

## A note on multigrid methods for nonlinear problems

***Citation for published version (APA):***

Reusken, A. A. (1995). *A note on multigrid methods for nonlinear problems*. (RANA : reports on applied and numerical analysis; Vol. 9507). Technische Universiteit Eindhoven.

***Document status and date:***

Published: 01/01/1995

***Document Version:***

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

***Please check the document version of this publication:***

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

***General rights***

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

***Take down policy***

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

EINDHOVEN UNIVERSITY OF TECHNOLOGY  
Department of Mathematics and Computing Science

RANA 95-07  
May 1995

A Note on Multigrid Methods  
for Nonlinear Problems

by

A.A. Reusken



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  
ISSN: 0926-4507

# A Note on Multigrid Methods for Nonlinear Problems

by

Arnold Reusken  
Eindhoven University of technology  
Department of Mathematics and Computing Science  
P.O. Box 513  
5600 MB Eindhoven, The Netherlands

## Introduction

Well-known methods for solving discretized nonlinear partial differential equations using multigrid techniques are:

- Nonlinear Multi-Grid Method (NMGM; Hackbusch [3]).
- The Full Approximation Scheme (FAS; Brandt [2]).
- Newton iteration combined with a linear multigrid method.

In practice all three methods are used. A systematic comparison, theoretically or numerically, between these methods is still lacking. In [5] the behaviour of the three methods applied to a test problem is shown. In the present paper we give an elementary theoretical comparison. Furthermore, we derive the NMGM from a non-standard point of view.

In §1 we describe the Nonlinear Two-Grid Method (NTGM), the two-grid FAS, and the relation between these two. In §2 we introduce a model class of locally affine problems, to which we restrict ourselves in the remainder of the paper. In §3 we recall the principle of Nested Iteration, which is useful for generating acceptable starting vectors. In §4 it is shown how for our model class of nonlinear problems the NTGM can be derived in a natural way, such that the role of the parameters occurring in the method is clear. In §5 the Newton iteration combined with a linear multigrid method is given. Finally in §6 methods are compared.

## 1 The Nonlinear Two-Grid Method and the Full Approximation Scheme

We assume a standard setting for discretized boundary value problems. Then we have a sequence  $(U_k)_{k \geq 0}$  of spaces of grid functions corresponding to a decreasing sequence  $(h_k)_{k \geq 0}$  of mesh sizes. On every  $U_k$  we have a discrete operator  $\mathbf{F}_k : U_k \rightarrow U_k$ . We recall the NTGM (see [3]) for solving  $\mathbf{F}_k(\mathbf{u}_k^*) = \mathbf{f}_k$ . In the algorithm below  $\mathcal{S}_k(\mathbf{u}, \mathbf{f})$  is a suitable smoother,  $\mathbf{p} : U_{k-1} \rightarrow U_k$  is a prolongation,  $\mathbf{r} : U_k \rightarrow U_{k-1}$  is a restriction and  $\tilde{\mathbf{u}}_{k-1} \in U_{k-1}$ ,  $s \in \mathbb{R} \setminus \{0\}$  are parameters that will be discussed further on. The number of pre- and post-smoothing iterations is denoted by  $\nu_1, \nu_2$  respectively.

```

Procedure NTGM( $k, \mathbf{u}, \mathbf{f}$ ); integer  $k$ ; array  $\mathbf{u}, \mathbf{f}$ ;
  Begin array  $\mathbf{d}, \mathbf{v}$ ;
     $\mathbf{u} := \mathcal{S}_k^{\nu_1}(\mathbf{u}, \mathbf{f})$ ;
     $\mathbf{d} := \mathbf{r}(\mathbf{f} - \mathbf{F}_k(\mathbf{u}))$ ;
     $\mathbf{d} := \mathbf{F}_{k-1}(\tilde{\mathbf{u}}_{k-1}) + s\mathbf{d}$ ;
     $\mathbf{v} := \mathbf{F}_{k-1}^{-1}(\mathbf{d})$ ;
     $\mathbf{u} := \mathbf{u} + \frac{1}{s}\mathbf{p}(\mathbf{v} - \tilde{\mathbf{u}}_{k-1})$ ;
     $\mathbf{u} := \mathcal{S}_k^{\nu_2}(\mathbf{u}, \mathbf{f})$ 
  End;

```

We obtain the FAS algorithm if in the NTGM we take the following choice of the parameters:  $\tilde{\mathbf{u}}_{k-1} = \mathbf{r}\mathbf{u}$ ,  $s = 1$  (more general:  $\tilde{\mathbf{u}}_{k-1} = \tilde{\mathbf{r}}\mathbf{u}$  with a restriction  $\tilde{\mathbf{r}}$  which may differ from  $\mathbf{r}$ ).

