (*) is not solved exactly, but approximately by using one ("V-cycle") or two ("W-cycle") iterations of the two-grid method on level $l-1$; then systems on level $l-2$ occur which are again solved approximately by one or two iterations of the two-grid method on level $l-2$, etc. until $\mathbf{L}_0^{-1}\,\mathbf{d}_0$.
It is well-known that these multigrid methods can be very efficient for solving discretized boundary value problems.
There is an extensive literature concerning the convergence analysis of multigrid methods. We refer to [1], [3] and the references given there. We briefly discuss two important approaches.

In "symmetric multigrid" one assumes that the matrices $\mathbf{L}_l$ are symmetric (w.r.t. to the Euclidean inner product) and satisfy a Galerkin relation: $\mathbf{L}_{l-1} = \mathbf{r}\,\mathbf{L}_l\,\mathbf{p} = \mathbf{p}^T\,\mathbf{L}_l\,\mathbf{p}$. Furthermore the iteration matrix $\mathbf{S}_l$ of the smoothing method is assumed to be symmetric (or symmetrizable) with respect to the energy inner product. Also an "approximation property" (in which regularity of the boundary value problem is used) should hold. Then one can prove l-independent convergence for the multigrid V-cycle with only one smoothing iteration ($\nu = 1$). For h_l small enough results for a (slightly) broader class of problems can be obtained by using perturbation arguments.
Another approach, applicable to a larger class of problems, has been introduced by Hackbusch (cf. [1]). In this theory one first proves convergence of the two-grid method and then deals with the multigrid (W-cycle) method by means of a perturbation argument. The convergence of the two-grid method is based on the "Approximation Property" and "Smoothing Property". For a detailed explanation we refer to [1]. We summarize the main points. The

Smoothing Property holds if there is a function $\eta(\nu)$, independent of l and with $\lim_{\nu\to\infty} \eta(\nu) = 0$, such that

(1.2) $$\|\mathbf{L}_l\, \mathbf{S}_l^{\nu}\| \le \eta(\nu)\, \|\mathbf{L}_l\|\ ;$$

the Approximation Property holds if there is a constant c_A, independent of l, such that

(1.3) $$\|\mathbf{L}_l^{-1} - \mathbf{p}\, \mathbf{L}_{l-1}^{-1}\, \mathbf{r}\| \le c_A\, \|\mathbf{L}_l\|^{-1}\ .$$

If both properties hold then for the iteration matrix of the two-grid method we have $\|\mathbf{T}_l(\nu)\| \le c_A\, \eta(\nu)$, so l-independent convergence for ν large enough. The Approximation Property can be verified by using discretization error estimates. The verification of the Smoothing Property is based on the following fundamental lemma:

> If $\mathbf{A}$ is symmetric positive definite, with $\sigma(\mathbf{A}) \subseteq [0,1]$, then $\|\mathbf{A}(\mathbf{I}-\mathbf{A})^{\nu}\|_2 \le \eta_0(\nu)$ holds with $\eta_0(\nu) := \nu^{\nu}/(\nu+1)^{\nu+1}$ (note: $\eta_0(\nu) = (e\nu)^{-1} + O(\nu^{-2})$ for $\nu \to \infty$).

Using this lemma and suitable perturbation arguments the Smoothing Property can be proved for different relaxation methods and for a fairly large class of symmetric and nonsymmetric problems.
However, even with the latter, more general, approach it is not clear how to analyse convergence in certain (interesting) situations. For example serious problems occur if one wants to prove results in a norm different from the Euclidean or energy norm (e.g. the maximum norm), or if one tries to analyse strongly nonsymmetric problems (e.g. convection-diffusion with strong convection). One important source of problems is that in the fundamental lemma for the Smoothing Property an orthogonal eigenvector basis is essential.
In this paper we introduce a new approach for verifying the Smoothing Property. Roughly our main result is that if $\|\mathbf{I} - 2\,\mathbf{M}_l^{-1}\,\mathbf{L}_l\| \le 1$ holds for all l then the Smoothing Property holds for $\mathbf{S}_l := \mathbf{I} - \mathbf{M}_l^{-1}\,\mathbf{L}_l$ in the same norm $\|\cdot\|$. The norm $\|\cdot\|$ may be any submultiplicative matrix norm (e.g. the maximum norm). Note that there are no symmetry conditions involved.
In this paper we only discuss the Smoothing Property. Of course, for two grid convergence this should be combined with an analysis of the Approximation Property. This will be done in forthcoming papers (cf. Remark 5.1 below).

