Numer. Math. 52, 251–277 (1988)

Numerische Mathematik

Convergence of the Multigrid Full Approximation Scheme for a Class of Elliptic Mildly Nonlinear Boundary Value Problems

Arnold Reusken

Mathematical Institute, State University of Utrecht, Budapestlaan 6, P.O. Box 80.010, 3508 Utrecht, The Netherlands

Summary. The multigrid full approximation scheme (FAS MG) is a well-known solver for nonlinear boundary value problems. In this paper we restrict ourselves to a class of second order elliptic mildly nonlinear problems and we give local conditions, e.g. a local Lipschitz condition on the derivative of the continuous operator, under which the FAS MG with suitably chosen parameters locally converges. We prove quantitative convergence statements and deduce explicit bounds for important quantities such as the radius of a ball of guaranteed convergence, the number of smoothings needed, the number of coarse grid corrections needed and the number of FAS MG iterations needed in a nested iteration. These bounds show well-known features of the FAS MG scheme.

Subject Classifications: AMS(MOS): 65N20; CR: 5.15.

1. Introduction

We consider the following class of second order elliptic mildly nonlinear boundary value problems on an open, connected, bounded domain $\Omega \subset \mathbb{R}^2$ with smooth boundary $\partial\Omega$:

$$\begin{cases} -\nabla.(a\nabla u) + b\, g \circ u = f & \text{on } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

with $a \in C^1(\bar{\Omega})$, $\min\{a(x) \mid x \in \bar{\Omega}\} > 0$, $b \in C(\bar{\Omega})$, $\min\{b(x) \mid x \in \bar{\Omega}\} \geqq 0$, $f \in L^2(\Omega)$, $g \in C^1(\mathbb{R})$, $g'(t) \geqq 0$ for all $t \in \mathbb{R}$.

Existence and uniqueness of a solution of such a boundary value problem is proved in literature using the global monotonicity of the problem. We reformulate the above problem in an equivalent variational form and give a finite element discretization. Now, after discretization we have a nonlinear system of equations to be solved. Two methods frequently used in a multigrid context to solve such a system of equations are the following. One can use an iterative linearisation process like (a modified) Newton's method combined with a linear multigrid method to solve the linear equations occurring in this Newton process. Another method is a nonlinear multigrid method, known in literature as Full Approximation Scheme (FAS; Brandt in [2]) or Nonlinear Multigrid Iteration (Hackbusch in [6]). With respect to the first method we note that for the Newton iteration and the linear multigrid solvers seperately many theoretical analyses of convergence properties can be found in literature (e.g. [9, 1, 6]). However the only analysis of the first method known to the author is [10]. One might expect (for experimental results see ch. I in [7]) that asymptotically, i.e. in a small enough neighbourhood of the solution, the two methods (Newton + lin. MG and FAS) behave similarly.

A convergence proof of a FAS-iteration scheme is given (only) by Hackbusch in e.g. [6]. Hackbusch imposes general conditions on for example the derivative of the discrete operator and the derivative of the relaxation operator used. Under these general conditions he deduces the qualitative result that on a fine enough discretization level in a small enough neighbourhood of the discrete solution the FAS scheme with a bounded number of smoothings per iteration and a bounded number of coarse grid corrections converges. He also proves that asymptotically, under conditions, the convergence factor agrees with what one expects.

In this paper we only consider FAS. We abandon generality and restrict ourselves to the class of nonlinear problems defined above. We do note that in our analysis of the FAS iteration all conditions and results are local ones and thus our analysis can be applied to nonlinear problems which locally (in a neighbourhood of the solution) behave like the above stated (monotone) problems. Having restricted to the above nice class of nonlinear problems we define a class of suitable nonlinear Jacobi-like smoothing operators and a suitable FAS iteration and prove a convergence statement in which quantitative statements about e.g. a domain of guaranteed convergence, the number of smoothings needed, the number of coarse grid corrections needed and the coarsest acceptable grid are given. As Hackbusch does, we prove convergence by linearising the FAS iteration. In essence the only assumption we make about the continuous operator (apart from the monotonicity) is that its derivative satisfies a local Lipschitz condition and a regularity condition. We discuss properties of g which induce that these conditions are fulfilled.

Because of our more quantitative convergence statement we can prove among other things the following features of a FAS iteration with suitably (we will specify what is "suitable") chosen parameters applied to the above mentioned nonlinear system of equations: (more precise statements can be found further on)

- the FAS two grid scheme (FAS TG) with one pre-smoothing and one post-smoothing locally converges (the smoothers are nonlinear Jacobi-like smoothers).
- under natural conditions the FAS multigrid (FAS MG) with W-cycles locally converges.
- on fine enough grids (levels) there is a ball of guaranteed convergence (of FAS TG) with a radius that is about inversely proportional to a local Lip-

schitz constant of the derivative of the continuous operator (just as in many Newton like iterations). So this radius is in essence independent of the level.
- for FAS MG convergence, the coarsest level (grid) used should be such that the discretization error on that level is smaller than the radius of the ball of guaranteed TG convergence.
- if FAS MG is used in a nested iteration, we need only a good enough starting vector on a coarsest (but fine enough) level to generate acceptable starting vectors for MG on finer levels. Suppose we have suitable finite element spaces $(S_k)_{k=0,1,\dots}$ on a sequence of ever refining grids (indexed by k). Now suitable FAS MG generates a sequence of approximations $u_k \in S_k$, $k=1, 2, \dots$ (approximating the discrete solution on level k) with error smaller than the relative discretization error on level k. In this nested iteration we need fewer FAS MG iterations for larger k, and in a FAS MG iteration on level k we need fewer coarse grid corrections for larger k. Moreover the number of FAS MG iterations needed and the number of the coarse grid corrections needed tend ot expected lower bounds (for $k \to \infty$).

In this paper we use the convergence theory for linear TG methods for symmetric elliptic problems as in [1]. The idea of linearisation is easy but our proofs are somewhat technical because we need uniformity in some parameters and we have to deal with (complicated) higher order terms occurring after linearisation of the FAS TG iteration.

The paper is organized as follows.

In §2 we consider the continuous problem and a finite element discretization. We discuss existence and uniqueness of solutions.

In §3 we collect notations and conventions and prove relations between different energynorms. In §4 we introduce a class of linear two grid operators and prove a uniform convergence statement.

In §5 a FAS two grid iteration is defined and after linearisation, using estimates for higher order terms and the convergence thm. of §4, a local convergence statement is proved. In §6 we deal with a FAS multigrid iteration and prove convergence using the FAS two grid convergence theorem of §5. Finally in §7 we consider a nested iteration using FAS multigrid (Full Multigrid Algorithm). In §8 we discuss simple conditions on the function g such that the assumptions we make in §3–7 about the continuous problem are fulfilled.

2. A Class of Mildly Nonlinear Differential Equations and their Discretization

In this section we consider a class of second order elliptic mildly nonlinear boundary value problems and their finite element discretizations. The continuous and discrete problems have unique solutions and assuming a Lipschitz condition a natural bound for the discretization error can be given.

Definitions 2.1. Let $\Omega \subset \mathbb{R}^2$ be an open, connected, bounded area with smooth boundary. By $H^k(\Omega)$ $(k \in \mathbb{N})$ we denote the Sobolev space of all functions $u \in L^2(\Omega)$ whose distributional derivatives $D^\alpha u$ for $|\alpha| \leq k$ are elements of the space $L^2(\Omega)$.

For $u \in H^k(\Omega)$ its norm $\|u\|_k$ is defined by $\|u\|_k = (\sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^2}^2)^{\frac{1}{2}}$. We also define $H_0^k(\Omega)$ to be the closure in $H^k(\Omega)$ of $D(\Omega) := \{\phi \in C^\infty(\Omega) | \operatorname{supp}(\phi) \subset \Omega\}$. On $H_0^1(\Omega)$ the norm $\|.\|_1$ is equivalent to $\|.\|$ which is given by $\|u\| = (u, u)^{\frac{1}{2}}$ where $(u, v) = \int \nabla u . \nabla v \, dx \, (u, v \in H_0^1(\Omega))$.

2.2. *The Problem Considered and its Discretization*

We consider the following class of second order elliptic nonlinear boundary value problems:

$$(1) \qquad \begin{cases} -\nabla.(a \nabla u) + b g \circ u = f & \text{on } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

Here $a \in C^1(\bar{\Omega})$, $\min\{a(x) | x \in \bar{\Omega}\} > 0$, $b \in C(\bar{\Omega})$, $\min\{b(x) | x \in \bar{\Omega}\} \geq 0$, $f \in L^2(\Omega)$ and $g \in C^1(\mathbb{R})$, $g'(t) \geq 0$ for all $t \in \mathbb{R}$.

We also give a variational formulation of (1):

(2) Find $u \in H_0^1(\Omega)$ such that: $a(u, v) + (b g \circ u, v)_{L^2} = (f, v)_{L^2}$ for all $v \in H_0^1(\Omega)$. Here $(.,.)_{L^2}$ is the L^2 inner product and $a(u, v) = \int_\Omega a \nabla u . \nabla v \, dx$. The operator $u \to a(u, .) + (b g \circ u, .)_{L^2}$ will be denoted by n.

We assume a sequence of standard linear finite element spaces (i.e., a regular affine family of continuous bilinear finite elements; see e.g. §3.2 in [5]), denoted by $S_0 \subset S_1 \subset \ldots \subset H_0^1(\Omega)$ with corresponding "stepsizes" $h_0 > h_1 > \ldots$. We assume $0 < K_1 \leq \frac{h_{i+1}}{h_i} \leq K_2 < 1$ for $i \geq 0$. Let $\{\psi_1, \psi_2, \ldots, \psi_{n_k}\}$ be the standard bilinear finite element basis of S_k, then we define $U_k = \mathbb{R}^{n_k}$ and the isomorphism $P_k \colon U_k \to S_k$, $P_k \alpha = \sum_{j=1}^{n_k} \alpha^{(j)} \psi_j$ where $\alpha^{(j)}$ is the j-th coordinate of $\alpha \in U_k$.

Now a finite element discretization of (2) is given by:

(3) Find $u_k \in S_k$ such that: $a(u_k, v_k) + (b g \circ u_k, v_k)_{L^2} = (f, v_k)_{L^2}$ for all $v_k \in S_k$.

And an equivalent formulation is given by:

(4) Find $\alpha_k \in U_k$ such that $N_k(\alpha_k) := A_k \alpha_k + \bar{g}_k(\alpha_k) = b_k$, where A_k is a standard Poisson discretization: $(A_k)_{i,j} = h_k^{-2} a(\psi_i, \psi_j)$, and

$$(\bar{g}_k(\alpha_k))^{(j)} = h_k^{-2} \int_\Omega b(x) \, g(P_k \alpha_k) \, \psi_j(x) \, dx, \qquad b_k^{(j)} = h_k^{-2} \int_\Omega f(x) \, \psi_j(x) \, dx.$$

2.3. *Existence and Uniqueness of Solutions*

From [4] one can deduce that there exists a unique solution of (1) in $H^2(\Omega) \cap H_0^1(\Omega)$, say u^*.

Clearly u^* is a solution of (2). Monotonicity, i.e. $a(u_1 - u_2, u_1 - u_2) + (b(g \circ u_1 - g \circ u_2), u_1 - u_2)_{L^2} > 0$ for all $u_1 \neq u_2 \in H_0^1(\Omega)$, guarantees uniqueness, so u^* is also the unique solution of (2).

Existence and uniqueness of a solution of (3) (and thus (4)) can be shown by using monotonicity arguments as in the proof of existence and uniqueness of the continuous solution (actually: N_k is maximal monotone and coercive, see [3]). Another possibility, which in addition yields information about the smoothness of the inverse, is to apply a suitable (global) version of the implicit function theorem. Using Theorem 15.4 and 15.2 in [12] results in: N_k is a C^1-diffeomorphism on U_k. The solution of (3) is denoted by u_k^*.

Remark 2.4. In the remainder of this paper we will often use the following Sobolev embedding theorem (see e.g. [5]):

$$\forall q \in [1, \infty[\ \exists d_q \in \mathbb{R}: \forall u \in H_0^1(\Omega): \|u\|_{L^q} \leqq d_q \|u\|.$$

We define the function I: $[1, \infty[\to]0, \infty[$ by:

$$I(q) = \sup\{\|u\|_{L^q} \|u\|^{-1} \mid u \in H_0^1(\Omega), u \neq 0\}.$$

Lemma 2.5. *Let* $a_- := \min\{a(x) \mid x \in \bar{\Omega}\}$, $b_+ := \max\{b(x) \mid x \in \bar{\Omega}\}$. $H_0^1(\Omega)$ *is denoted by* H. $U := I(2)\, a_-^{-1}(b_+ \|g(0)\|_{L^2} + \|f\|_{L^2})$ *and* $B := \{v \in H \mid \|v\| \leqq U\}$.