**Remark 1.1.**

- a) With respect to the derivation of the coarse grid correction we note that the approach of Hackbusch differs from the approach of Brandt: for NTGM a linearization is used, whereas for FAS an equation for the correction on the fine grid is "transferred" to the coarse grid.
- b) In the presentation of the NTGM in [3] the role of  $\tilde{\mathbf{u}}_{k-1}$  and of  $s$  is not very clear.

## 2 A sequence of locally affine problems

Let  $(U_k)_{k \geq 0}$  be a sequence of spaces of grid functions corresponding to a decreasing sequence of mesh sizes  $(h_k)_{k \geq 0}$ . On each  $U_k$  we have a (e.g. scaled Euclidean) norm  $\|\cdot\| = \|\cdot\|_k$ .

Consider problems which are discretizations of a given continuous problem  $F(u^*) = f$ :

$$\mathbf{F}_k(\mathbf{u}_k^*) = \mathbf{f}_k \quad (k = 0, 1, 2, \dots)$$

with  $\mathbf{F}_k : U_k \supseteq D_k \rightarrow U_k$ . We assume  $\mathbf{F}_k \in C^1(D_k)$ .

In general, for nonlinear problems, computations should be restricted to some neighbourhood of  $\mathbf{u}_k^*$ . We define  $B_k = B(\mathbf{u}_k^*; \varepsilon_k) := \{\mathbf{u} \in U_k \mid \|\mathbf{u} - \mathbf{u}_k^*\| \leq \varepsilon_k\}$ , and consider:

$$\mathbf{F}_k : B_k \rightarrow U_k \quad (k = 0, 1, 2, \dots) \text{ for given } (\varepsilon_k)_{k \geq 0} .$$

The  $\varepsilon_k$  are determined by e.g.:

- domain of  $\mathbf{F}_k$
- uniqueness of  $\mathbf{u}_k^*$
- nonlinearity of  $\mathbf{F}_k$  (one should avoid "too strong nonlinearities")
- limitations on computational range.

All three situations  $\varepsilon_k/\varepsilon_{k+1} \approx 1$ ,  $\varepsilon_k/\varepsilon_{k+1} \ll 1$ ,  $\varepsilon_k/\varepsilon_{k+1} \gg 1$  might occur. We are interested in iterative methods for solving

$$\mathbf{F}_k(\mathbf{u}_k^*) = \mathbf{f}_k, \text{ with } \mathbf{F}_k : B_k \rightarrow U_k, \quad (2.1)$$

for given  $k \in \mathbb{N}$ .

Related to the problem in (2.1) we now introduce problems that are *locally affine* at  $\mathbf{u}_k^*$ . We assume given linear operators  $\mathbf{C}_k : U_k \rightarrow U_k$ ,  $k \in \mathbb{N}$ , (for example  $\mathbf{C}_k = D\mathbf{F}_k(\mathbf{u}_k^*)$ ) and define:

$$\mathbf{A}_k : B_k \rightarrow U_k, \quad \mathbf{A}_k(\mathbf{u}) = \mathbf{f}_k + \mathbf{C}_k(\mathbf{u} - \mathbf{u}_k^*). \quad (2.2)$$

We consider the problem of solving

$$\mathbf{A}_k(\mathbf{u}_k^*) = \mathbf{f}_k \text{ with } \mathbf{A}_k \text{ as in (2.2)}, \quad (2.3)$$

for given  $k \in \mathbb{N}$ . In the remainder we study multigrid methods for (2.1) by applying them to the problem in (2.3). This simplification is based on the following motivation:

- Reasonable methods for the problem in (2.1) should at the least give reasonable results for the problem in (2.3) with  $\mathbf{C}_k \approx D\mathbf{F}(\mathbf{u}_k^*)$ .
- If  $\varepsilon_k$  is "small enough" (can be quantified, cf. convergence domain of Newton) the results for a method applied to the problem in (2.1) and applied to the problem in (2.3) with  $\mathbf{C}_k \approx D\mathbf{F}(\mathbf{u}_k^*)$  are comparable.