2. A new lemma

In Lemma 2.1 below our main result is given. The norm $\|\cdot\|$ we use may be any submultiplicative matrix norm.

Lemma 2.1. Let $\mathbf{A}$ be an $n \times n$-matrix with $\|\mathbf{A}\| \le 1$. Then the following holds:

(2.2) $$\|(\mathbf{I}-\mathbf{A})\,(\mathbf{I}+\mathbf{A})^{\nu}\| \le 2 \binom{\nu}{\left[\frac{1}{2}\nu\right]} \le 2^{\nu+1}\sqrt{\frac{2}{\pi\nu}} \qquad (\nu \ge 1)\ .$$

Proof.

$$(\mathbf{I}-\mathbf{A})\,(\mathbf{I}+\mathbf{A})^{\nu} = (\mathbf{I}-\mathbf{A}) \sum_{k=0}^{\nu} \binom{\nu}{k} \mathbf{A}^k$$

$$= \mathbf{I} - \mathbf{A}^{\nu+1} + \sum_{k=1}^{\nu} \left(\binom{\nu}{k} - \binom{\nu}{k-1} \right) \mathbf{A}^k .$$

So

$$\text{(2.3)} \qquad \|(\mathbf{I}-\mathbf{A})\,(\mathbf{I}+\mathbf{A})^{\nu}\| \leq 2 + \sum_{k=1}^{\nu} \left| \binom{\nu}{k} - \binom{\nu}{k-1} \right| .$$

Using $\binom{\nu}{k} \geq \binom{\nu}{k-1} \iff k \leq \frac{1}{2}(\nu+1)$, and $\binom{\nu}{k} = \binom{\nu}{\nu-k}$ we get

$$\sum_{k=1}^{\nu} \left| \binom{\nu}{k} - \binom{\nu}{k-1} \right| = \sum_{1}^{[\frac{1}{2}(\nu+1)]} \left(\binom{\nu}{k} - \binom{\nu}{k-1} \right) + \sum_{[\frac{1}{2}(\nu+1)]+1}^{\nu} \left(\binom{\nu}{k-1} - \binom{\nu}{k} \right)$$

$$= \sum_{1}^{[\frac{1}{2}\nu]} \left(\binom{\nu}{k} - \binom{\nu}{k-1} \right) + \sum_{m=1}^{[\frac{1}{2}\nu]} \left(\binom{\nu}{m} - \binom{\nu}{m-1} \right) \qquad (m = \nu + 1 - k)$$

$$= 2 \sum_{k=1}^{[\frac{1}{2}\nu]} \left(\binom{\nu}{k} - \binom{\nu}{k-1} \right) = 2 \left(\binom{\nu}{[\frac{1}{2}\nu]} - \binom{\nu}{0} \right) .$$

Combined with (2.3) this yields the first inequality in (2.2). Now define the sequences

$$a_k := \binom{2k}{k} \sqrt{k}\, 2^{-2k} \quad \text{and} \quad b_k := \int_0^{\frac{\pi}{2}} \sin^k(x)\, dx \qquad (k \geq 0) .$$

Elementary analysis yields that $(b_k)_{k\geq 0}$ is monotonically decreasing,

$$b_k = \frac{k-1}{k}\, b_{k-2} \,, \quad b_{2k+1} = (2^k\, k!)^2 \,/\, (2k+1)! \,, \quad b_{2k} = \tfrac{1}{2}\pi (2k)! \,/\, (2^k\, k!)^2 \,.$$