Assume n maps B into H' and $|(n(v) - n(w))(u)| \leqq C_B \|v - w\| \, \|u\|$ for all $v, w \in B$, all $u \in H$ (cf. 8.4 (a_2)). Then with u^* and u_k^* as in 2.3 the following holds:

$$\exists c > 0: \forall k \geqq 0 \qquad \|u^* - u_k^*\| \leqq c h_k.$$

Proof. Our proof runs as the proof of Theorem 5.3.4 in [5]. We have:

$$\begin{aligned}\|u^*\|^2 &\leqq a_-^{-1} a(u^*, u^*) = a_-^{-1}\{n(u^*)(u^*) - (b(g \circ u^* - g(0)), u^*)_{L^2} - (b g(0), u^*)_{L^2}\} \\ &\leqq a_-^{-1}\{(f, u^*)_{L^2} - (b g(0), u^*)_{L^2}\} \leqq a_-^{-1} I(2)(\|f\|_{L^2} + b_+ \|g(0)\|_{L^2}) \|u^*\|.\end{aligned}$$

So: $\|u^*\| \leqq U$. The same bound holds for u_k^*.
Now for arbitrary $w \in S_k$ we have:

$$\begin{aligned}\|u^* - u_k^*\|^2 &\leqq a_-^{-1}(n(u^*) - n(u_k^*))(u^* - u_k^*) = a_-^{-1}(n(u^*) - n(u_k^*))(u^* - w) \\ &\leqq a_-^{-1} C_B \|u^* - u_k^*\| \, \|u^* - w\|.\end{aligned}$$

This implies $\|u^* - u_k^*\| \leqq a_-^{-1} C_B \inf\{\|u^* - w\| \mid w \in S_k\}$. Now the use of standard finite element estimates and the fact that $u^* \in H^2(\Omega) \cap H_0^1(\Omega)$ proves the lemma. □

3. Assumptions, Definitions and Fundamental Relations

In the remainder of this paper we assume that the nonlinearity of the problem as stated in 2.2.(2) fulfils certain conditions. These conditions are stated in 3.1 and 4.7. We comment on these in 3.3, 4.8 and 5.15.

In this section we also define norms and operators that we will frequently use further on and we prove equivalence of a class of norms. In 3.5 we give a bound for the error made by linearizing the nonlinear problem 2.2.(2).

Assumption 3.1. We denote $H_0^1(\Omega)$ by H. In Sect. 3–7 we assume that the operator $n: u \to a(u, .) + (b g \circ u, .)_{L^2}$ satisfies:

(1) n maps H into H' (the space of all continuous linear functionals on H).
(2) n is Fréchet differentiable on H (Fréchet derivative denoted by Dn).

(3) for all $u \in H$ the operator $(v, w) \to Dn(u)(v)(w)$ defines a symmetric, bilinear form on $H \times H$.
(4) for every bounded subset B of H there exists a constant $\tilde{\Gamma}_B$ such that: $\forall v, w \in B \; \forall u, z \in H$: $|(Dn(v) - Dn(w))(u)(z)| \leqq \tilde{\Gamma}_B \|v - w\| \, \|u\| \, \|z\|$.

Sufficient conditions on the function g implying assumption 3.1 will be given in §8.

Using (1) and (2) of 3.1 and because $(n(u+h) - n(u))(h) \geqq a(h, h)$ for all $u, h \in H_0^1(\Omega)$ (results from $g' \geqq 0$) we obtain that $Dn(u)(h)(h) \geqq a(h, h)$. And thus $(v, w) \to Dn(u)(v)(w)$ is coercive on $H_0^1(\Omega) \times H_0^1(\Omega)$.

Definitions and Conventions 3.2.

– The space $H_0^1(\Omega)$ is denoted by H (with inner product $(.,.)$ and norm $\|.\|$ as in 2.1).
– We recall the definition of the bilinear form a: $H \times H \to \mathbb{R}$: $a(u, v) = \int_\Omega \mathrm{a} \nabla u . \nabla v \, dx$. We define a norm $\|.\|_*$ on H by $\|u\|_* := \sqrt{a(u, u)}$.
– For $u \in H$ the symmetric continuous coercive bilinear form on $H \times H$ given by $(v, w) \to Dn(u)(v)(w)$ is denoted by b_u (also used as operator $H \to H'$). We also define the norm $\|.\|_u$ on H by $\|v\|_u := \sqrt{b_u(v, v)}$.
– We use an obvious notation for operator norms, e.g.: j: $H \to H'$, $u \in H$ then $\|j\|_u = \sup\{|j(v)| \mid v \in H, \; \|v\|_u \leqq 1\}$.
– For j: $H \to H'$ we denote $j_{|S_k}$: $S_k \to S'_k$ by j_k and $j_{|S_k}$: $S_k \to S'_{k-1}$ by $j_{k,k-1}$.
– We define the reals a_-, a_+, b_-, b_+ by:

$$a_- = \min\{a(x) \mid x \in \bar{\Omega}\}, \; a_+ = \max\{a(x) \mid x \in \bar{\Omega}\},$$
$$b_- = \min\{b(x) \mid x \in \bar{\Omega}\}, \; b_+ = \max\{b(x) \mid x \in \bar{\Omega}\}.$$

– Let w^* be an element of $H \cap C(\bar{\Omega})$. *In Sect. 3, 4, 5, 8 we keep this w^* fixed and in 5.13, § 6 and § 7 we will specify a suitable w^*.*
– For $v \in H$, $r \in [0, \infty[$ we define $B(v; r) := \{u \in H \mid \|u - v\|_v \leqq r\}$.
Also: $B^k(v; r) := B(v; r) \cap S_k$, $B(r) := B(w^*; r)$, $B^k(r) := B(r) \cap S_k$.
– With $\tilde{\Gamma}_B$ from 3.1.(4) we define $\Gamma_r := a_-^{-1\frac{1}{2}} \tilde{\Gamma}_{B(r)}$.

Remark 3.3. With respect to assumption 3.1 we note the following. The conditions in 3.1 may be weakened in the sense that there is *some* open ball B containing w^* on which n is differentiable (with derivative as in 3.1(3)) and on which Dn satisfies a Lipschitz condition as in 3.1.(4).

Our analysis is then still applicable and our results do not really change (they only become "more local"). However these weaker conditions cause additional technical difficulties because we have to make sure that, after choosing a starting vector (for the FAS algorithm) within B, "everything remains within B". These additional difficulties we have wanted to avoid.

We also note that the analysis and results can be adapted in a straightforward manner if we only assume a Hölder condition on Dn instead of a Lipschitz condition. The condition (3) we consider to be reasonable because in standard convergence theory for linear multigrid methods [1, 6, 11] assuming symmetry is common practice (to make possible and analysis using energy norms).

We comment on the condition (4) in 5.15.

The following lemma compares different inner products and all sorts of equivalent norms. In the remainder we will often use estimates from this lemma.

Lemma 3.4. *The following holds:*

(1) $|b_v(u,z)-b_w(u,z)| \leqq \tilde{\Gamma}_{B(r)} \|v-w\| \, \|u\| \, \|z\| \leqq \Gamma_r \|v-w\|_{y_1} \|u\|_{y_2} \|z\|_{y_3}$ *for all* $v, w \in B(r)$, *all* $y_i, u, z \in H$.

There exists a function $G: H \to [0, \infty[$ *such that the following holds:*

(2) $|b_v(u,z)-a(u,z)| \leqq G(v) \|u\| \, \|z\|$ *for all* $u, v, z \in H$;

(3) $a_- \|u\|^2 \leqq \|u\|_*^2 \leqq \|u\|_v^2 \leqq (a_+ + G(v)) \|u\|^2$ *for all* $u, v \in H$;

(4) $\|u\|_v^2 \leqq (1 + a_-^{\frac{1}{2}} \Gamma_r \|v-w\|) \|u\|_w^2 \leqq (1 + \Gamma_r \|v-w\|_y) \|u\|_w^2$ *for all* $v, w \in B(r)$, *all* $u, y \in H$;

(5) $\|u\|_v^2 \leqq (a_+ + G(w) + a_- \Gamma_r \|v-w\|_y) \|u\|^2$ *for all* $v, w \in B(r)$, *all* $u, y \in H$;

(6) $\|u\|_v^2 \leqq (1 + a_-^{-1} G(w) + \Gamma_r \|v-w\|_y) \|u\|_*^2$ *for all* $v, w \in B(r)$, *all* $u, y \in H$.

Proof. The proof of (2) is easy using the fact that b_v and a are bounded bilinear forms on $H \times H$. The first two inequalities of (3) are also easy (cf. remark after 3.1), the third follows from (2); (1) follows from assumption 3.1.(4) and from (3); (4) is easy using (1) and (3). Now note that using (1), (2) and (3):

$$\forall v, w \in B(r)\ \forall u, y \in H: \ \|u\|_v^2 - \|u\|_*^2 \leqq |\|u\|_v^2 - \|u\|_w^2| + |\|u\|_w^2 - \|u\|_*^2| \leqq (a_- \Gamma_r \|v-w\|_y + G(w)) \|u\|^2. \qquad (*)$$

Now (5) is easy using (*); (6) follows using (*) and (3). □

The following lemma gives a bound for the error made by linearizing n.

Lemma 3.5. *For* $v \in H$ *let* $d_v: H \times H \to H'$ *be defined by* $d_v(u,w) = n(u) - n(w) - b_v(u-w)$. *For all* $r \geqq 0$ *and all* $u, v, w \in B(r)$ *the following holds:*

$$\|d_v(u,w)\|_v \leqq \tfrac{1}{2} \Gamma_r (\|u-v\|_v + \|w-v\|_v) \|u-w\|_v.$$

Proof.

$$\begin{aligned}
\|d_v(u,w)\|_v &= \|n(u) - n(w) - b_v(u-w)\|_v \\
&= \left\| \int_0^1 Dn(w + t(u-w))(u-w) - b_v(u-w)\, dt \right\|_v \\
&\leqq \int_0^1 \left\| Dn(w + t(u-w))(u-w) - b_v(u-w) \right\|_v dt \\
&= \int_0^1 \|b_{w+t(u-w)}(u-w) - b_v(u-w)\|_v \, dt \\
&\leqq \int_0^1 \Gamma_r \|(1-t)(w-v) + t(u-v)\|_v \, dt \, \|u-w\|_v \quad \text{(using 3.4.(1))} \\
&\leqq \tfrac{1}{2} \Gamma_r (\|w-v\|_v + \|u-v\|_v) \|u-w\|_v. \quad \square
\end{aligned}$$

4. A Class of Linear two Grid Operators

In this section we define a class of linear two grid operators and using the convergence theory of [1] we prove uniform convergence for this class of operators.

Definition 4.1. Let l_k: $S_k \times S_k \to \mathbb{R}$ be a continuous, symmetric, positive definite bilinear form. Let $v \in H$, for $w \in S_k$ let $\bar{w}$ be the unique element in S_k such that: $l_k(\bar{w}, y) = l_k(w, y) - b_v(w, y)$ for all $y \in S_k$. We define R_v: $S_k \to S_k$ by $R_v w = \bar{w}$. If l_k, b_v are seen as operators $S_k \to S'_k$ we have $R_v = id_{S_k} - l_k^{-1} b_v$. Let $v \in H$, for $e \in S_k$ let $\bar{e} \in S_{k-1}$ be the element in S_{k-1} such that $b_v(\bar{e}, y) = b_v(e, y)$ for all $y \in S_{k-1}$. We define C_v: $S_k \to S_k$ by $C_v e = e - \bar{e}$. Note that C_v is the b_v-orthogonal projection upon $S_{k-1}^{\perp}$ ($\perp$ w.r.t. b_v).

Remark 4.2. The operator $R_v C_v R_v$: $S_k \to S_k$ can be seen as a so called two grid operator (cf. [1] and [7]). Classical convergence proofs give bounds smaller than one independent of k for $\|R_v C_v R_v\|_v$ for fixed v. Here, we want uniform bounds in v. Theorem 4.10 gives such a result.

Assumption 4.3. For all $k \in \mathbb{N}$, all $r \in]0, \infty[$ let $l_{k,r}$: $S_k \times S_k \to \mathbb{R}$ be defined with the following properties:

(1) $\forall k \forall r$: $l_{k,r}$ is a continuous, symmetric, positive definite bilinear form.
(2) $\exists n \exists c$: $\forall r \forall k \forall v \in B(r)$ $\forall y \in S_k$: $b_v(y, y) \leqq l_{k,r}(y, y) \leqq c(1 + \Gamma_r r)^n h_k^{-2} \|y\|_{L^2}^2$ (with Γ_r as in 3.2).

Notation 4.4. R_v as defined above, with $l_k = l_{k,r}$ is denoted by $R_{v,r}$.

Examples 4.5. Possible choices for $l_{k,r}$ are: Richardson relaxation where $l_{k,r}(\psi_i, \psi_j) = 0$, $i \neq j$, $l_{k,r}(\psi_i, \psi_i) = c(1 + \Gamma_r r)$ with suitable c (depending on b_{w^*}), and "damped Jacobi relaxation" where $l_{k,r}(\psi_i, \psi_j) = 0$, $i \neq j$, $l_{k,r}(\psi_i, \psi_i) = c(1 + \Gamma_r r)$ $(1 + \Gamma_{r_0} r_0)\, b_{v_0}(\psi_i, \psi_i)$ with $v_0 \in B(r_0)$ and suitable c (depending on b_{w^*}).