The following simple relations hold for all  $\mathbf{u} \in B_k$ :

$$D\mathbf{A}_k(\mathbf{u}) = \mathbf{C}_k \quad (2.4a)$$

$$\mathbf{u}_k^* = \mathbf{u} - (D\mathbf{A}_k(\mathbf{u}))^{-1}(\mathbf{A}_k(\mathbf{u}) - \mathbf{f}_k) \text{ (one Newton iteration results in } \mathbf{u}_k^*) \quad (2.4b)$$

$$D\mathbf{A}_k(\mathbf{u})(\mathbf{v}) = \mathbf{A}_k(\mathbf{u} + \mathbf{v}) - \mathbf{A}_k(\mathbf{u}) \text{ whenever } \mathbf{u} + \mathbf{v} \in B_k \quad (2.4c)$$

### 3 Nested Iteration

When solving nonlinear problems, the importance of good starting vectors is clear. In our situation here we want starting vectors in  $B(\mathbf{u}_k^*; \varepsilon_k)$ . The well-known principle, of Nested Iteration is very useful for achieving this. A Nested Iteration procedure for  $\mathbf{A}_m \mathbf{u}_m^* = \mathbf{f}_m$  is as follows:

```

 $\mathbf{u}_0 := \mathbf{A}_0^{-1} \mathbf{f}_0;$ 
For  $\ell = 1$  step 1 until  $m$  do
  Begin
     $\mathbf{u}_\ell := \mathbf{p} \mathbf{u}_{\ell-1};$ 
     $\text{Meth}^{(i)}(\ell, \mathbf{u}_\ell, \mathbf{f}_\ell)$  (*  $i$ : iteration count *)
  End;

```

This Nested Iteration approach is illustrated in Fig. 1.

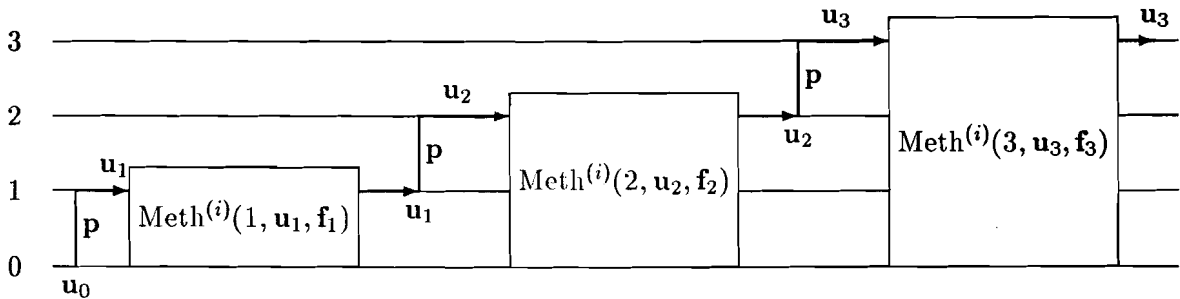


Fig 1.

In §§4-6, for  $\text{Meth}^{(i)}$  we will consider the following two algorithms:

- $\text{Meth}_I^{(i)}$  :  $i$  iterations of the NMGM,
- $\text{Meth}_{II}^{(i)}$ : a Newton iteration combined with  $i$  iterations of a linear multigrid solver.

## 4 Derivation of the Nonlinear Multi-Grid Method

We assume a given finest level, with index  $\ell$ , i.e. we want to solve the problem

$$\mathbf{A}_\ell(\mathbf{u}_\ell^*) = \mathbf{f}_\ell . \quad (4.1)$$

For the nonlinear multi-grid method (NMGM) we need a starting vector on level  $\ell$ . Furthermore, we will use "suitable" (explained below) approximations of  $\mathbf{u}_k^*$  on level  $k < \ell$ . We consider NMGM in a Nested Iteration approach. Therefore it is reasonable to assume that the problems on level  $k < \ell$  have been solved sufficiently accurate, such that we have approximations  $\mathbf{u}_k$  with

$$\mathbf{u}_k \in B(\mathbf{u}_k^*; \beta_k \varepsilon_k), \quad 0 < \beta_k < 1, \quad 0 < k < \ell . \quad (4.2)$$

We also assume that

$$\mathbf{p}\mathbf{u}_{\ell-1}^* \in B(\mathbf{u}_\ell^*, \alpha\varepsilon_\ell), \quad \text{with } 0 < \alpha < 1, \quad (4.3)$$

holds. This imposes a restriction on the coarseness of the grid on level  $\ell - 1$  compared to the grid on level  $\ell$ . For the starting vector on level  $\ell$  we take

$$\mathbf{u}_\ell := \mathbf{p}\mathbf{u}_{\ell-1}.$$

Using (4.2), (4.3) we then have

$$\begin{aligned} \|\mathbf{u}_\ell - \mathbf{u}_\ell^*\| &\leq \|\mathbf{p}\mathbf{u}_{\ell-1} - \mathbf{p}\mathbf{u}_{\ell-1}^*\| + \|\mathbf{p}\mathbf{u}_{\ell-1}^* - \mathbf{u}_\ell^*\| \\ &\leq (\|\mathbf{p}\|\beta_{\ell-1}\varepsilon_{\ell-1}\varepsilon_\ell^{-1} + \alpha)\varepsilon_\ell \end{aligned}$$