From this it follows that

$$1 \leq \frac{b_{2k}}{b_{2k+1}} \leq 1 + \frac{1}{2k}$$

and thus

$$\lim_{k\to\infty} \frac{b_{2k}}{b_{2k+1}} = 1 .$$

This then yields

$$(2.4) \qquad \lim_{k\to\infty} a_k = \lim_{k\to\infty} \frac{(2k)!\,\sqrt{k}}{(2^k\,k!)^2} = \frac{1}{\sqrt{\pi}} .$$

Furthermore $(a_k)_{k\geq 0}$ is monotonically increasing, so $a_k \leq \frac{1}{\sqrt{\pi}}$ for all k, and

$$(2.5) \qquad \binom{2k}{k} \leq 2^{2k}\sqrt{\frac{1}{\pi k}} \qquad \text{for all} \quad k .$$

If ν is even then applying (2.5) with $k = \frac{1}{2}\nu$ yields $\binom{\nu}{[\frac{1}{2}\nu]} \leq 2^\nu \sqrt{\frac{2}{\pi\nu}}$. If ν is uneven then we use (2.5) with $k = \frac{1}{2}(\nu+1)$ and we get

$$\binom{\nu}{[\frac{1}{2}\nu]} = \frac{1}{2}\binom{\nu+1}{\frac{1}{2}(\nu+1)} \leq \frac{1}{2}2^{\nu+1}\sqrt{\frac{2}{\pi(\nu+1)}} \leq 2^\nu\sqrt{\frac{2}{\pi\nu}} .$$

□

If the norm we want to use is the Euclidean norm $\|\cdot\|_2$, then the condition $\|\mathbf{A}\|_2 \leq 1$ in Lemma 2.1 can be weakened by using the *numerical radius* $r(\mathbf{A})$:

$$r(\mathbf{A}) := \max\{|\mathbf{x}^*\mathbf{A}\,\mathbf{x}| \;\mid\; \mathbf{x}\in \boldsymbol{C}^n ,\;\; \|\mathbf{x}\|_2 = 1\} .$$

Some well-known properties are given below:

$$(2.6) \qquad \rho(\mathbf{A}) \leq r(\mathbf{A}) \leq \|\mathbf{A}\|_2 \leq 2\,r(\mathbf{A})$$

$$(2.7) \qquad r(\alpha\,\mathbf{A}) = |\alpha|\;r(\mathbf{A}) \quad (\alpha\in\boldsymbol{C})$$

$$(2.8) \qquad r(\mathbf{A}+\mathbf{B}) \leq r(\mathbf{A}) + r(\mathbf{B})$$

$$(2.9) \qquad r(\mathbf{A}^k) \leq r(\mathbf{A})^k .$$

Using the numerical radius results in the following variant of Lemma 2.1:

Lemma 2.10. Let $\mathbf{A}$ be an $n\times n$-matrix with $r(\mathbf{A}) \leq 1$.
Then the following holds:

$$\|(\mathbf{I}-\mathbf{A})\,(\mathbf{I}+\mathbf{A})^\nu\|_2 \leq 4\binom{\nu}{[\frac{1}{2}\nu]} \leq 2^{\nu+2}\sqrt{\frac{2}{\pi\nu}} \quad (\nu\geq 1) .$$

Proof. The proof of Lemma 2.1 remains valid with (2.3) replaced by

$$\|(\mathbf{I}-\mathbf{A})\,(\mathbf{I}+\mathbf{A})^{\nu}\|_2 \le 2\,r(\mathbf{I}-\mathbf{A}^{\nu+1}+\sum_{k=1}^{\nu}\left(\binom{\nu}{k}-\binom{\nu}{k-1}\right)\mathbf{A}^k)$$

$$\le 2\,\{r(\mathbf{I})+r(\mathbf{A})^{\nu+1}+\sum_{k=1}^{\nu}\left|\binom{\nu}{k}-\binom{\nu}{k-1}\right| r(\mathbf{A})^k\}$$