Lemma 4.6. *Let k be given. With the definitions of $R_{v,r}$ and C_v as in* 4.1, 4.4 *the following holds: $\|R_{v,r}\|_v \leqq 1$ for all $v \in B(r)$, $\|C_v\|_v \leqq 1$ for all $v \in H$.*

Proof. For all $w, y \in S_k$: $l_{k,r}(R_{v,r} w, y) = l_{k,r}(w, y) - b_v(w, y) = l_{k,r}(y, w) - b_v(y, w) = l_{k,r}(R_{v,r} y, w)$ so $R_{v,r}$ is $l_{k,r}$-symmetric and consequently b_v-symmetric. Since $0 \leqq l_{k,r}(R_{v,r} w, w)(l_{k,r}(w, w))^{-1} = 1 - b_v(w, w)(l_{k,r}(w, w))^{-1} \leqq 1$ for all $v \in B(r)$, all $w \in S_k$ (use 4.3) we have that for the spectrum $\sigma(R_{v,r})$ of $R_{v,r}$: $\sigma(R_{v,r}) \subseteq [0, 1]$ and thus $\|R_{v,r}\|_v \leqq 1$. Because C_v is the b_v-orthogonal projection upon $S_{k-1}^{\perp}$ ($\perp$ w.r.t. b_v) we have $\|C_v\|_v \leqq 1$. □

Assumption 4.7. In the remainder of this paper (except §8) we assume that for all $r \geqq 0$ there exists a constant $d_{B(r)}$ such that for all $v \in B(r)$ and for all $m \in L^2(\Omega) \subset H'$ the unique solution u of $b_v u = m$ belongs to $H^2(\Omega)$ and satisfies $\|u\|_2 \leqq d_{B(r)} \|m\|_{L^2}$.

Remark 4.8. In line with Remark 3.3 we note here that our analysis is still applicable and the results do not really change if in 4.7 "for all $r \geqq 0$" is replaced by "for some $r > 0$". Such a local uniform regularity condition we consider to be reasonable because it is a natural generalization for our nonlinear problem of regularity conditions that are common practice in standard convergence theory for linear methods (e.g. [1, 6]).

In §8 we give conditions on g under which assumption 4.7 is fulfilled.

In [1] convergence of a two grid method is determined by the so called "generalized condition number", denoted by κ. Using arguments deduced from [1] we prove existence of (and give a bound for) a uniform κ for a whole class of problems. This uniform κ implies a uniform convergence statement.

Lemma 4.9. *For* $v\in H$ *define* $S_{k-1}^v:=\{u\in S_k \mid b_v(y,u)=0$ *for all* $y\in S_{k-1}\}$. *With* $l_{k,r}$ *as in* 4.3 *the following holds:*

$$\forall r>0\,\exists\,\kappa_{B(r)}: \forall k\,\forall v\in B(r)\,\forall u\in S_{k-1}^v: l_{k,r}(u,u)\leqq \kappa_{B(r)} b_v(u,u).$$

Proof. First we note that, using standard finite element spaces, the following holds ([5] theorem 3.2.1): for all $v\in H^2(\Omega)$ there exists a constant e (independent of k) such that $\min\{\|v-y\| \mid y\in S_k\}\leqq e h_k \|v\|_2$. (*)

Recall that $h_{k-1}\leqq K_1^{-1} h_k$ (2.2). Now take $k\in\mathbb{N}$, $v\in B(r)$ and $u\in S_{k-1}^v$. For $m\in L^2(\Omega)$ let $z\in H^2(\Omega)$ be the solution of $b_v z=m$ (cf. 4.7). Now using lemma 3.4(5), assumption 4.7 and (*) we get, with $x\in S_{k-1}$ suitably chosen and $k_1(r):=a_+ + G(w^*)+a_- \Gamma_r r$ (cf. 3.4):

$$\begin{aligned}(m,u)_{L^2}&=b_v(z,u)=b_v(z-x,u)\leqq \|z-x\|_v \|u\|_v\leqq k_1(r)^{\frac12}\|z-x\|\,\|u\|_v\\ &\leqq k_1(r)^{\frac12} e K_1^{-1} h_k \|z\|_2 \|u\|_v \leqq k_1(r)^{\frac12} e K_1^{-1} h_k d_{B(r)} \|m\|_{L^2}\|u\|_v.\end{aligned}$$

So we have:

$$\|u\|_{L^2}=\max\{(m,u)_{L^2} \mid m\in L^2(\Omega),\ \|m\|_{L^2}=1\}\leqq k_1(r)^{\frac12} e K_1^{-1} h_k d_{B(r)}\|u\|_v.$$

Finally using assumption 4.3 we get:

$$l_{k,r}(u,u)\leqq c(1+\Gamma_r r)^n h_k^{-2}\|u\|_{L^2}^2\leqq c K_1^{-2} e^2 k_1(r)(1+\Gamma_r r)^n d_{B(r)}^2\|u\|_v^2=:\kappa_{B(r)}\|u\|_v^2. \quad \square$$

Now using the convergence proof of Bank and Douglas [1] in combination with lemma 4.9 results in the following uniform convergence statement:

Theorem 4.10. *With* $\kappa_{B(r)}$ *from lemma* 4.9 *and* $\delta_{B(r)}<1$ *defined by:*

$$\delta_{B(r)}=\begin{cases}(1-\kappa_{B(r)}^{-1})^2 & \text{if } \kappa_{B(r)}\geqq 3\\ \kappa_{B(r)}\frac{4}{27} & \text{if } \kappa_{B(r)}<3\end{cases}$$

we have for $R_{v,r}C_vR_{v,r}: S_k\to S_k$ *as defined in* 4.1–4.4*:*

$$\forall r>0: L_{B(r)}:=\sup_{k;\,v\in B(r)}(\|R_{v,r}C_vR_{v,r}\|_v)\leqq\delta_{B(r)}<1.$$

Remark 4.11. The reader may check that $\delta_{B(r)}$ is an increasing function of r. The uniform contraction factor $L_{B(r)}$ need not be an increasing function of r. When we restrict to $r\leqq r_{\max}$ and take $l_{k,r}=l_{k,r_{\max}}$ for all $r\leqq r_{\max}$, then $R_{v,r}=R_{v,r_{\max}}$ for all $r\leqq r_{\max}$ and $L_{B(r)}$ is an increasing function of r for $r\leqq r_{\max}$.

5. Two Grid Full Approximation Scheme Convergence

In this section we prove local convergence of the well-known two grid FAS-iteration (see e.g. [2]). By linearisation and using the convergence theorem of §4 we prove a local uniform convergence statement. The uniformity makes it possible to prove multigrid convergence in the next section. As in the Newton-Kantorovich theorem a domain of guaranteed convergence is determined by a Lipschitz constant.

Given $m \in S_k'$ we want to approximate the solution $v^* \in S_k$ of the Galerkin discretization: $n(v^*, y) = m(y)$ for all $y \in S_k$.

Definition 5.1. For $r > 0$ let $l_{k,r}\colon S_k \times S_k \to \mathbb{R}$ be as in 4.3. For $w \in S_k$ let $\bar{w} \in S_k$ be the unique element in S_k such that:

$$l_{k,r}(\bar{w}, y) = l_{k,r}(w, y) + m(y) - n(w, y) \qquad \text{for all } y \in S_k.$$

We define $R_r\colon S_k \to S_k$ by $R_r(w) = \bar{w}$. With the notation of 3.2 we also have:

$$R_r = id_{S_k} + l_{k,r}^{-1} m - l_{k,r}^{-1} n_k.$$

Choices of $l_{k,r}$ as in 4.5 correspond to Richardson relaxation and nonlinear damped Jacobi with a modified Newton iteration.

For every $\tilde{u} \in S_{k-1}$, $s \in]0, \infty[$, $m \in S_k'$ we define $C\colon S_k \to S_k$ by $C(e) = e + \frac{1}{s}(\bar{e} - \tilde{u})$ where $\bar{e} \in S_{k-1}$ is the unique element in S_{k-1} such that:

$$n(\bar{e}, x) = n(\tilde{u}, x) + s(m(x) - n(e, x)) \qquad \text{for all } x \in S_{k-1}$$

(for existence of such an $\bar{e}$ see 2.3). With the notation of 3.2 we also have:

$$C(e) = e + \frac{1}{s}\left[n_{k-1}^{-1}(n_{k-1}(\tilde{u}) + s(m - n_{k,k-1}(e))) - \tilde{u}\right].$$

We define $F_k\colon]0, \infty[\times S_k \times S_{k-1} \times]0, \infty[\times S_k' \to S_k$ by: $F_k(r, w, \tilde{u}, s, m) = R_r C R_r(w)$ with C as above (corresponding to $\tilde{u}$, s and m). For given $r^{(j)}, \tilde{u}^{(j)}, s^{(j)}$ and m a two grid Full Approximation Scheme iteration for solving $n_k(u) = m$ is defined by:

$$w^{(j+1)} = F_k(r^{(j)}, w^{(j)}, \tilde{u}^{(j)}, s^{(j)}, m).$$

We will indicate suitable choices of $r^{(j)}$, $\tilde{u}^{(j)}$ and $s^{(j)}$ later on.

Remark 5.2. The above definition in variational formulation of the FAS two grid iteration is equivalent to the more familiar and practical definition in matrix formulation (cf. (3), (4) in 2.2) as given in e.g. [2] and [6].

Clearly for convergence of the FAS iteration we are interested in the ratio of $|||F_k(r, w, \tilde{u}, s, m) - v^*|||$ and $|||w - v^*|||$ for a suitably chosen norm $|||.|||$ on S_k. Theorem 5.10 gives a bound for this ratio. We first give some lemmas, trying to make the technical proof of Theorem 5.10 more transparant.

Lemma 5.3. *The following holds (for definition of $G(w^*)$ see 3.4):*

(a) For all $w \in H$ and all $v \in B(r)$:

$$\|b_w^{-1} b_v\|_v \leq \hat{\Gamma}_r \quad \textit{with} \quad \hat{\Gamma}_r := 1 + a_-^{-1} G(w^*) + \Gamma_r r.$$

(b) For all $v, w \in B(r)$:

$$\|b_w^{-1} b_v\|_v \leqq 1 + 2\Gamma_r r$$

(c) For all $v, w \in B(r)$, *for all* $y \in H$:

$$\|b_w^{-1} b_v - id_H\|_v \leqq \exp(\Gamma_r \|v-w\|_y) - 1 \leqq \exp(2\Gamma_r r) - 1$$

Proof. Notice that for $v \in B(r)$, $w, z \in H$: $b_v(z,z) \leqq (1+\gamma)\, b_w(z,z)$ with $\gamma = a_-^{-1} G(w^*) + \Gamma_r r$ (follows from 3.4 (6) and (3)); we may take $\gamma = 2\Gamma_r r$ if $w \in B(r)$ (follows from 3.4 (4)). Now note that $b_w^{-1} b_v u = z$ is equivalent to: $b_v(u, y) = b_w(z, y)$ for all $y \in H$, and thus

$$\|z\|_v^2 = b_v(z,z) \leqq (1+\gamma)\, b_w(z,z) = (1+\gamma)\, b_v(u,z) \leqq (1+\gamma)\, \|u\|_v \|z\|_v.$$

This results in $\|z\|_v \leqq (1+\gamma)\|u\|_v$ and thus

$$\|b_w^{-1} b_v\|_v = \max\{\|b_w^{-1} b_v u\|_v \mid u \in H,\ \|u\|_v = 1\} \leqq 1+\gamma.$$

This proves (a) and (b). For (c) note that for $v, w \in B(r)$, $u, z \in H$ with $b_w^{-1} b_v u = z$ we have:

$$\|z-u\|_v^2 = b_v(z-u, z-u) = b_v(z, z-u) - b_w(z, z-u) \leqq \Gamma_r \|v-w\|_y \|z\|_w \|z-u\|_v$$

(using 3.4 (1) with $y_1 = y \in H$, $y_2 = w$, $y_3 = v$). This implies:

$$\|z-u\|_v \leqq \Gamma_r \|v-w\|_y \|z\|_w \leqq \Gamma_r \|v-w\|_y (1 + \Gamma_r \|v-w\|_y)^{\frac{1}{2}} \|u\|_v;$$

in the last inequality we used:

$$b_w(z,z) = b_v(u,z) \leqq \|u\|_v \|z\|_v \leqq \|u\|_v (1 + \Gamma_r \|v-w\|_y)^{\frac{1}{2}} \|z\|_w$$

(use 3.4 (4)). Conclusion:

$$\begin{aligned}\|b_w^{-1} b_v u - u\|_v &\leqq \Gamma_r \|v-w\|_y (1+\Gamma_r\|v-w\|_y)^{\frac{1}{2}} \|u\|_v \leqq (\exp(\Gamma_r \|v-w\|_y) - 1)\|u\|_v \\ &\leqq (\exp(2\Gamma_r r) - 1)\|u\|_v \qquad (\text{take } y = w^*). \quad \square\end{aligned}$$

The following lemma will be used further on to linearize n^{-1}. The nice looking property 5.4 (a) will be used several times; the more general statement in 5.4 (b) will be used only once.

Lemma 5.4. *For* $t \in [0,1]$ *we define* ω_t: $H \times H' \to H$ *by* $\omega_t(u, \varphi) = n^{-1}(n(u) + t\varphi)$. *The following holds:*

(a) $\forall r > 0\ \forall u \in H\ \forall \varphi \in H'\ \forall v \in B(r)$: $\|\omega_1(u,\varphi) - u\|_v \leqq \hat{\Gamma}_r \|\varphi\|_v$ ($\hat{\Gamma}_r$ as in 5.3).