We assume that the problem on level  $\ell - 1$  has been solved sufficiently accurate, such that for  $\beta_\ell := \|\mathbf{p}\|\beta_{\ell-1}\varepsilon_{\ell-1}\varepsilon_\ell^{-1} + \alpha$  we have

$$\beta_\ell < 1. \quad (4.4)$$

Now define

$$\hat{\beta}_\ell := \max_{0 < k \leq \ell} \beta_k. \quad (4.5)$$

Then  $\hat{\beta}_\ell < 1$  holds and approximations  $\mathbf{u}_k$  of  $\mathbf{u}_k^*$  are available with

$$\mathbf{u}_k \in B(\mathbf{u}_k^*; \hat{\beta}_\ell \varepsilon_k), \quad 0 < k \leq \ell. \quad (4.6)$$

In the NMGM coarse grid problems will appear that are *perturbations* of the discrete problem  $\mathbf{A}_k(\mathbf{u}_k^*) = \mathbf{f}_k$ ,  $k < \ell$ . Therefore, instead of the problem in (4.1) we consider a more general situation with a problem

$$\mathbf{A}_k(\mathbf{v}_k^*) = \mathbf{g}_k, \quad 0 < k \leq \ell, \quad (4.7a)$$

in which  $\mathbf{v}_k^*$  is such that we have

$$\|\mathbf{v}_k^* - \mathbf{u}_k\| + \|\mathbf{v}_k^* - \mathbf{u}_k^*\| \leq \varepsilon_k, \quad 0 < k \leq \ell. \quad (4.7b)$$

(Note that due to (4.6) the estimate in (4.7b) holds if we take  $\mathbf{g}_k = \mathbf{f}_k$ ,  $\mathbf{v}_k^* = \mathbf{u}_k^*$ ). For the problem in (4.7a) we use a "smoothing" operator

$$\mathcal{S}_k(\cdot, \mathbf{g}_k) : B_k \rightarrow U_k, \quad \text{with } \mathcal{S}_k(\mathbf{v}_k^*, \mathbf{g}_k) = \mathbf{v}_k^*.$$

We assume (cf. also Remark 4.1d) that the following holds:

$$\|\mathcal{S}_k(\mathbf{u}, \mathbf{g}_k) - \mathbf{v}_k^*\| \leq \|\mathbf{u} - \mathbf{v}_k^*\| \text{ for all } \mathbf{u} \in B_k. \quad (4.8)$$



As in the multigrid approach for linear problems, we need a reasonable coarse grid approximation of  $(D\mathbf{A}_k(\mathbf{u}_k^*))^{-1}$ . Smoothing and coarse grid approximation can be made explicit by the following *smoothing property* (here on  $B_k$  instead of on  $U_k$ ) and *approximation property* (cf. [3]):

There are constants  $c_A$  and  $\alpha$  and a function  $\eta(\nu)$  with  $\eta(\nu) \downarrow 0$  for  $\nu \rightarrow \infty$  such that for all  $0 < k \leq \ell$  we have:

$$\|D\mathbf{A}_k(\mathbf{u}_k^*)(\mathcal{S}_k^\nu(\mathbf{u}, \mathbf{g}_k) - \mathbf{v}_k^*)\| \leq \eta(\nu)h_k^{-\alpha}\|\mathbf{u} - \mathbf{v}_k^*\| \text{ for all } \mathbf{u} \in B_k ,$$

and

$$\|(D\mathbf{A}_k(\mathbf{u}_k^*))^{-1} - \mathbf{p}(D\mathbf{A}_{k-1}(\mathbf{u}_{k-1}^*))^{-1}\mathbf{r}\| \leq c_A h_k^\alpha .$$

The constants  $c_A$  and  $\alpha$  and the function  $\eta(\nu)$  should be independent of  $\ell$  and of  $\mathbf{v}_k^*$ .

Here, for simplicity, we used only one norm  $\|\cdot\|$  on  $U_k$ . Using several different norms, a more general setting can be obtained (cf. [3]).

We now derive the NTGM for solving  $\mathbf{A}_k(\mathbf{v}_k^*) = \mathbf{g}_k$  (cf. (4.7)) following the same line of reasoning as in the linear two-grid approach. For the starting vector we take  $\mathbf{u}_k$  as in (4.6).