$$\le 2\,(2+\sum_{k=1}^{\nu}\left|\binom{\nu}{k}-\binom{\nu}{k-1}\right|)\,.$$

□

Remark 2.11. The result of Lemma 2.1 (or 2.10) holds without any symmetry conditions on the matrix $\mathbf{A}$. We compare this result with a result for the symmetric situation $\mathbf{A}=\mathbf{A}^*$. For the norm $\|\cdot\|$ we take the Euclidean norm $\|\cdot\|_2$; note that due to the symmetry $\|\mathbf{A}\|_2 = r(\mathbf{A})$. Using an orthogonal eigenvector basis it is easy to prove the following result:

If $\|\mathbf{A}\|_2 \le 1$ then with $\eta_0(\nu)=\nu^{\nu}/(\nu+1)^{\nu+1}$ the following holds:

(2.12) $\quad \|(\mathbf{I}-\mathbf{A})\,(\mathbf{I}+\mathbf{A})^{\nu}\|_2 \le 2^{\nu+1}\,\eta_0(\nu) \le 2^{\nu+1}\,/\,(e\nu+1) \qquad (\nu \ge 1)\,.$

Moreover, the second inequality in (2.12) is sharp for ν "large" ($(e\nu+1)\,\eta_0(\nu)\to 1$ for $\nu\to\infty$) and the first inequality is sharp for $\mathbf{A}=\lambda_\nu \mathbf{I}$ with $\lambda_\nu = 1-\frac{2}{\nu+1}$.
Comparing the result in (2.12) with the result of Lemma 2.1 (or 2.10) we see that if the symmetry condition is released the upperbound $C\,2^{\nu}/\nu$ (in (2.12)) is replaced by $\tilde{C}\,2^{\nu}/\sqrt{\nu}$.

Remark 2.13. The result of Lemma 2.1 is sharp in the following sense. For ν "large" the second inequality in (2.2) is sharp:

$$\lim_{\nu\to\infty} 2\binom{\nu}{\left[\frac{1}{2}\nu\right]}\left(2^{\nu+1}\sqrt{\frac{2}{\pi\nu}}\right)^{-1} = 1\,.$$

With respect to the first inequality in (2.2) we note the following. For $\|\cdot\|$ we take the maximum norm $\|\cdot\|_\infty$ and we take a fixed ν. For $n \ge \nu+2$ we define the $n\times n$-matrix $\mathbf{A}$ by

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & & \theta \\ & \ddots & \ddots & \\ \theta & & \ddots & 1 \\ & & & 0 \end{pmatrix}.$$

Then $\|(\mathbf{I}-\mathbf{A})\,(\mathbf{I}+\mathbf{A})^{\nu}\|_\infty = 2\binom{\nu}{\left[\frac{1}{2}\nu\right]}$.

3. The Smoothing Property

Using Lemma 2.1 (or 2.10) we now prove a theorem about the Smoothing Property. We consider a sequence of matrices $\mathbf{L}_l$ as in §1, and $\mathbf{S}_l := \mathbf{I} - \alpha_l \mathbf{M}_l^{-1} \mathbf{L}_l$ denotes the iteration matrix of a (damped) basic iterative method (e.g. Jacobi, Gauss-Seidel). We assume $\alpha_l \neq 0$.

Theorem 3.1. The following holds:
If

$$(3.1) \qquad \|\mathbf{I} - 2\alpha_l \mathbf{M}_l^{-1} \mathbf{L}_l\| \leq 1 \ ,$$

then

$$(3.2) \qquad \|\mathbf{L}_l \mathbf{S}_l^\nu\| \leq |\alpha_l|^{-1} \sqrt{\tfrac{2}{\pi\nu}} \, \|\mathbf{M}_l\| \qquad (\nu \geq 1) \ .$$