(b) We define $\tilde{r}$: $H \times H' \to \mathbb{R}$ *by*

$\tilde{r}(u,\varphi) = \inf\{s \mid \omega_t(u,\varphi) \in B(s)$ *for all* $t \in [0,1]\}$

then: $\forall r > 0\ \forall u \in H\ \forall \varphi, \psi \in H'\ \forall v \in B(r)$: $\|\omega_1(u,\varphi) - u - b_v^{-1}\psi\|_v$

$\leqq \{\exp[\Gamma_{\max(r, \tilde{r}(u,\varphi))}(\|v-u\|_v + \hat{\Gamma}_r \|\varphi\|_v)] - 1\}\, \|\psi\|_v + \hat{\Gamma}_r \|\varphi - \psi\|_v.$

Proof. The reader may check that n: $H \to H'$ is a C^1-diffeomorphism (as in 2.3 for N_k). Now take $u \in H$, $\varphi, \psi \in H'$, $v \in B(r)$ and for $t \in [0,1]$: $z(t) := n^{-1}(n(u)$

$+t\varphi)-u-b_v^{-1}\psi$ so $n(z(t)+u+b_v^{-1}\psi)=n(u)+t\varphi$ and $z(0)=-b_v^{-1}\psi$. By differentiating w.r.t. t we get:

$$Dn(z(t)+u+b_v^{-1}\psi)(z'(t))=\varphi \Leftrightarrow b_{\omega_t(u,\varphi)}z'(t)=\varphi \Leftrightarrow z'(t)=b^{-1}_{\omega_t(u,\varphi)}\varphi$$

implying:

$$z(1)=\int_0^1 z'(t)\,dt+z(0)=\int_0^1 b^{-1}_{\omega_t(u,\varphi)}\varphi\,dt-b_v^{-1}\psi$$
$$=\int_0^1 (b^{-1}_{\omega_t(u,\varphi)}b_v-id)\,b_v^{-1}\psi\,dt+\int_0^1 b^{-1}_{\omega_t(u,\varphi)}b_v b_v^{-1}(\varphi-\psi)\,dt$$

Taking norms, using the continuity in t and $\|b_v^{-1}\|_v=1$ results in:

$$\|z(1)\|_v \leqq \int_0^1 \|b^{-1}_{\omega_t(u,\varphi)}b_v-id\|_v dt\,\|\psi\|_v+\int_0^1 \|b^{-1}_{\omega_t(u,\varphi)}b_v\|_v dt\,\|\varphi-\psi\|_v.$$

Taking $\psi=0$ and using 5.3 (a) proves (a).

For the term $\|b^{-1}_{\omega_t(u,\varphi)}b_v-id\|_v$ we note that using 5.3(c) we get:

$$\begin{aligned}\|b^{-1}_{\omega_t(u,\varphi)}b_v-id\|_v &\leqq \exp(\Gamma_{\max(r,\tilde r(u,\varphi))}\|v-\omega_t(u,\varphi)\|_v)-1\\ &\leqq \exp(\Gamma_{\max(r,\tilde r(u,\varphi))}(\|v-u\|_v+\|\omega_t(u,\varphi)-u\|_v))-1\\ &\leqq \exp(\Gamma_{\max(r,\tilde r(u,\varphi))}(\|v-u\|_v+\hat{\Gamma}_r t\,\|\varphi\|_v))-1,\end{aligned}$$

where in the last inequality we used (a). Now (b) follows. □

Remark 5.5. By using 5.3 (b) instead of 5.3 (a) in the proof of 5.4 it may be seen that $\hat{\Gamma}_r$ in 5.4 may be replaced by $1+2\Gamma_{\max(r,\tilde r(u,\varphi))}\cdot\max(r,\tilde r(u,\varphi))$.

Suppose in 5.4 we are interested in the following situation:

for every $r>0$: $u, v\in B(r)$; for every $r>0$: $\varphi\in H'$ with $\|\varphi\|_{w*}\leqq f(r)$

where $f(r)$ is a continuous increasing function of r with $f(0)=0$. Then $\tilde r(u,\varphi)\leqq f(r)\,\hat{\Gamma}_r+r$ and thus $\hat{\Gamma}_r$ in (a) and (b) can be replaced by the factor $1+2\Gamma_{f(r)\hat{\Gamma}_r+r}\cdot(f(r)\,\hat{\Gamma}_r+r)$, which goes to 1 if $r\downarrow 0$ (note that $\hat{\Gamma}_r \downarrow X$ if $r\downarrow 0$).

Remark 5.6. In lemma 3.5, 5.3, 5.4 the results do not change if H is replaced by some closed linear subspace $\tilde H$ of H and n, b_v are seen as operators $\tilde H\to\tilde H'$.

In the remainder of this section we assume some fixed $S>0$.

Definition 5.7. We now define Lipschitz constants on different balls. For $r>0$, $r^{(0)}\in[r,\infty[$ is such that for all $v\in B(r)$: $B(v;r)\subseteq B(r^{(0)})$. For $r, r^{(0)}>0$, $r^{(1)}\in[r^{(0)},\infty[$ is such that for all k, all $v\in B^k(r)$ the relaxation operator R_r: $S_k\to S_k$ as defined in 5.1 with $m=n_k(v)$ satisfies $R_r(B^k(v;r))\subseteq B(r^{(1)})$. For r, $r^{(0)}$, $r^{(1)}>0$, $r^{(2)}\in[r^{(1)},\infty[$ is such that for all k, all $v\in B^k(r)$, all $\tilde u\in B^{k-1}(r)$, all $s\leqq S$ the operators R_r and C as defined in 5.1 with $m=n_k(v)$ satisfy: $CR_r(B^k(v;r))\subseteq B(r^{(2)})$.

Also define $r^{(3)}:=r+S(r^{(1)}+r^{(2)})$.

The Lipschitz constants $\Gamma_{r^{(i)}}$ are denoted by Γ_i ($i=0,\dots,3$). Finally we define $\gamma_r:=\max_{i=0,1,2,3}\Gamma_i=\max(\Gamma_2,\Gamma_3)$ (Lipschitz constant on a large enough ball).

Remark 5.8. Actual bounds for $r^{(0)}, r^{(1)}, r^{(2)}$ in terms of r and the Lipschitz constant Γ_r can be deduced from lemma 3.4 and proposition 5.9. It turns out that $r^{(i)}$ $i=0,1,2$ can be chosen such that they are continuous increasing functions of r with $\lim_{r\downarrow 0} \frac{r^{(i)}}{r}=2$. Also note that for S small enough $\gamma_r=\Gamma_2$.

Proposition 5.9 gives bounds for higher order terms that occur when we linearize the FAS two grid iteration (cf. 5.10).

Proposition 5.9. For $r>0$ and $v\in S_k$ we define $\Delta R_{v,r}$: $S_k \to S_k$ by

$$\Delta R_{v,r}(u)=R_r(v)-R_r(u)-R_{v,r}(v-u)$$

where R_r is defined as in 5.1 with $m=n_k(v)$ and $R_{v,r}$ as in 4.4.

For $v\in S_k$ define ΔC_v: $S_k \to S_k$ by

$$\Delta C_v(u)=C(v)-C(u)-C_v(v-u)$$

with C as defined in 5.1 with $m=n_k(v)$ and C_v as in 4.1.

Then for all $v\in B^k(r)$, all $\tilde{u}\in B^{k-1}(r)$, all $u\in B^k(v;r)$ and all $s\leqq S$ we have, with R_r and C as in 5.1 with $m=n_k(v)$:

(1) $\|\Delta R_{v,r}(u)\|_v \leqq \frac{1}{2}\Gamma_0 \|v-u\|_v^2 \leqq \frac{1}{2}\gamma_r \|v-u\|_v^2$.

(2) $\|\Delta C_v(R_r(u))\|_v \leqq A_r(\gamma_r \|v-u\|_v, \gamma_r \|v-\tilde{u}\|_v)\|v-u\|_v$ where A_r: $[0,\infty[\times[0,\infty[\to [0,\infty[$ is given by $A_r(x,y)=(\exp(y+Sz_r(x))-1)(1+\frac{1}{2}x)+\frac{1}{2}z_r(x)$ with $z_r(x)=\hat{\Gamma}_r x e^x$ and $\hat{\Gamma}_r$ from 5.3.

(3) $\|\Delta R_{v,r}(CR_r(u))\|_v \leqq \tilde{A}_r(\gamma_r \|v-u\|_v, \gamma_r \|v-\tilde{u}\|_v)\|v-u\|_v$ where $\tilde{A}_r$: $[0,\infty[\times[0,\infty[\to[0,\infty[$ is given by $\tilde{A}_r(x,y)=(1+\frac{1}{2}x+A_r(x,y))^2\,\frac{1}{2}x$.

Proof. Take $k\in\mathbb{N}$, $r>0$, $v\in B^k(r)$, $\tilde{u}\in B^{k-1}(r)$, $u\in B^k(v;r)$, $s\leqq S$ and let R_r and C be as in 5.1 with $m=n_k(v)$.

Proof of (1): let $\mathsf{E}:=\{y\in S_k \mid l_{k,r}(y,y)=1\}$.

Now

$$\begin{aligned}\|\Delta R_{v,r}(u)\|_v &= b_v(\Delta R_{v,r}(u), \Delta R_{v,r}(u))^{\frac{1}{2}} \leqq l_{k,r}(\Delta R_{v,r}(u), \Delta R_{v,r}(u))^{\frac{1}{2}}\\ &=\max_{y\in\mathsf{E}}|l_{k,r}(R_r(v)-R_r(u)-R_{v,r}(v-u),y)|\\ &=\max_{y\in\mathsf{E}}|b_v(v-u,y)-n(v,y)-n(u,y)|=\max_{y\in\mathsf{E}}|d_v(v,u)(y)|\\ &\leqq \max_{\|y\|_v=1}|d_v(v,u)(y)|=\|d_v(v,u)\|_v \leqq \tfrac{1}{2}\Gamma_0\|v-u\|_v^2 \quad (\text{use } 3.5) \leqq \tfrac{1}{2}\gamma_r\|v-u\|_v^2.\end{aligned}$$

This proves (1).

Proof of (2): let

$$\varphi:=s(n_{k,k-1}(v)-n_{k,k-1}(R_r(u))) \qquad \psi:=s(b_v)_{k,k-1}(v-R_r(u))$$

and for $0\leqq t\leqq 1$:

$$\omega_t:=n_{k-1}^{-1}(n_{k-1}(\tilde{u})+t\varphi).$$

The definition of C with st instead of s implies:

$$C(R_r(u))=R_r(u)+\frac{1}{st}\{\omega_t-\tilde{u}\} \qquad (t\neq 0)$$

and so for all $t\in]0,1]$ we have:

$$\|\omega_t - w^*\|_{w^*} \leqq S\cdot 1\cdot \|C(R_r(u)) - R_r(u)\|_{w^*} + \|\tilde{u} - w^*\|_{w^*}$$
$$\leqq S(r^{(1)} + r^{(2)}) + r = r^{(3)} \quad \text{(cf. 5.7)}.$$

Easy writing out using the definitions results in:

$$\Delta C_v(R_r(u)) = -\frac{1}{s}(\omega_1 - \tilde{u} - (b_v)_{k-1}^{-1}\psi),$$

now using lemma 5.4 (b) (with S_{k-1} instead of H) results in:

$$\|\Delta C_v(R_r(u))\|_v \leqq \frac{1}{s}\{\exp[\Gamma_3(\|v-\tilde{u}\|_v + \hat{\Gamma}_r\|\varphi\|_v)] - 1\}\|\psi\|_v + \frac{1}{s}\hat{\Gamma}_r\|\varphi - \psi\|_v. \quad (*)$$

Now note that:

$$\|\varphi\|_v \leqq S(\|v - R_r(u)\|_v + \|d_v(v, R_r(u))\|_v)$$
(definition of d_v, see 3.5; use $\|(b_v)_{k,k-1}\|_v \leqq 1$)
$$\|\psi\|_v \leqq s\|v - R_v(u)\|_v, \qquad \|\varphi - \psi\|_v \leqq s\|d_v(v, R_r(u))\|_v$$
$$\|d_v(v, R_r(u))\|_v \leqq \tfrac{1}{2}\Gamma_1\|v - R_r(u)\|_v^2 \quad \text{(see 3.5)}$$
$$\|v - R_r(u)\|_v \leqq \|v - u\|_v + \tfrac{1}{2}\Gamma_0\|v - u\|_v^2 \quad \text{(using (1) and } \|R_{v,r}\|_v \leqq 1\text{, see 4.6)}.$$

Using these estimates in (*) and tedious writing out results in:

$$\|\Delta C_v(R_r(u))\|_v \leqq A_r(\gamma_r\|u - v\|_v, \gamma_r\|\tilde{u} - v\|_v)\|u - v\|_v \quad \text{with}$$
$$A_r(x, y) = (\exp(y + S z_r(x)) - 1)(1 + \tfrac{1}{2}x) + \tfrac{1}{2}z_r(x) \qquad \text{where } z_r(x) = \hat{\Gamma}_r x e^x.$$

This proves (2).

Proof of (3): as in (1) one can prove:

$$\|\Delta R_{v,r}(CR_r(u))\|_v \leqq \tfrac{1}{2}\Gamma_2\|CR_r(u) - v\|_v^2. \quad (**)$$

Also:

$$\|CR_r(u) - v\|_v = \|CR_r(u) - CR_r(v)\|_v \leqq \|C_v\|_v\|R_r(u) - R_r(v)\|_v + \|\Delta C_v(R_r(u))\|_v$$
$$\leqq \|u - v\|_v + \tfrac{1}{2}\Gamma_0\|v - u\|_v^2 + A_r(\gamma_r\|u - v\|_v, \gamma_r\|\tilde{u} - v\|_v)\|u - v\|_v$$

(using 4.6 and (1) and (2) above)

$$\leqq (1 + \tfrac{1}{2}\gamma_r\|v - u\|_v + A_r(\gamma_r\|u - v\|_v, \gamma_r\|\tilde{u} - v\|_v))\|u - v\|_v.$$

Combining this with (**) leads to the result. □

Now we are able to prove the theorem about uniform FAS two grid convergence for a whole class of problems that we announced in the beginning of this section. We will also use this theorem in the next section to prove FAS MG convergence. For comments concerning this theorem see 5.13, 5.14.