We apply  $\nu_1$  smoothing iterations:  $\tilde{\mathbf{u}} := \mathcal{S}_k^{\nu_1}(\mathbf{u}_k, \mathbf{g}_k)$ . Using (4.7b) and (4.8) we have

$$\begin{aligned} \|\tilde{\mathbf{u}} - \mathbf{u}_k^*\| &\leq \|\mathcal{S}_k^{\nu_1}(\mathbf{u}_k, \mathbf{g}_k) - \mathbf{v}_k^*\| + \|\mathbf{v}_k^* - \mathbf{u}_k^*\| \\ &\leq \|\mathbf{u}_k - \mathbf{v}_k^*\| + \|\mathbf{v}_k^* - \mathbf{u}_k^*\| \leq \varepsilon_k , \end{aligned}$$

so  $\tilde{\mathbf{u}} \in B_k$  holds. Now note that with  $s \in \mathbb{R}$ ,  $s \neq 0$ , we have

$$\begin{aligned} \mathbf{A}_k(\mathbf{v}_k^*) - \mathbf{A}_k(\tilde{\mathbf{u}}) &= \mathbf{g}_k - \mathbf{A}_k(\tilde{\mathbf{u}}) \Leftrightarrow \\ D\mathbf{A}_k(\mathbf{u}_k^*)(\mathbf{v}_k^* - \tilde{\mathbf{u}}) &= \mathbf{g}_k - \mathbf{A}_k(\tilde{\mathbf{u}}) \Leftrightarrow \\ D\mathbf{A}_k(\mathbf{u}_k^*)s(\mathbf{v}_k^* - \tilde{\mathbf{u}}) &= s(\mathbf{g}_k - \mathbf{A}_k(\tilde{\mathbf{u}})) . \end{aligned}$$

So for the scaled error  $\mathbf{e}_k := s(\mathbf{v}_k^* - \tilde{\mathbf{u}})$  we have the *linear* equation

$$D\mathbf{A}_k(\mathbf{u}_k^*)\mathbf{e}_k = s(\mathbf{g}_k - \mathbf{A}_k(\tilde{\mathbf{u}})) . \quad (4.9)$$

As in the case of linear multigrid, the smooth error  $\mathbf{e}_k$  can be approximated on the coarse grid. Let  $\mathbf{e}_{k-1}$  satisfy the corresponding coarse grid equation

$$D\mathbf{A}_{k-1}(\mathbf{u}_{k-1}^*)\mathbf{e}_{k-1} = sr(\mathbf{g}_k - \mathbf{A}_k(\tilde{\mathbf{u}})) . \quad (4.10)$$

Now take  $|s|$  small enough such that

$$\mathbf{u}_{k-1} + \mathbf{e}_{k-1} \in B(\mathbf{u}_{k-1}^*, \tilde{\beta}_\ell \varepsilon_{k-1}), \quad \text{with } \tilde{\beta}_\ell := \frac{1}{2}(1 + \hat{\beta}_\ell), \quad (4.11)$$

holds ( $\hat{\beta}_\ell$  as in (4.5)). Sufficient conditions are given by

$$|s| \|(D\mathbf{A}_{k-1}(\mathbf{u}_{k-1}^*))^{-1} \mathbf{r}(\mathbf{g}_k - \mathbf{A}_k(\tilde{\mathbf{u}}))\| \leq \frac{1}{2}(1 - \hat{\beta}_\ell) \varepsilon_{k-1} \quad (4.12a)$$

or

$$|s| \leq \|(D\mathbf{A}_{k-1}(\mathbf{u}_{k-1}^*))^{-1} \mathbf{r} D\mathbf{A}_k(\mathbf{u}_k^*)\|^{-1} \frac{1}{2}(1 - \hat{\beta}_\ell) \varepsilon_{k-1} / \varepsilon_k. \quad (4.12b)$$

Then for  $\mathbf{e}_{k-1}$  we have the following equivalent equations (use (2.4))

$$\begin{aligned} D\mathbf{A}_{k-1}(\mathbf{u}_{k-1}^*) \mathbf{e}_{k-1} &= s \mathbf{r}(\mathbf{g}_k - \mathbf{A}_k(\tilde{\mathbf{u}})) \Leftrightarrow \\ \mathbf{A}_{k-1}(\mathbf{u}_{k-1} + \mathbf{e}_{k-1}) - \mathbf{A}_{k-1}(\mathbf{u}_{k-1}) &= s \mathbf{r}(\mathbf{g}_k - \mathbf{A}_k(\tilde{\mathbf{u}})) \Leftrightarrow \\ \mathbf{e}_{k-1} &= \mathbf{A}_{k-1}^{-1}[\mathbf{A}_{k-1}(\mathbf{u}_{k-1}) + s \mathbf{r}(\mathbf{g}_k - \mathbf{A}_k(\tilde{\mathbf{u}}))] - \mathbf{u}_{k-1}. \end{aligned}$$