Proof. Define $\mathbf{A} := \mathbf{I} - 2\alpha_l \mathbf{M}_l^{-1} \mathbf{L}_l$. Then $\mathbf{I} - \mathbf{A} = 2\alpha_l \mathbf{M}_l^{-1} \mathbf{L}_l$ and $\mathbf{I} + \mathbf{A} = 2\mathbf{S}_l$. Using Lemma 2.1 we get

$$\begin{aligned}
\|\mathbf{L}_l \mathbf{S}_l^\nu\| &= \|\tfrac{1}{2} \alpha_l^{-1} \mathbf{M}_l (\mathbf{I} - \mathbf{A}) \, (\tfrac{1}{2}(\mathbf{I} + \mathbf{A}))^\nu\| \\
&\leq \tfrac{1}{2} |\alpha_l|^{-1} \, \|\mathbf{M}_l\| \, (\tfrac{1}{2})^\nu \, \|(\mathbf{I} - \mathbf{A}) \, (\mathbf{I} - \mathbf{A})^\nu\| \\
&\leq \tfrac{1}{2} |\alpha_l|^{-1} \, \|\mathbf{M}_l\| \, (\tfrac{1}{2})^\nu \, 2^{\nu+1} \sqrt{\frac{2}{\pi\nu}} \\
&= |\alpha_l|^{-1} \sqrt{\frac{2}{\pi\nu}} \, \|\mathbf{M}_l\| \ .
\end{aligned}$$

□

For the Smoothing Property we want results uniformly in l (cf. §1). Such a result is given in the following

Corollary 3.3. Assume that for all l the following conditions are fulfilled:

$$(3.4) \qquad \|\mathbf{I} - 2\alpha_l \mathbf{M}_l^{-1} \mathbf{L}_l\| \leq 1$$

$$(3.5) \qquad |\alpha_l|^{-1} \, \|\mathbf{M}_l\| \leq c \, \|\mathbf{L}_l\| \qquad (c \text{ independent of } l) \ .$$

Then

$$(3.6) \qquad \|\mathbf{L}_l \mathbf{S}_l^\nu\| \leq \frac{\tilde{c}}{\sqrt{\nu}} \, \|\mathbf{L}_l\| \ \text{ holds for all } \ l \ , \ \text{ with } \ \tilde{c} = c \sqrt{\frac{2}{\pi}} \ .$$

Remark 3.7. In view of Lemma 2.10 it is clear that if $\|\cdot\| = \|\cdot\|_2$ then the condition (3.1) (or (3.4)) can be replaced by the weaker condition $r(\mathbf{I} - 2\alpha_l \mathbf{M}_l^{-1} \mathbf{L}_l) \leq 1$. In the upperbound in (3.2) (or (3.6)) an additional factor 2 then occurs.

4. Applications

In this section we show three applications of Theorem 3.1. The results in the subsections 4.1 and 4.2 are new. The Smoothing Property in subsection 4.3 was already known (cf. [1]), however, our approach here is essentially different from the one used in [1]. We think that the examples in this section are a starting point for more new results concerning the Smoothing Property, based on Theorem 3.1.

4.1. Damped Jacobi and Gauss-Seidel for weakly diagonally dominant matrices

Assume that the matrices $\mathbf{L}_l$ are weakly diagonally dominant. Let $\mathbf{L}_l = \mathbf{M}_l - \mathbf{N}_l$ be the splitting corresponding to the Jacobi or Gauss-Seidel iteration. Then $\mathbf{S}_l := \mathbf{I} - \frac{1}{2}\mathbf{M}_l^{-1}\mathbf{L}_l$ is the iteration matrix of a damped (Jacobi or Gauss-Seidel) iteration. Note that $\|\mathbf{I} - \mathbf{M}_l^{-1}\mathbf{L}_l\|_\infty = \|\mathbf{M}_l^{-1}\mathbf{N}_l\|_\infty \le 1$ holds, so condition (3.1) is fulfilled with $\alpha_l = \frac{1}{2}$, and thus we get

$$(4.2) \qquad \|\mathbf{L}_l\,\mathbf{S}_l^\nu\|_\infty \le 2\sqrt{\frac{2}{\pi\nu}}\,\|\mathbf{M}_l\|_\infty \,.$$

Clearly if $2\|\mathbf{M}_l\|_\infty \le c\|\mathbf{L}_l\|_\infty$ holds with c independent of l, then we have the Smoothing Property (3.6) *in the maximum norm.* Note that no symmetry conditions are used.