Theorem 5.10. FAS two grid convergence. *(For definitions see 5.1 and 5.7). Let $r>0$ be given; suppose k_0 such that $B^{k_0}(w^*; r) \neq \phi$. Then for all $k>k_0$, all $m\in S'_k$ with $v^* := n_k^{-1}(m) \in B^k(w^*; r)$, all $\tilde{u}\in B^{k-1}(w^*; r)$, all $w\in B^k(v^*; r)$ and all $s\leqq S$ the following holds:*

$$\|F_k(r, w, \tilde{u}, s, m) - v^*\|_{v^*} \leqq [L_{B(r)} + O_{B(r)}(\gamma_r \|w - v^*\|_{v^*}, \gamma_r \|\tilde{u} - v^*\|_{v^*})] \|w - v^*\|_{v^*}$$

with $L_{B(r)} < 1$ *as in* 4.10 *and:*

$$O_{B(r)}(x, y) := \tfrac{1}{2}x + A_r(x, y) + (1 + \tfrac{1}{2}x + A_r(x, y))^2 \tfrac{1}{2}x,$$

where A_r *is as in* 5.9(2), *i.e.:*

$$A_r(x, y) = (\exp(y + S z_r(x)) - 1)(1 + \tfrac{1}{2}x) + \tfrac{1}{2}z_r(x), \quad \textit{with} \ z_r(x) = \hat{\Gamma}_r x e^x.$$

Proof. Take $\tilde{u} \in B^{k-1}(r)$, $s \leqq S$, $m \in S'_k$ with $v^* := n_k^{-1}(m) \in B^k(w^*; r)$ and $w \in B^k(v^*; r)$. Simple writing out using the definitions of 5.1 and 5.9 results in:

$$\begin{aligned} \|F_k(r, w, \tilde{u}, s, m) - v^*\|_{v^*} &= \|R_r C R_r(w) - v^*\|_{v^*} \leqq \|R_{v^*,r} C_{v^*} R_{v^*,r}\|_{v^*} \|w - v^*\|_{v^*} \\ &+ \|R_{v^*,r}\|_{v^*} \|C_{v^*}\|_{v^*} \|\Delta R_{v^*,r}(w)\|_{v^*} + \|R_{v^*,r}\|_{v^*} \|\Delta C_{v^*}(R_r(w))\|_{v^*} \\ &+ \|\Delta R_{v^*,r}(C R_r(w))\|_{v^*}. \end{aligned}$$

Now using Lemma 4.6, Theorem 4.10 and Proposition 5.9, the proof is completed. □

In the first order terms of the function $A_r(x, y)$ (see theorem) the factor $\hat{\Gamma}_r$ occurs. We can replace $\hat{\Gamma}_r$ by a better factor if we make a suitable assumption about S. The following lemma makes this clear.

Lemma 5.11. *The factor* $\hat{\Gamma}_r$ *in* A_r *(see* 5.9(2) *and* 5.10*) can be replaced by:*
(a) $1 + 2\Gamma_{c(r)r} c(r)\, r$ with $c(r) := \hat{\Gamma}_r S \exp(1\tfrac{1}{2}\Gamma_1 r) + 1$.
This new factor has the desirable property that it tends to 1 if $r \downarrow 0$. However to analyze its behaviour for $\gamma_r r$ small is much more complicated than it is for $\hat{\Gamma}_r$. *For some* $\tilde{r} > 0$ *assume* $S = S_1 \hat{\Gamma}_{\tilde{r}}^{-1} \exp(-1\tfrac{1}{2}\Gamma_1 \tilde{r})$ *with* $S_1 \leqq 1$, *then for all* $r \leqq \tilde{r} \hat{\Gamma}_r$ *in* A_r *can be replaced by:*
(b) $1 + 2\Gamma_{(S_1+1)r}(S_1 + 1)\, r \leqq 1 + 2(S_1 + 1)\, \gamma_r r$.

Proof. Inspection of the proof of 5.9(2) shows that the factor $\hat{\Gamma}_r$ is caused by the fact that we use 5.4 where the factor $\hat{\Gamma}_r$ occurs. We now use the variant of 5.4 as mentioned in 5.5. We use the notation as in the proof of 5.9(2).

$$\|\omega_t - w^*\|_{w^*} \leqq \|\omega_t - \tilde{u}\|_{w^*} + r \leqq (1 + \Gamma_r r)^{\frac{1}{2}} \hat{\Gamma}_r t \|\varphi\|_v + r$$

(using lemma 3.4(4) and 5.4(a)).

Using estimates as in the proof of Lemma 5.9(2) one can easily show:

$$\|\varphi\|_v \leqq S \exp[\Gamma_1 \|v - u\|_v] \|v - u\|_v;$$

using this and $t \in [0, 1]$, $u \in B(v; r)$ we get:

$$\begin{aligned} \|\omega_t - w^*\|_{w^*} &\leqq (1 + \Gamma_r r)^{\frac{1}{2}} \hat{\Gamma}_r S \exp(\Gamma_1 r)\, r + r \leqq \exp(\tfrac{1}{2}\Gamma_1 r)\, \hat{\Gamma}_r S \exp(\Gamma_1 r)\, r + r \\ &= (\hat{\Gamma}_r S \exp(1\tfrac{1}{2}\Gamma_1 r) + 1)\, r. \end{aligned}$$

Now using 5.5 results in (a). The second statement is easy (use $\Gamma_{2r} \leqq \gamma_r$). □

Remark 5.12. In the next section we will again use the following estimate, that occurs in the proof of 5.11 (also in the proof of 5.9(2)):

$$\forall k\, \forall v \in B^k(w^*; r)\, \forall u \in B^k(v; r): \|n_{k,k-1}(v) - n_{k,k-1}(R_r(u))\|_v$$
$$\leqq \exp[\Gamma_1 \|v-u\|_v] \|v-u\|_v \leqq \exp[\Gamma_1 r]\, r,$$

where R_r as in 5.1 with $m = n_k(v)$.

Remark 5.13. If we want to prove two grid convergence only, we do not need the uniformity in m in 5.10. Suppose for $f \in S'_k$ we use a FAS two grid iteration for the problem $n_k(u) = f$ (solution denoted by u_k^*). Taking $w^* = v^* = u_k^*$ in theorem 5.10 shows that convergence will occur if we take w (starting vector) within the ball $B^k(u_k^*; r)$ and $\tilde{u} \in B^{k-1}(u_k^*; r)$ with r such that $O_{B(r)}(\gamma_r r, \gamma_r r) < 1 - L_{B(r)}$, and k large enough such that $B^{k-1}(u_k^*; r) \neq \phi$.

With respect to $O_{B(r)}(\gamma_r r, \gamma_r r)$ we note the following:

$$O_{B(r)}(\gamma_r r, \gamma_r r) \approx 2\tfrac{1}{2} \gamma_r r + \text{h.o.}(\gamma_r r)$$

for all r with $\gamma_r r \leqq \frac{2}{5}$ and S small enough (see below). Here h.o. denotes higher order terms, and as an example: h.o.(0.3)$\approx$1.40, h.o.(0.2)$\approx$0.25, h.o.(0.1)$\approx$0.06. To make this clear we note the following. If we take $S=0$ in the definition of $O_{B(r)}$ (so $\hat{\Gamma}_r$ in A_r may be replaced by $1 + 2\gamma_r r$; see 5.11) and neglect higher order terms $O_{B(r)}(\gamma_r r, \gamma_r r)$ results in $2\frac{1}{2}\gamma_r r$. So in order to get $O_{B(r)}(\gamma_r r, \gamma_r r) < 1$ we restrict to $r \in]0, r_0]$ with r_0 such that $\gamma_{r_0} r_0 = \frac{2}{5}$. Now assume S small enough such that $S\hat{\Gamma}_{r_0} \ll 1$, then $1 + S\hat{\Gamma}_r \exp(\gamma_r r) \approx 1$ $(r \leqq r_0)$ and for $r \leqq r_0$ $\hat{\Gamma}_r$ in A_r may be replaced by $\approx 1 + 2\gamma_r r$ (see 5.11). Using this in the definition of $O_{B(r)}$ we get the above result.

Remark 5.14. In order to assess the sharpness of the theorem we consider two extreme situations. Clearly, if we take $\gamma_r = 0$ (linear problem) the theorem results in the expected convergence statement. Another extreme situation is the following: in the two grid situation of 5.13 we take the $k-1$ and k level identical, resulting in $C_{u_k^*} = 0$, and we also take $S \downarrow 0$. Now the "two" grid FAS iteration is a modified Newton iteration proceded and followed by a relaxation iteration that is not locally convergent ($\|R_r(u) - u_k^*\|_{u_k^*} \leqq \|u - u_k^*\|_{u_k^*} + \frac{1}{2}\Gamma_r \|u - u_k^*\|_{u_k^*}^2$). Let A_r^0 be the function A_r of Theorem 5.10 with $S=0$. Inspection of the proof of the theorem and of 5.9(3) shows that because $C_{u_k^*} = 0$ we can get a somewhat better (but analogous) result than in theorem 5.10: convergence in $B^k(u_k^*; r)$ is guaranteed if r is such that $A_r^0(\gamma_r r, \gamma_r r) + (A_r^0(\gamma_r r, \gamma_r r))^2 \frac{1}{2}\gamma_r r < 1$ (use $L_{B(r)} = 0$). We have that $A_r^0(\gamma_r r, \gamma_r r) + (A_r^0(\gamma_r r, \gamma_r r))^2 \frac{1}{2}\gamma_r r = 1\frac{1}{2}\gamma_r r + \text{h.o.}(\gamma_r r)$ with (as an example) h.o.(0.4)$\approx$0.78, h.o.(0.3)$\approx$0.35, h.o.(0.2)$\approx$0.13. The above kind of condition for a domain of guaranteed convergence is well-known for Newton iterations. Moreover, if we also assume $\tilde{u} = w$, resulting in a non-modified Newton iteration, we see that as a bound for the convergence factor we get (with $e_k := \|w - u_k^*\|_{u_k^*}$):

$$A_r^0(\gamma_r e_k, \gamma_r e_k) + (A_r^0(\gamma_r e_k, \gamma_r e_k))^2 \tfrac{1}{2}\gamma_r e_k = 1\tfrac{1}{2}\gamma_r e_k + \text{h.o.}(\gamma_r e_k).$$

This exhibits the quadratic convergence of the Newton iteration.

Remark 5.15. Finally, we reflect on the dispensability of the Lipschitz condition that we assumed in 3.1.(4). We first recall that this Lipschitz condition could be weakened to a local Hölder condition (see 3.3).

In this section our goal is to prove a quantitative convergence statement for a FAS TG scheme for solving $n_k(u_k)-f=0$ which is uniform in k (that is, we wish to obtain explicit expressions for a domain of guaranteed convergence and for a bound on the contraction factor which are independent of k; cf. Theorem 5.10). In Remark 5.14 it is noticed that for an extreme choice of parameters ($h_{k-1}\to h_k$, $S\downarrow 0$) the FAS TG scheme reduces to a Newton kind of algorithm. Thus our conditions should certainly warrant a quantitative convergence statement that is uniform for this set of Newton algorithms. Now for a quantitative convergence analysis of Newton iterations a Lipschitz condition (a local Hölder condition) on the derivative is a common requirement (cf. [9]), and it is, indeed, not clear how one could do without. This leads then to Lipschitz conditions on the Dn_k uniform in k. Finally, recall that n_k is just $n_{|S_k}$: $S_k\to S'_k$ and that $\overline{\bigcup_k S_k}=H$. Hence, in our opinion, the Lipschitz condition on Dn in 3.1 is rather a minimal condition.

6. Multigrid Full Approximation Scheme

In this section we define a multigrid FAS iteration with a bounded number of coarse grid corrections on each level and prove convergence with contraction number smaller than one, independent of the level, provided we start in a ball with radius small enough (independent of the level) and provided the coarsest level is fine enough.

Assume $p\in\mathbb{N}$ and $f\in L^2(\Omega)\subset S'_p$. We want to approximate the solution u_p^* of the Galerkin discretization $n_p(u)=f$, using FAS multigrid on level p.

In 6.1 we define a FAS multigrid iteration on level p and in 6.5 and 6.6 we will prove a convergence statement.

Definition 6.1. If $\tilde{F}_k$: $S_k\times S'_k\to S_k$ is defined, then for $n\in\mathbb{N}$, $n\geqq 2$ we define $\tilde{F}_k^n$: $S_k\times S'_k\to S_k$ by $\tilde{F}_k^n(u,m)=\tilde{F}_k(\tilde{F}_k^{n-1}(u,m),m)$.

We will now define a FAS multigrid iteration on level $p\in\mathbb{N}$.

For $k=k_0,\ldots,p$ let there be given $r_k>0$, $\tilde{u}_{k-1}\in S_{k-1}$, σ_k: $S_k\times S_k\to]0,\infty[$ and $\tau_k\in\mathbb{N}$.

We define $\tilde{F}_k$: $S_k\times S'_k\to S_k$ by the following (cf. 5.1):

(1) $k=k_0$: $\tilde{F}_{k_0}(w,m)=F_{k_0}(r_{k_0},w,\tilde{u}_{k_0-1},\sigma_{k_0}(v^*,w),m)$ with F_{k_0} from 5.1 and $v^*:=n_{k_0}^{-1}(m)$.