Based on  $\mathbf{p}\mathbf{e}_{k-1} \approx \mathbf{e}_k = s(\mathbf{v}_k^* - \tilde{\mathbf{u}})$ , we define the new iterate  $\mathbf{u}_{\text{new}}$  for approximating  $\mathbf{v}_k^*$  by

$$\mathbf{u}_{\text{new}} = \tilde{\mathbf{u}} + \frac{1}{s} \mathbf{p}\mathbf{e}_{k-1} = \tilde{\mathbf{u}} + \frac{1}{s} \mathbf{p}\{\mathbf{A}_{k-1}^{-1}[\mathbf{A}_{k-1}(\mathbf{u}_{k-1}) + s \mathbf{r}(\mathbf{g}_k - \mathbf{A}_k(\tilde{\mathbf{u}}))] - \mathbf{u}_{k-1}\}. \quad (4.13)$$

The iteration in (4.13) is the same as one iteration of the NTGM defined in §1. with  $\nu_2 = 0$  (no post-smoothing). From the derivation it is clear that for the parameters  $\tilde{\mathbf{u}}_{k-1}, s$  in the NTGM in §1 we can take  $\tilde{\mathbf{u}}_{k-1} := \mathbf{u}_{k-1}$  and  $s$  "sufficiently small", e.g. as in (4.12).

**Remark 4.1.**

- a) From the derivation of the method above it is clear that for a suitable choice of the parameters  $\tilde{\mathbf{u}}_{k-1}$  and  $s$  we can guarantee that all (intermediate) results remain "close" to  $\mathbf{u}_k^*$ , i.e. in  $B(\mathbf{u}_k^*; \varepsilon_k)$ . This is not guaranteed if we use the FAS algorithm (i.e.  $s = 1, \tilde{\mathbf{u}}_{k-1} = \mathbf{r}\mathbf{u}$ ). This explains why for the NMGM a convergence analysis is available (in [6], [7]), whereas for the FAS algorithm no theoretical convergence results are known. The difference between NMGM and FAS may cause a very different convergence behaviour for "hard" nonlinear problems, as is shown in [5].
- b) The condition for  $s$  in (4.12a) is satisfied if the following holds:

$$|s| \|\mathbf{u}_k - \mathbf{v}_k^*\| \leq \|(D\mathbf{A}_{k-1}(\mathbf{u}_{k-1}^*))^{-1} \mathbf{r} D\mathbf{A}_k(\mathbf{u}_k^*)\|^{-1} \frac{1}{2}(1 - \hat{\beta}_\ell) \varepsilon_{k-1}.$$

So asymptotically, i.e. for  $\mathbf{u}_k \rightarrow \mathbf{v}_k^*, s = 1$  (as in FAS) will suffice.

- c) For  $\mathbf{u}_{\text{new}}$  as in (4.13) the following holds

$$\|\mathbf{u}_{\text{new}} - \mathbf{v}_k^*\| \leq \|(D\mathbf{A}_k(\mathbf{u}_k^*))^{-1} - \mathbf{p}(D\mathbf{A}_{k-1}(\mathbf{u}_{k-1}^*))^{-1} \mathbf{r}\| \|D\mathbf{A}_k(\mathbf{u}_k^*)(S_k^{\nu_1}(\mathbf{u}_k, \mathbf{g}_k) - \mathbf{v}_k^*)\|.$$

So the smoothing property and approximation property defined above, yield the usual two-grid convergence results.

d) Similar results can be obtained if the condition in (4.8) (which may be rather severe for nonlinear problems) is replaced by the condition

$$\|\mathcal{S}_k^\nu(\mathbf{u}, \mathbf{g}_k) - \mathbf{v}_k^*\| \leq C\|\mathbf{u} - \mathbf{v}_k^*\| \quad \text{for } \mathbf{u} \in B_k, \text{ with } C \text{ independent of } \nu .$$

We now discuss the extension of the two-grid method to a multi-grid method. Clearly the given problem on the highest level in (4.1) is of the type as in (4.7) with  $k = \ell$ . The derivation of the NTGM shows that for a problem as in (4.7) on level  $k$  and using the starting vector  $\mathbf{u}_k$  the algorithm is well-defined and we obtain the following coarse grid problem on level  $k - 1$  (cf. (4.13)):

$$\mathbf{A}_{k-1}(\mathbf{v}_{k-1}^*) = \mathbf{A}_{k-1}(\mathbf{u}_{k-1}) + s\mathbf{r}(\mathbf{g}_k - \mathbf{A}_k(\tilde{\mathbf{u}})) , \quad (4.14a)$$

with

$$\mathbf{v}_{k-1}^* := \mathbf{u}_{k-1} + \mathbf{e}_{k-1} \quad (\mathbf{e}_{k-1} \text{ as in (4.10)}) . \quad (4.14b)$$