4.2. Damped Gauss-Seidel for nonsymmetric $1-D$ problems

Let

$$\mathbf{L}_l := \beta\,h_l^{-2}\begin{pmatrix} 1 & -\delta_l & & & \\ -\gamma_l & 1 & \ddots & & \theta \\ & \ddots & \ddots & \ddots & \\ & \theta & \ddots & \ddots & -\delta_l \\ & & & -\gamma_l & 1 \end{pmatrix} \quad \text{with } \beta > 0\,,\ \ 0 \le \gamma_l < 1\,,\ \ 0 \le \delta_l \le 1\,,\ \ \gamma_l + \delta_l \le 1\,.$$

We use a Gauss-Seidel splitting $\mathbf{L}_l = \mathbf{M}_l - \mathbf{N}_l$; so

$$\mathbf{N}_l = \beta\,h_l^{-2}\,\delta_l\,\mathbf{J} \quad \text{with } \mathbf{J} = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \theta \\ \theta & & \ddots & 1 \\ & & & 0 \end{pmatrix} \qquad (\text{note} : \|\mathbf{J}\|_2 = 1)$$

and $\mathbf{M}_l = \beta\,h_l^{-2}(\mathbf{I} - \gamma_l\,\mathbf{J}^T)$.

Then $\|\mathbf{M}_l^{-1}\|_2 \le \beta^{-1}\,h_l^2(1-\gamma_l)^{-1}$ holds, and

$$\|\mathbf{M}_l^{-1}\,\mathbf{N}_l\|_2 \le \|\mathbf{M}_l^{-1}\|_2\,\|\mathbf{N}_l\|_2 \le \beta^{-1}\,h_l^2(1-\gamma_l)^{-1}\,\beta\,h_l^{-2}\,\delta_l \le 1\,.$$

So (3.4) is fulfilled with $\alpha_l = \frac{1}{2}$.
Note that for $2 \le j \le n-1$:

$$\|\mathbf{L}_l\|_2 \ge \sqrt{(\mathbf{L}_l\,\mathbf{e}_j)^T\,(\mathbf{L}_l\,\mathbf{e}_j)} = \beta\,h_l^{-2}\,\sqrt{\delta_l^2+1+\gamma_l^2} \ge \beta\,h_l^{-2}\,\delta_l = \|\mathbf{N}_l\|_2$$

and thus $\|\mathbf{M}_l\|_2 = \|\mathbf{L}_l+\mathbf{N}_l\|_2 \le 2\,\|\mathbf{L}_l\|_2$. So (3.5) is fulfilled with $c = 4$.
We conclude that the Smoothing Property (3.6) holds:

$$(4.3) \qquad \|\mathbf{L}_l(\mathbf{I}-\tfrac{1}{2}\,\mathbf{M}_l^{-1}\,\mathbf{L}_l)^\nu\|_2 \le 4\,\sqrt{\frac{2}{\pi\nu}}\,\|\mathbf{L}_l\|_2\,.$$

Note that the upperbound in (4.3) does not depend on γ_l, δ_l.
In (4.3) we have the Smoothing Property for the damped Gauss-Seidel iteration in the Euclidean norm (the Smoothing Property in $\|\cdot\|_\infty$ follows from subsection 4.1). Similar arguments can be used to prove the Smoothing Property for the damped Jacobi iteration in the Euclidean norm.