(2) $k_0<k\leqq p$: choose $w\in S_k$, $m\in S'_k$ and let $v^*:=n_k^{-1}(m)$; define R_{r_k}: $S_k\to S_k$ as in 5.1. Define

$$\tilde{C}: S_k\to S_k \quad \text{by} \quad \tilde{C}(e)=e+\sigma_k(v^*,w)^{-1}\{\hat{e}-\tilde{u}_{k-1}\}$$

with

$$\hat{e}=\tilde{F}_{k-1}^{\tau_k}(\tilde{u}_{k-1},n_{k-1}(\tilde{u}_{k-1})+\sigma_k(v^*,w)\{m-n_{k,k-1}(e)\}).$$

Now $\tilde{F}_k(w,m):=R_{r_k}\tilde{C}R_{r_k}(w)$.

A FAS multigrid iteration on level p for solving $n_p(u)=f$ is defined by:

$$w^{(j+1)}=\tilde{F}_p(w^{(j)},f).$$

In 6.2–6.5 we assume that there is some fixed $w^* \in C(\bar{\Omega}) \cap H$. The reader may think of the following two possibilities for w^* (cf. 6.6): $w^* = u^*$ (u^* the solution in H of $n(u)=f$), $w^* = u_p^*$ (u_p^* the solution of the Galerkin discretization $n_p(u)=f$). Also in the remainder of this section we assume some fixed $S>0$.
Given w^* as above we define $k_-:]0, \infty[\to \mathbb{N}$ by $k_-(r) = \min\{n \mid B^n(w^*; r) \neq \phi\}$.

Conditions 6.2. We state the following conditions, which we will use in 6.5. We will discuss the feasibility of these conditions in 6.3.

(A) ("choice of r") Define (cf. 5.1, 5.10):

$$N_{B(r)} = N_{B(w^*; r)} := \sup\{\|F_k(r, w, \tilde{u}, s, m) - v^*\|_{v^*} \|w - v^*\|_{v^*}^{-1} \mid k > k_-(r), \quad m \in S'_k$$
$$\text{with } v^* := n_k^{-1}(m) \in B^k(r), \tilde{u} \in B^{k-1}(r), w \in B^k(v^*; r) \text{ with } w \neq v^*, s \leq S\}.$$

Let $N<1$ and $r>0$ be such that $N_{B(t)} \leq N$ for all $t \in]0, r]$.

(B) ("choice of σ_k"). For $\eta \in]0, 1[$, $t>0$ let $\sigma_k: S_k \times S_k \to]0, S]$ satisfy:

$$\|n_{k-1}^{-1}\{n_{k-1}(\tilde{u}_{k-1}) + \sigma_k(v, u)(n_{k,k-1}(v) - n_{k,k-1}(R_t(u)))\} - \tilde{u}_{k-1}\|_{w^*}$$
$$\leq \eta t (1 + \Gamma_t t)^{-\frac{1}{2}}$$

for all $\tilde{u}_{k-1} \in S_{k-1}$; here R_t is as defined in 5.1 with $m = n_k(v)$.

(C) ("choice of τ_k"). Suppose we have sequences $(a_n)_{n \geq 0}$, $(b_n)_{n \geq 0}$ in $\mathbb{R}$ such that:

(1) $\exists A>0$: $A \leq a_n \leq a_0 < 1$ for all n, and $\lim_{n \to \infty} a_n = A$.

(2) $\exists B \geq 1$: $B \leq b_n \leq b_0$ for all n, and $\lim_{n \to \infty} b_n = B$.

Now let $(\tau_k)_{k \geq 1}$ be a sequence in $\mathbb{N}$ such that with $\varphi_n(x) := a_n + b_n x^{\tau_n}$ ($x \in [0, 1]$) the sequence $x_1 := a_1$, $x_{n+1} := \varphi_{n+1}(x_n)$ ($n \geq 1$) satisfies: $\sup_n (x_n) < 1$.

Remarks 6.3.

(A 1) Theorem 5.10 implies that $N_{B(r)} \leq L_{B(r)} + O_{B(r)}(\gamma_r r, \gamma_r r(1+\gamma_r r)^{\frac{1}{2}})$. From 4.10 we deduce that $L_{B(r)} \leq \delta_{B(r)} < 1$ for all $r>0$ and $\delta_{B(r)}$ is an increasing function of r. Using the definition of $O_{B(r)}$ it is clear that there exist r, N: $N_{B(t)} \leq N < 1$ for all $t \in]0, r]$.

(B 1) Possible choices for σ_k are:

(a) σ_k related to the defect: if $v \in B(w^*; t)$ then lemma 5.4(a) combined with lemma 3.4 implies that with $\tilde{\varphi} := n_{k,k-1}(v) - n_{k,k-1}(R_t(u))$ we have:

$$\|n_{k-1}^{-1}\{n_{k-1}(\tilde{u}_{k-1}) + \sigma_k(v, u)\,\tilde{\varphi}\} - \tilde{u}_{k-1}\|_{w^*} \leq (1 + \Gamma_t t)^{\frac{1}{2}} \sigma_k(v, u)\, \hat{\Gamma}_t \|\tilde{\varphi}\|_v \qquad (*)$$

$$\text{so} \quad \sigma_k(v, u) = \begin{cases} \min\{S, \eta t (1 + \Gamma_t t)^{-1} \hat{\Gamma}_t^{-1} \|\tilde{\varphi}\|_v^{-1}\} & \text{if } \tilde{\varphi} \neq 0 \\ \text{arbitrary in }]0, S] & \text{if } \tilde{\varphi} = 0 \end{cases}$$

suffices.

(b) σ_k constant: if $v \in B^k(w^*; t)$, $u \in B^k(v; t)$ then 5.12 results in $\|\tilde{\varphi}\|_v \leq \exp(\frac{1}{2}\Gamma_1 t)\, t$. So using this and the estimate (*) in (a) results in the choice: $\sigma_k(v, u) = \min\{S, \eta \hat{\Gamma}_t^{-1} \exp(-1\frac{1}{2}\Gamma_1 t)\}$. Note that with $\Gamma_1 t$ small enough and S suitable (cf. 5.11) this results in $\sigma_k(v, u) \approx \min(S, \eta)$.

(C 1) Possible choices for $(\tau_k)_{k\geq 1}$ are:

(a) take $\tau_0 \in \mathbb{N}$ such that $\varphi_0(x) := a_0 + b_0 x^{\tau_0}$ has a fixed point $x^* \in]0,1[$. Now $\tau_k = \tau_0$ for all $k \geq 1$ is a possible choice.

(b) let $\tau \in \mathbb{N}$ be the minimal element in $\mathbb{N}$ such that $\varphi(x) := A + Bx^\tau$ has fixed points $x_L, x_R \in]0,1[$ with $x_L < x_R$. Suppose $a_0 < x_R$. Now choose $\bar{y}$ such that $\max(x_L, a_0) < \bar{y} < x_R$. Define $\tau_k := \min\{m \in \mathbb{N} \mid a_k + b_k \bar{y}^m \leq \bar{y}\}$. The smallest fixed point of $\varphi_n(x) = a_n + b_n x^{\tau_n}$ is denoted by x_n^*. It is easy to verify that $\tau_k = \tau$ for k large enough, $x_n^* \leq \bar{y}$ for all n, $\lim_{n\to\infty} x_n^* = x_L$ and $x_{n+1} \leq \max(x_{n+1}^*, x_n)$ for all $n \geq 1$ ($(x_n)_{n\geq 1}$ as in 6.2 (C)). So the sequence $(x_n)_{n\geq 1}$ satisfies $\sup_n(x_n) < 1$. Note that if $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ are decreasing sequences then for $(\tau_n)_{n\geq 1}$ we have $\tau_{n+1} \leq \tau_n$, and $\lim_{n\to\infty} x_n = x_L$.

Definitions 6.4. In the following we will need a Lipschitz constant on a large enough ball. For this we define $r^{(4)} := r^{(2)} + (1 + \Gamma_r r)^{\frac{1}{2}} r$ with $r^{(2)}$ as in 5.7. And with γ_r as defined in 5.7: $\tilde{\gamma}_r := \max(\Gamma_{B(w^*; r^{(4)})}, \gamma_r)$.

We also define the function M_r: $]0, \infty[\to]0, \infty[$ by:

$$M_r(x) = 1\tfrac{1}{2}x + \tfrac{1}{2}x^2 + A_r(x, 2x(1+x)^{\frac{1}{2}})x \qquad (A_r \text{ as in 5.9(2)}).$$

This function will arise in the proof of 6.5.

Theorem 6.5. FAS multigrid convergence. *Let $\beta \in]0,1[$ be given, let r be such that 6.2 (A) is satisfied. Define $k_0 := k_-(\beta r)$ (k_- as defined above); assume sequences $(r_k)_{k\geq k_0}$, $(\tilde{u}_k)_{k\geq k_0}$ with $r_{k_0} := r$ and for $k > k_0$ r_k, $\tilde{u}_{k-1}$ such that $0 < r_k \leq r_{k-1}$ and $\tilde{u}_{k-1} \in B^{k-1}(w^*; \beta r_k)$ (this is possible; cf. definition of k_-). Also for $k > k_0$ assume σ_k as in 6.2 (B) with $\eta = 1 - \beta$, $t = r_k$ and let τ_k be as in 6.2 (C) with $a_k = N_{B_{r_k}}$ and $b_k = (1 + M_{r_k}(\tilde{\gamma}_{r_k} r_k))\, \hat{\Gamma}_{r_k} \exp(3\gamma_{r_k} r_k)$.*

Now let $\tilde{F}_k$ be as defined in 6.1 with $r_n, \tilde{u}_{n-1}, \sigma_n, \tau_n$ ($k_0 < n \leq k$) as above, then the following holds:

for all $k > k_0$, all $m \in S_k'$ with $v^* := n_k^{-1}(m) \in B^k(w^*; r_k)$, all $w \in B^k(v^*; r_k)$:

$$\|\tilde{F}_k(w,m) - v^*\|_{v^*} \leq \chi_k \|w - v^*\|_{v^*} \quad \textit{with}$$

$$\chi_{k_0+1} = a_{k_0+1}, \qquad \chi_{k+1} = a_{k+1} + b_{k+1}\chi_k^{\tau_{k+1}}, \qquad k > k_0 \quad (\textit{note that } \chi_k \textit{ depends on } w^*).$$

Proof. We will prove the assertion of the theorem by induction with respect to k. Note that if the assertion holds for some k, then also $\|\tilde{F}_k^n(w,m) - v^*\|_{v^*} \leq \chi_k^n \|w - v^*\|_{v^*}$ for $n \in \mathbb{N}$.

$k = k_0 + 1$: choose $m \in S_k$ with $v^* := n_k^{-1}(m) \in B^k(w^*; r_k)$ and $w \in B^k(v^*; r_k)$ (these exist). Then:

$$\begin{aligned}\|\tilde{F}_k(w,m) - v^*\|_{v^*} &= \|F_k(r_k, w, \tilde{u}_{k-1}, \sigma_k(v^*, w), m) - v^*\|_{v^*}\\ &\leq N_{B(r_k)} \|w - v^*\|_{v^*} = \chi_k \|w - v^*\|_{v^*}\end{aligned}$$

(use definition of $N_{B(r)}$ in 6.2 (A)).

Suppose that the assertion is true for $k-1$; we now consider k. We have:

$$\begin{aligned}\|\tilde{F}_k(w,m) - v^*\|_{v^*} \leq\, &\|F_k(r_k, w, \tilde{u}_{k-1}, \sigma_k(v^*, w), m) - v^*\|_{v^*}\\ &+ \|F_k(r_k, w, \tilde{u}_{k-1}, \sigma_k(v^*, w), m) - \tilde{F}_k(w,m)\|_{v^*}.\end{aligned}$$

Using the definition of $N_{B(r)}$ (6.2 (A)) we have:

$$\|F_k(r_k, w, \tilde{u}_{k-1}, \sigma_k(v^*, w), m) - v^*\|_{v^*} \leqq N_{B(r_k)} \|w - v^*\|_{v^*}.$$

We also have, using the notation as in 5.1 and 6.1:

$$\begin{aligned}\|\tilde{F}_k(w, m) - F_k(r_k, w, \tilde{u}_{k-1}, \sigma_k(v^*, w), m)\|_{v^*} &= \|R_{r_k} \tilde{C} R_{r_k}(w) - R_{r_k} C R_{r_k}(w)\|_{v^*} \\ &\leqq \|R_{v^*, r_k}\|_{v^*} \|\tilde{C} R_{r_k}(w) - C R_{r_k}(w)\|_{v^*} + \|d_{v^*}(\tilde{C} R_{r_k}(w), C R_{r_k}(w))\|_{v^*}. \qquad (1)\end{aligned}$$

The last inequality can be shown as in the proof of 5.9(1). Now use that $\|R_{v^*, r_k}\|_{v^*} \leqq 1$ (see 4.6); denoting $R_{r_k}(w)$ by y and $\sigma_k(v^*, w)$ by σ_k we get:

$$\begin{aligned}&\|\tilde{C} R_{r_k}(w) - C R_{r_k}(w)\|_{v^*} \\ &\quad = \|\tilde{C} y - C y\|_{v^*} = \|(y + \sigma_k^{-1}(\hat{y} - \tilde{u}_{k-1})) - (y + \sigma_k^{-1}(\bar{y} - \tilde{u}_{k-1}))\|_{v^*} \\ &\quad = \sigma_k^{-1} \|\tilde{F}_{k-1}^{\tau_k}(\tilde{u}_{k-1}, \alpha) - n_{k-1}^{-1}(\alpha)\|_{v^*} \quad \text{with} \\ &\quad \alpha = n_{k-1}(\tilde{u}_{k-1}) + \sigma_k(m - n_{k,k-1}(y)).\end{aligned}$$