Assume that  $s$  is such that (4.12a) or (4.12b) is satisfied. Then we have that  $\|\mathbf{e}_{k-1}\| \leq \frac{1}{2}(1 - \hat{\beta}_\ell)\varepsilon_{k-1}$  holds. Thus we obtain using (4.11) that

$$\|\mathbf{v}_{k-1}^* - \mathbf{u}_{k-1}\| + \|\mathbf{v}_{k-1}^* - \mathbf{u}_{k-1}^*\| \leq (\frac{1}{2}(1 - \hat{\beta}_\ell) + \tilde{\beta}_\ell)\varepsilon_{k-1} = \varepsilon_{k-1} \quad (4.15)$$

holds. From (4.14a) and (4.15) we conclude that the coarse grid problem is again of the form (4.7) but now on level  $k - 1$ . Thus we can apply the NTGM on level  $k - 1$  with starting vector  $\mathbf{u}_{k-1}$  for solving the problem in (4.14a) approximately. An induction argument now yields that the multi-grid extension of the NTGM in §1 for solving the problem in (4.1) is well-defined, if on level  $k \leq \ell$  we use  $\mathbf{u}_k$  as the starting vector, take  $\tilde{\mathbf{u}}_{k-1} = \mathbf{u}_{k-1}$  and  $s$  such that (4.12a) or (4.12b) holds.

This results in the Nonlinear Multi-Grid Method (NMGM). This NMGM can be used in a Nested Iteration for solving the problem in (4.1) for increasing values of  $\ell$  (cf. §3).

## 5 Newton iteration and linear multigrid (NewtonMGM)

The Newton iteration, with starting vector  $\mathbf{u} \in B_k$ , for solving the problem in (4.1) is given by

$$\mathbf{N}_\ell(\mathbf{u}) = \mathbf{u} - (D\mathbf{A}_\ell(\mathbf{u}))^{-1}(\mathbf{A}_\ell(\mathbf{u}) - \mathbf{f}_\ell) . \quad (5.1)$$

So, a linear problem of the form

$$D\mathbf{A}_\ell(\mathbf{u})\mathbf{w}_\ell^* = \mathbf{b}_\ell \quad (5.2)$$

has to be solved. Note that due to (2.4a) ( $\mathbf{A}_\ell$  is locally affine) the identity  $D\mathbf{A}_\ell(\mathbf{u}) = D\mathbf{A}_\ell(\mathbf{u}_\ell^*)$  holds. For solving the linear problem in (5.2) we use a standard linear multigrid method with smoothers denoted by  $S_\ell(\cdot, \mathbf{b}_\ell)$  and coarse grid operators  $D\mathbf{A}_{k-1}(\mathbf{u}_{k-1}^*)$  ( $1 \leq k \leq \ell$ ). The corresponding smoothing and approximation properties are:

There are constants  $c_A$  and  $\alpha$  and a function  $\eta(\nu)$  with  $\eta(\nu) \downarrow 0$  if  $\nu \rightarrow \infty$  such that for all  $0 < k \leq \ell$  we have

$$\|DA_k(\mathbf{u}_k^*)(S_k^\nu(\mathbf{v}, \mathbf{b}_k) - \mathbf{w}_k^*)\| \leq \eta(\nu)h_k^{-\alpha}\|\mathbf{v} - \mathbf{w}_k^*\| \text{ for all } \mathbf{v} \in U_k$$

and

$$\|(DA_k(\mathbf{u}_k^*))^{-1} - \mathbf{p}(DA_{k-1}(\mathbf{u}_{k-1}^*))^{-1}\mathbf{r}\| \leq c_A h_k^\alpha .$$

The constants  $c_A$  and  $\alpha$  and the function  $\eta(\nu)$  should be independent of  $\ell$  and of  $\mathbf{w}_k^*$ .

**Remark 5.1.**

- a) In the NewtonMGM the Jacobians  $DA_k(\mathbf{u}_k)$  are used. Note that for  $\mathbf{u} \in B(\mathbf{u}_k^*, \varepsilon_k)$ ,  $\mathbf{w} \in U_k$  and  $s \in \mathbb{R} \setminus \{0\}$  with  $|s|$  small enough such that  $\mathbf{u} + s\mathbf{w} \in B(\mathbf{u}_k^*, \varepsilon_k)$  we have (cf. (2.4))

$$DA_k(\mathbf{u})(\mathbf{w}) = \frac{1}{s}DA_k(\mathbf{u})(s\mathbf{w}) = \frac{1}{s}(\mathbf{A}_k(\mathbf{u} + s\mathbf{w}) - \mathbf{A}_k(\mathbf{u})) . \quad (5.3)$$

Using (5.3) to evaluate a Jacobian, will result in conditions on  $s$  that are similar to the conditions on  $s$  in §4.

- b) The NewtonMGM can be used in a Nested Iteration for solving the problem in (4.1) for increasing values of  $\ell$  (cf. §3).

## 6 Comparison of NMGM and NewtonMGM

We briefly compare the two methods discussed in §4 and in §5.