4.3. Damped Richardson for nonsymmetric $\mathbf{L}_l$

We use the following result about the convergence of a damped Richardson iteration (cf. [4], [2]):

If for a matrix $\mathbf{A}$ there are constants $0 < \lambda \le \Lambda$, $\tau \ge 0$ such that

$$(4.4) \qquad \lambda\mathbf{I} \le \tfrac{1}{2}(\mathbf{A}+\mathbf{A}^T) \le \Lambda\mathbf{I}$$

$$(4.5) \qquad \|\tfrac{1}{2}(\mathbf{A}-\mathbf{A}^T)\|_2 \le \tau\,,$$

then $\|\mathbf{I}-\theta\,\mathbf{A}\|_2 \le 1$ if $0 \le \theta \le 2\lambda(\lambda\Lambda+\tau^2)^{-1}$.

Now for our matrices $\mathbf{L}_l$ we assume that there are constants $c_1 > 0$ and c_2 independent of l such that the following holds:

$$(4.6) \qquad c_1\,\mathbf{I} \le \tfrac{1}{2}(\mathbf{L}_l+\mathbf{L}_l^T)$$

$$(4.7) \qquad \|\tfrac{1}{2}(\mathbf{L}_l-\mathbf{L}_l^T)\|_2 \le c_2\,\|\mathbf{L}_l\|_2^{\frac{1}{2}}\,.$$

The condition (4.7) corresponds to the property that the nonsymmetric part of the operator is a "lower order term" (cf. also [1]).

Applying the above-mentioned result for damped Richardson yields (with $\lambda = c_1$, $\Lambda = \|\mathbf{L}_l\|_2$, $\tau = c_2\,\|\mathbf{L}_l\|_2^{\frac{1}{2}}$):

$$(4.8) \qquad \|\mathbf{I} - 2\alpha_l \mathbf{L}_l\|_2 \le 1 \quad \text{if} \quad \alpha_l = \gamma\, c_1 (c_1 \|\mathbf{L}_l\|_2 + c_2^2\, \|\mathbf{L}_l\|_2)^{-1}\ , \quad \text{with} \quad \gamma \in]0,1] \text{ fixed.}$$

So condition (3.1) is satisfied and (3.2) yields the Smoothing Property for $\mathbf{S}_l := \mathbf{I} - \alpha_l\, \mathbf{L}_l$:

$$(4.9) \qquad \|\mathbf{L}_l\, \mathbf{S}_l^\nu\|_2 \le |\alpha_l|^{-1} \sqrt{\frac{2}{\pi\nu}} = \gamma^{-1}(1 + c_2^2 c_1^{-1}) \sqrt{\frac{2}{\pi\nu}}\ \|\mathbf{L}_l\|_2\ .$$

Our approach yields the Smoothing Property in the Euclidean norm for the nonsymmetric case as in [1]. However, in our approach here we avoid perturbation arguments as used in [1].

5. Remark 5.1.

Example 4.1 shows that the Smoothing Property can be proved in the maximum norm. It remains to investigate the Approximation Property in this norm. The Approximation Property is closely related to a discretization error estimate. If we consider an elliptic second order boundary value problem and use linear finite elements then for the $1 - D$ case an "optimal" result holds: the discretization error in the maximum norm is $O(h_l^2)$ $(l \to \infty)$; for $2 - D$ problems a "nearly optimal" result holds: the discretization error in the maximum norm is $O(h_l^2 \log|h_l|)$ $(l \to \infty)$. These results for the discretization error probably induce similar results for the Approximation Property. This will be analyzed in a forthcoming paper. Another interesting question is whether this new approach can be used to prove a suitable (i.e. related to the Approximation Property) Smoothing Property for (stable) discretizations of singularly perturbed convection-diffusion problems. This is a subject of current research.

References

[1] Hackbusch, W.: Multi-grid methods and applications. Springer, Berlin (1985).

[2] Hackbusch, W.: Iterative Lösung großer schwachbesetzter Gleichungssysteme. Teubner, Stuttgart (1991).

[3] McCormick, S. (ed.): Multigrid methods. SIAM, Philadelphia (1987).

[4] Samarskii, A.A. and Nikolaev, E.S.: Numerical methods for grid equations. Vol. II: Iterative methods. Birkhäuser, Basel (1989).