Now note that because of 6.2(B) and $\tilde{u}_{k-1} \in B^{k-1}(w^*; \beta r_k)$:

$$\|n_{k-1}^{-1}(\alpha) - w^*\|_{w^*} \leqq \|n_{k-1}^{-1}(\alpha) - \tilde{u}_{k-1}\|_{w^*} + \beta r_k \leqq r_k \leqq r_{k-1}$$

so $$n_{k-1}^{-1}(\alpha) \in B^{k-1}(w^*; r_{k-1}). \qquad (2)$$

Also:

$$\|n_{k-1}^{-1}(\alpha) - \tilde{u}_{k-1}\|_{n_{k-1}^{-1}(\alpha)} \leqq (1 + \Gamma_{r_k} r_k)^{\frac{1}{2}} \|n_{k-1}^{-1}(\alpha) - \tilde{u}_{k-1}\|_{w^*} \leqq (1 - \beta) r_k \leqq r_{k-1},$$

so $$\tilde{u}_{k-1} \in B^{k-1}(n_{k-1}^{-1}(\alpha); r_{k-1}). \qquad (3)$$

Now (2) and (3) make that we may use the induction hypothesis, implying:

$$\begin{aligned}\|\tilde{C} R_{r_k}(w) - C R_{r_k}(w)\|_{v^*} &= \sigma_k^{-1} \|\tilde{F}_{k-1}^{\tau_k}(\tilde{u}_{k-1}, \alpha) - n_{k-1}^{-1}(\alpha)\|_{v^*} \\ &\leqq \sigma_k^{-1}(1 + 2\Gamma_{r_k} r_k)^{\frac{1}{2}} \|\tilde{F}_{k-1}^{\tau_k}(\tilde{u}_{k-1}, \alpha) - n_{k-1}^{-1}(\alpha)\|_{n_{k-1}^{-1}(\alpha)} \\ &\leqq \sigma_k^{-1}(1 + 2\Gamma_{r_k} r_k)^{\frac{1}{2}} \chi_{k-1}^{\tau_k} \|\tilde{u}_{k-1} - n_{k-1}^{-1}(\alpha)\|_{n_{k-1}^{-1}(\alpha)} \\ &\leqq \sigma_k^{-1}(1 + 2\Gamma_{r_k} r_k) \chi_{k-1}^{\tau_k} \|\tilde{u}_{k-1} - n_{k-1}^{-1}(\alpha)\|_{v^*} \\ &\leqq \chi_{k-1}^{\tau_k}(1 + 2\Gamma_{r_k} r_k) \hat{\Gamma}_{r_k} \|n_{k,k-1}(v^*) - n_{k,k-1}(y)\|_{v^*} \\ &\quad \text{(using 5.4(a))} \\ &\leqq \chi_{k-1}^{\tau_k}(1 + 2\Gamma_{r_k} r_k) \hat{\Gamma}_{r_k} \exp[\gamma_{r_k} r_k] \|v^* - w\|_{v^*} \\ &\quad \text{(use 5.12)} \\ &\leqq \chi_{k-1}^{\tau_k} \hat{\Gamma}_{r_k} \exp[3\gamma_{r_k} r_k] \|v^* - w\|_{v^*} \leqq r_k.\end{aligned}$$

The last inequality can be deduced from $\hat{\Gamma}_{r_k} \exp[3\gamma_{r_k} r_k] \chi_{k-1}^{\tau_k} < b_k \chi_{k-1}^{\tau_k} < 1$ (cf. definition of b_k and 6.2(C)).

Summarizing:

$$\|\tilde{C} R_{r_k}(w) - C R_{r_k}(w)\|_{v^*} \leqq \chi_{k-1}^{\tau_k} \hat{\Gamma}_{r_k} \exp[3\gamma_{r_k} r_k] \|w - v^*\|_{v^*} \leqq r_k. \qquad (4)$$

Now note that $\|C R_{r_k}(w) - w^*\|_{w^*} \leqq r_k^{(2)}$ (definition of $r^{(2)}$, see 5.7) and that

$$\begin{aligned}\|\tilde{C} R_{r_k}(w) - w^*\|_{w^*} &\leqq \|C R_{r_k}(w) - w^*\|_{w^*} + \|C R_{r_k}(w) - \tilde{C} R_{r_k}(w)\|_{w^*} \\ &\leqq r_k^{(2)} + (1 + \Gamma_{r_k} r_k)^{\frac{1}{2}} r_k = r_k^{(4)}\end{aligned}$$

(see definition of $r^{(4)}$ in 6.4 and use (4)).

Now returning to (1) and using Lemma 3.5 and (4) we get:

$$\|R_{r_k}\tilde{C}R_{r_k}(w)-R_{r_k}CR_{r_k}(w)\|_{v*}\leqq\|CR_{r_k}(w)-\tilde{C}R_{r_k}(w)\|_{v*}$$
$$+\tfrac{1}{2}\tilde{\gamma}_{r_k}(\|\tilde{C}R_{r_k}(w)-v^*\|_{v*}+\|CR_{r_k}(w)-v^*\|_{v*})\|CR_{r_k}(w)-\tilde{C}R_{r_k}(w)\|_{v*}$$
(definition $\tilde{\gamma}_r$: see 6.4)
$$\leqq[1+\tfrac{1}{2}\tilde{\gamma}_{r_k}(\|\tilde{C}R_{r_k}(w)-CR_{r_k}(w)\|_{v*}+2\|CR_{r_k}(w)-v^*\|_{v*})]\,\chi_{k-1}^{\tau_k}\hat{\Gamma}_{r_k}$$
$$\cdot\exp[3\gamma_{r_k}r_k]\,\|v^*-w\|_{v*}.$$

The reader may check (cf. proof of 5.9(3)) that:

$$\tfrac{1}{2}\tilde{\gamma}_{r_k}(\|\tilde{C}R_{r_k}(w)-CR_{r_k}(w)\|_{v*}+2\|CR_{r_k}(w)-v^*\|_{v*})$$
$$\leqq\tfrac{1}{2}\tilde{\gamma}_{r_k}r_k+\tilde{\gamma}_{r_k}\|CR_{r_k}(w)-v^*\|_{v*}\leqq M_{r_k}(\tilde{\gamma}_{r_k}r_k)$$

with $\quad M_r(x)=1\tfrac{1}{2}x+\tfrac{1}{2}x^2+A_r(x,2x(1+x)^{\frac{1}{2}})x \quad$ (A_r as in 5.9(2)).

Finally we get:

$$\|\tilde{F}_k(w,m)-v^*\|_{v*}\leqq N_{B(r_k)}\|w-v^*\|_{v*}+(1+M_{r_k}(\tilde{\gamma}_{r_k}r_k))\,\chi_{k-1}^{\tau_k}\hat{\Gamma}_{r_k}\exp[3\gamma_{r_k}r_k]\,\|v^*-w\|_{v*}$$
$$=[N_{B(r_k)}+\chi_{k-1}^{\tau_k}(1+M_{r_k}(\tilde{\gamma}_{r_k}r_k))\,\hat{\Gamma}_{r_k}\exp[3\gamma_{r_k}r_k]]\,\|v^*-w\|_{v*}$$
$$=(a_k+b_k\chi_{k-1}^{\tau_k})\,\|v^*-w\|_{v*}=\chi_k\|v^*-w\|_{v*}.\quad\square$$

Remark 6.6. Take $w^*=u^*$ $(=n^{-1}(f))$ and with β and r as in the theorem take $r_k=r$ for $k\geqq 0$, and the 0-level fine enough such that $\|u_k^*-u^*\|_{u*}\leqq\beta r$ for all $k\geqq 0$. Then $k_0=0$ and the theorem guarantees convergence of the FAS MG iteration on level $k\geqq 1$ for solving $n_k(u)=f$ when starting within $B^k(u_k^*;r)$ and the contraction number is bounded by $\max\{\chi_n|0<n\leqq k\}\leqq\sup\{\chi_n|n>0\}<1$. If in addition we assume α with $0<\alpha<\beta$ and the 0-level fine enough such that there exists $\tilde{u}_0\in S_0$ with $\|\tilde{u}_0-u_0^*\|_{u*}\leqq(\beta-\alpha)\,r$ and $\|u_0^*-u^*\|_{u*}\leqq\alpha r$, then the following choice for $\tilde{u}_n$ is possible: take $\tilde{u}_n=\tilde{u}_0$ for $n\geqq 1$, then:

$$\|\tilde{u}_n-u^*\|_{u*}\leqq\|\tilde{u}_0-u_0^*\|_{u*}+\|u_0^*-u^*\|_{u*}\leqq(\beta-\alpha)\,r+\alpha r=\beta r.$$

Another choice for $\tilde{u}_n$ will be mentioned in the next section.

Take $w^*=u_p^*$ for some p. Again take β and r as in the theorem and $r_k=r$ for $k\geqq k_0$ (k_0 as in the theorem). Assume $k_0<p$. Then the theorem guarantees convergence of the FAS MG iteration when starting within $B^p(u_p^*;r)$ and the contraction factor is bounded by $\max\{\chi_k|k_0<k\leqq p\}<1$ (a bound depending on p). This holds even if u_p^* is "far away" from u^* (coarse level). However, an important condition is $k_0<p$ which means that there are coarser levels than the p-level on which we can approximate u_p^* sufficiently accurate: $B^k(u_p^*;\beta r)\neq\phi$ for $k_0\leqq k<p$.

Remark 6.7. Let r be as in the theorem and take $r_k=r$ for all $k\geqq k_0$ (cf. 6.6). Then with $(a_k)_{k\geqq k_0}$, $(b_k)_{k\geqq k_0}$, $(\tau_k)_{k>k_0}$ as defined in the theorem we have for all $k>k_0$:

$$a_k=a_{k_0},\quad b_k=b_{k_0},\quad \tau_k=\tau_0,\quad \text{with } \tau_0 \text{ such that } a_{k_0}+b_{k_0}x^{\tau_0} \text{ has a fixed point.}$$

Suppose $a_{k_0}=N_{B(r)}<\frac{1}{4}$. Now if $\tilde{\gamma}_r r$ small enough and S small enough (cf. 5.11) we have that $b_{k_0}a_{k_0}<\frac{1}{4}$ so $\tau_0=2$. So we have local convergence of the FAS MG algorithm using W-cycles.

7. Nested Iteration

From the theorems in §5, §6 it is seen that a better starting vector (for FAS MG) not only results in a smaller initial error but also in a smaller bound for the contraction factor. Also better coarse grid approximations $\tilde{u}_k$ result in a smaller bound for the contraction factor. By using Nested Iteration we can make an algorithm that "automatically" generates good starting vectors and good coarse grid approximations $\tilde{u}_k$. In this section we specify a nested iteration algorithm and prove a convergence statement which induces nice (expected) properties of the algorithm.

Definition 7.1. We are now going to define a sequence of *nested iteration approximations* $v_k \in S_k$ of u_k^* $(k \geqq 0)$; u_k^* is the solution (in S_k) of $n_k(u_k)=f$.

For all $k \geqq 1$ let there be given $r_k > 0$, σ_k: $S_k \times S_k \to]0, \infty[$ and τ_k, $i_k \in \mathbb{N}$.

We also assume $v_0 \in S_0$ to be given. Now for $k \geqq 1$ we define $v_k := \tilde{F}_k^{i_k}(v_{k-1}, f)$ with $\tilde{F}_k$ the FAS MG iteration on level k as defined in 6.1 with r_n, σ_n, τ_n $(1 \leqq n \leqq k)$ as above and $\tilde{u}_{n-1} = v_{n-1}$ $(1 \leqq n \leqq k)$.

Assumptions 7.2. In this section for w^* we take $w^* = u^*$. We assume $\|u_k^* - u_{k+1}^*\|_{u_{k+1}^*} \leqq e_k$ and $\|u_k^* - u^*\|_{u^*} \leqq \tilde{e}_k$ $(k \geqq 0)$ with $e_k, \tilde{e}_k$ such that there are constants $e_+, e_-, \tilde{e}_+, \tilde{e}_-$:

$$0 < e_- \leqq e_{k+1} e_k^{-1} \leqq e_+ < 1, \quad 0 < \tilde{e}_- \leqq \tilde{e}_{k+1} \tilde{e}_k^{-1} \leqq \tilde{e}_+ < 1 \qquad \text{for all } k$$

(cf. 2.5: discretization error $\|u_k^* - u^*\| \leqq c h_k$).