Both, in NMGM and in NewtonMGM, we need  $\mathbf{r}, \mathbf{p}, \mathbf{A}_k$ . In NMGM we need a *nonlinear* smoother (e.g. Jacobi, Gauss–Seidel), whereas in NewtonMGM we need a *linear* smoother (e.g. Jacobi, Gauss–Seidel, ILU). There are many (theoretical) results available for linear smoothers (cf. [8]), whereas for nonlinear smoothers there are hardly any results. In NewtonMGM we use (an approximation of) the Jacobian  $DA_k(\mathbf{u})$ , whereas in NMGM we (only) use the parameters  $\tilde{\mathbf{u}}_{k-1}$  and  $s$ .

The algorithmic structure of a NMGM is simpler than the algorithmic structure of a NewtonMGM. For the NMGM the structure is the same as in a linear multigrid algorithm. In NewtonMGM such a linear multigrid algorithm is used as an inner iteration and combined with a Newton outer iteration. Note that in the outer iteration the Jacobian will change (if the problem is not locally affine) and for each Jacobian we need a whole set of corresponding coarse grid problems in the linear multigrid method.

In the situation of our locally affine model problem the approximation property is the same for both methods. Also the smoothing property is the same if we use "related" smoothers  $\mathcal{S}_k$  and  $S_k$ . For example, for a given linear operator  $\mathbf{W}_k$  we define

$$\mathcal{S}_k(\mathbf{u}, \mathbf{g}_k) = \mathbf{u} - \mathbf{W}_k^{-1}\mathbf{A}_k(\mathbf{u}) + \mathbf{W}_k^{-1}\mathbf{g}_k \quad (\text{cf. §4})$$

and

$$S_k(\mathbf{v}, \mathbf{b}_k) = \mathbf{v} - \mathbf{W}_k^{-1} D\mathbf{A}_k(\mathbf{u}_k^*)\mathbf{v} + \mathbf{W}_k^{-1}\mathbf{b}_k \quad (\text{cf. §5})$$

Then, using  $\mathbf{g}_k = \mathbf{A}_k(\mathbf{v}_k^*)$ ,  $\mathbf{b}_k = D\mathbf{A}_k(\mathbf{u}_k^*)\mathbf{w}_k^*$  we obtain

$$\mathcal{S}_k^\nu(\mathbf{u}, \mathbf{g}_k) - \mathbf{v}_k^* = (\mathbf{I} - \mathbf{W}_k^{-1} D\mathbf{A}_k(\mathbf{u}_k^*))^\nu(\mathbf{u} - \mathbf{v}_k^*)$$

and

$$\mathcal{S}_k^\nu(\mathbf{v}, \mathbf{b}_k) - \mathbf{w}_k^* = (\mathbf{I} - \mathbf{W}_k^{-1} D\mathbf{A}_k(\mathbf{u}_k^*))^\nu(\mathbf{v} - \mathbf{w}_k^*),$$

so in both smoothers the error iteration is the same.

If the nonlinear problem is not locally affine, satisfactory convergence results can still be proved. In that case, however, the analysis is rather technical because one has to estimate the second order perturbations (cf. [6], [7]).

For both methods, NewtonMGM and NMGM, "damped" versions exist, which for a suitable class of problems enlarge the domain of converge (cf. [1], [4]). For both methods the combination with Nested Iteration is very useful to obtain suitable starting vectors. Furthermore, the Nested Iteration yields suitable coarse grid approximations ( $\tilde{\mathbf{u}}_{k-1}$ , cf. §1) that are needed in the NMGM.

## References

- [1] Bank, R.E. and Rose, D.J. *Analysis of a multilevel iterative method for nonlinear finite element equations*. Math Comp. 39 (1982), 453-465.
- [2] Brandt, A.: *Multi-level adaptive solutions to boundary-value problems*. Math. Comp. 31 (1977), 333-390.
- [3] Hackbusch, W.: *Multi-grid Methods and Applications*. Springer-Verlag, Berlin (1985).
- [4] Hackbusch, W. and Reusken, A.: *Analysis of a damped nonlinear multilevel method*. Numer. Math. 55 (1989), 225-246.
- [5] Hackbusch, W.: *Comparisons of Different Multi-grid Variants for Nonlinear Equations*. ZAMM 72 (1992), 148-151.
- [6] Reusken, A.: *Convergence of the Multigrid Full Approximation Scheme for a class of elliptic mildly nonlinear boundary value problems*. Numer. Math. 52 (1988), 251-277.
- [7] Reusken, A.: *Convergence of the Multilevel Full Approximation Scheme including the V-cycle*. Numer. Math. 53 (1988), 663-686.
- [8] Wesseling, P.: *An Introduction to Multigrid Methods*. Wiley, Chichester (1992).