Main-theorem 7.3. *Let* $\beta \in]0, 1[$. *For* $k \geqq 0$ *define (with* $e_{-1} = e_0$, $\tilde{e}_{-1} = \tilde{e}_0$*):*

$$r_k := \max\{((1 + \Gamma_{\tilde{e}_{k-1}} \tilde{e}_{k-1}) e_{k-1} + \tilde{e}_{k-1}) \beta^{-1}, (1 + \Gamma_{\tilde{e}_{k-1}} e_{k-1}) e_{k-1} + e_{k-1}\}.$$

Take r *such that* 6.2(A) *is satisfied. Assume* e_0, $\tilde{e}_0$ *small enough ("0-level fine enough") such that* $r_0 \leqq r$ *(then* k_0 *in theorem* 6.5 *equals zero).*

For all $k \geqq 1$ *let* σ_k, τ_k *and* χ_k *be as in* 6.5 *with* $(r_k)_{k \geqq 0}$ *as above.*

Take $i_k \in \mathbb{N}$, $i_k \geqq \log((2 + \Gamma_{\tilde{e}_{k-1}} e_{k-1}) e_-^{-1}) \log^{-1}(\chi_k^{-1})$ $(k \geqq 1)$. *Assume* $v_0 \in S_0$ *with* $\|v_0 - u_0^*\|_{u_0^*} \leqq e_0$. *Then the nested iteration approximations* v_k $(k \geqq 0)$ *as defined in* 7.1 *with* r_k, σ_k, τ_k *and* i_k *as above satisfy:*

$$\|v_k - u_k^*\|_{u_k^*} \leqq e_k.$$

Proof. $k=0$ is obvious. Now consider k, assuming $\|v_n - u_n^*\|_{u_n^*} \leqq e_n$ for all $n \leqq k-1$. Note that

$$\|u_n^* - u^*\|_{u^*} \leqq \tilde{e}_n \leqq r_n \quad \text{so} \quad u_n^* \in B^n(u^*; r_n) \text{ for all } n. \tag{1}$$

For $n \leqq k-1$:

$$\|v_n - u^*\|_{u^*} \leqq \|v_n - u_n^*\|_{u^*} + \|u_n^* - u^*\|_{u^*} \leqq \|v_n - u_n^*\|_{u^*} + \tilde{e}_n.$$

And because $u_n^* \in B^n(u^*; \tilde{e}_n)$, using 3.4(4) results in:

$$\|v_n - u_n^*\|_{u^*} \leqq (1 + \Gamma_{\tilde{e}_n} \|u_n^* - u^*\|_{u^*}) \|v_n - u_n^*\|_{u_n^*} \leqq (1 + \Gamma_{\tilde{e}_n} \tilde{e}_n) e_n$$
$$(n \leqq k-1).$$

So we get:
$$\|v_n - u^*\|_{u^*} \leqq (1+\Gamma_{\tilde{e}_n}\tilde{e}_n)\, e_n + \tilde{e}_n \leqq \beta r_{n+1} \qquad (n \leqq k-1). \tag{2}$$

Because $u_k^*, u_{k-1}^* \in B(u^*; \tilde{e}_{k-1})$ applying 3.4 (4) results in:

$$\|v_{k-1} - u_{k-1}^*\|_{u_k^*} \leqq (1+\Gamma_{\tilde{e}_{k-1}} \|u_k^* - u_{k-1}^*\|_{u_k^*})\, \|v_{k-1} - u_{k-1}^*\|_{u_{k-1}^*} \leqq (1+\Gamma_{\tilde{e}_{k-1}} e_{k-1})\, e_{k-1}$$

so

$$\|v_{k-1} - u_k^*\|_{u_k^*} \leqq \|v_{k-1} - u_{k-1}^*\|_{u_k^*} + \|u_{k-1}^* - u_k^*\|_{u_k} \leqq (1+\Gamma_{\tilde{e}_{k-1}} e_{k-1})\, e_{k-1} + e_{k-1} \leqq r_k. \tag{3}$$

Using (1), (2), (3) and the assumptions about σ_n, τ_n, we conclude that we may apply theorem 6.5 with

$$w^* = u^*, \tilde{u}_{n-1} = v_{n-1}\ (1 \leqq n \leqq k), v^* = u_k^*, w = v_{k-1},$$

resulting in:
$$\|\tilde{F}_k(v_{k-1}, f) - u_k^*\|_{u_k^*} \leqq \chi_k \|v_{k-1} - u_k^*\|_{u_k^*}$$

implying (see definition of v_k in 7.1):

$$\begin{aligned}
\|v_k - u_k^*\|_{u_k^*} &\leqq \chi_k^{i_k} \|v_{k-1} - u_k^*\|_{u_k^*} \\
&\leqq \chi_k^{i_k}(\|v_{k-1} - u_{k-1}^*\|_{u_k^*} + \|u_{k-1}^* - u_k^*\|_{u_k^*}) \leqq \chi_k^{i_k}(\|v_{k-1} - u_{k-1}^*\|_{u_k^*} + e_{k-1}) \\
&\leqq \chi_k^{i_k}((1+\Gamma_{\tilde{e}_{k-1}} e_{k-1})\, e_{k-1} + e_{k-1}) \leqq \chi_k^{i_k}(2+\Gamma_{\tilde{e}_{k-1}} e_{k-1})\, e_-^{-1} e_k \quad \text{(see 7.2)} \\
&\leqq e_k \quad \text{(using choice of } i_k). \quad \square
\end{aligned}$$

Concluding Remarks 7.4. Because $(r_k)_{k \geqq 0}$ in Theorem 7.3 is a decreasing sequence with $\lim_{k\to\infty} r_k = 0$ we have that $(b_k)_{k \geqq 0}$ (see 6.5) is a decreasing sequence with $\lim_{k\to\infty} b_k = 1$. We also have: $\lim_{k\to\infty} a_k = \lim_{r\downarrow 0} N_{B(r)} = \lim_{r\downarrow 0} L_{B(r)} =: L_0$ (linear two grid convergence factor). For ease we assume that $(a_k)_{k \geqq 0}$ is also a decreasing sequence (cf. definition of a_k in 6.5 and remark 4.11). In the following we use 6.3(C1)(b). Let $\varphi(x) = L_0 + x^\tau$ with $\tau \in \mathbb{N}$ minimal such that φ has two fixed points $x_L < x_R$. Now the $(\tau_k)_{k \geqq 1}$ in Theorem 7.3 are a decreasing sequence with $\tau_k = \tau$ for k large enough. As an example take $L_0 = \frac{3}{16}$, then $\tau_k = 2$ for k large enough (i.e., "W-cycles on high levels") and $x_L = \frac{1}{4}$, $x_R = \frac{3}{4}$. Assume that $r = r_0$ is small enough such that $a_0 = N_{B(r_0)} < \frac{3}{4}$. Now for the sequence $(\chi_k)_{k \geqq 1}$ we have that $\chi_k \leqq \frac{3}{4}$ for all k and $\lim_{k\to\infty} \chi_k = \frac{1}{4}$ (cf. 6.3(C1)(b)). Now for i_k in Theorem 7.3 we have that, assuming $\Gamma_{\tilde{e}_{k-1}} e_{k-1} \leqq 0.1$, $i_k = 5$ suffices for all k, but for $k \to \infty$ $i_k = 1$ already suffices ("approximately one FAS MG iteration on high levels"). The above shows that using a suitable nonlinear nested iteration (consisting of a bounded number of suitable FAS MG iterations on each level), starting on a coarsest level that is fine enough with a good enough approximation of the discrete solution on that level we get an approximation of the discrete solution within the relative discretization error on arbitrarily fine levels. Besides, this nested iteration has the nice property that we need fewer FAS MG iterations on finer levels and in the FAS MG iterations we need fewer coarse grid corrections on finer levels.

8. Simple Conditions on the Function g Instead of Conditions on the Operator n

The assumptions we made in the foregoing sections about the operator $n: u \to a(u,.)+(b g \circ u,.)_{L^2}$ are stated in 3.1 and 4.7. In this section we give conditions on g which induce that n has the properties as assumed in 3.1 and 4.7.

Definition 8.1. Let $h: \mathbb{R} \to \mathbb{R}$. For $p \geq 1$ we define: h has property $LBA(p)$ ("Lipschitz-continuous for bounded arguments") iff the following holds:
$u \to h \circ u$ maps H into $L^p(\Omega)$ and for every bounded subset B of H we have a finite Lipschitz constant

$$\Gamma(B;p) := \sup\{\|h \circ u - h \circ v\|_{L^p} \|u-v\|^{-1} \mid u, v \in B, u \neq v\}.$$

Remark 8.2. If h has property $LBA(p)$ then h has property $LBA(q)$ for all q with $1 \leqq q \leqq p$.

Lemma 8.3. *Let $h \in C^1(\mathbb{R})$. If for some $m \in [0, \infty[$ $\sup\{|h'(t)|(1+|t|^m)^{-1} \mid t \in \mathbb{R}\} < \infty$ then h has property $LBA(p)$ for all $p \in [1, \infty[$.*

Proof. Take $p \in [1, \infty[$ and let B be a bounded subset of H with $\|u\| \leqq M$ for all $u \in B$. For $\alpha, \beta \in \mathbb{R}$, $I_{\alpha,\beta}$ denotes the closed interval in $\mathbb{R}$ with endpoints α and β.

Now for $u, v \in B$ and for almost all $x \in \Omega$ we have, with suitable constant C:

$$|h(u(x)) - h(v(x))| = \left| \int_{v(x)}^{u(x)} h'(t)\, dt \right| \leqq \max_{t \in I_{u(x),v(x)}} |h'(t)|\, |u(x)-v(x)|$$
$$\leqq \max_{t \in I_{u(x),v(x)}} C(1+|t|^m)|u(x)-v(x)| \leqq C(1+|u(x)|^m + |v(x)|^m)\, |u(x)-v(x)|.$$

So we get, using the embedding theorem of 2.4:

$$\|h \circ u - h \circ v\|_{L^p} \leqq C(\|1\|_{L^{2p}} + \|u\|^m_{L^{2pm}} + \|v\|^m_{L^{2pm}})\, \|u-v\|_{L^{2p}}$$
$$\leqq C(\|1\|_{L^{2p}} + 2(d_{2pm} M)^m)\, d_{2p} \|u-v\|. \quad \square$$

Lemma 8.4. *The collection of all bounded subsets of H is denoted by $\mathscr{B}$. The following holds:*

(a) *if g has property $LBA(m)$ for some $m > 1$ then:*
  (a_1) *the operator $n: u \to a(u,.)+(b g \circ u,.)_{L^2}$ maps H into H';*
  (a_2) $\forall B \in \mathscr{B}\ \exists c: \forall v, w \in B\ \forall u \in H: |(n(v)-n(w))(u)| \leqq c\, \|v-w\|\, \|u\|$;

(b) *if g' has property $LBA(m)$ for some $m > 1$ then:*
  (b_1) *g has property $LBA(s)$ for all $s \in [1, m[$;*
  (b_2) *n is Frèchet differentiable on H with derivative*
  $Dn(w)(u)(z) = a(u,z) + (b(g' \circ w)\, u, z)_{L^2}$;
  (b_3) $\forall B \in \mathscr{B}\ \exists c: \forall v, w \in B\ \forall u, z \in H: |(Dn(v)-Dn(w))(u)(z)| \leqq c\, \|v-w\|\, \|u\|\, \|z\|$;
  (b_4) *if $m > 2$ then:*
  $\forall B \in \mathscr{B}\ \exists c: \forall v, w \in B\ \forall u, z \in H: |(Dn(v)-Dn(w))(u)(z)| \leqq c\, \|v-w\|\, \|u\|\, \|z\|_{L^2}$;

Proof. For $p \in]1, \infty[$ we denote $\left(1-\frac{1}{p}\right)^{-1} = \frac{p}{1-p}$ by p' $\left(\frac{1}{p'}+\frac{1}{p}=1\right)$.

(a) for $p \in [1, m]$, $B \in \mathscr{B}$ let $\Gamma(B;p)$ be as in 8.1 with $h = g$. Take $q \in]1, m]$ and $u, v \in H$. Now the following holds:

$$|a(u,v)+(b\,g\circ u,v)_{L^2}|\leqq a_+\,\|u\|\,\|v\|+b_+\,\|g\circ u\|_{L^q}\,\|v\|_{L^{q'}}$$
$$\leqq(a_+\,\|u\|+b_+\,\|g\circ u\|_{L^q}I(q'))\,\|v\|;$$

this proves (a_1).

Take $B\in\mathscr{B}$ and $v, w\in B$, $u\in H$. Now we have:

$$|(n(v)-n(w))(u)|\leqq a_+\,\|v-w\|\,\|u\|+b_+\,\|g\circ v-g\circ w\|_{L^q}\,\|u\|_{L^{q'}}$$
$$\leqq(a_+ + b_+\,\Gamma(B;q)\,I(q'))\,\|v-w\|\,\|u\|;$$

this proves (a_2).

(b) for $p\in[1,m]$ $B\in\mathscr{B}$ let $\Gamma'(B;p)$ be as in 8.1 with $h=g'$.

Take $B\in\mathscr{B}$, $p, q, r\in]1,\infty[$ with $p\leqq m$ and $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=1$, $v, w, y\in B$, $z\in L^r$. Then:

$$|(g\circ y-g\circ v-(g'\circ w)(y-v),z)_{L^2}|$$
$$\leqq\int_\Omega\int_0^1|g'(v(x)+t(y(x)-v(x)))-g'(w(x))|\,|y(x)-v(x)|\,|z(x)|\,dt\,dx$$
$$=\int_0^1(|g'\circ(v+t(y-v))-g'\circ w|,|y-v|\,|z|)_{L^2}\,dt\quad\text{(using Fubini's theorem)}$$
$$\leqq\int_0^1\|g'\circ(v+t(y-v))-g'\circ w\|_{L^p}\,\|y-v\|_{L^q}\,\|z\|_{L^r}\,dt$$
$$\leqq\Gamma'(\mathrm{conv}(B);p)\int_0^1\|(1-t)(v-w)+t(y-w)\|\,dt\,I(q)\,\|y-v\|\,\|z\|_{L^r}$$

(conv(B): convex hull of B)

$$\leqq\tfrac{1}{2}\Gamma'(\mathrm{conv}(B);p)\,I(q)(\|v-w\|+\|y-w\|)\,\|y-v\|\,\|z\|_{L^r}\qquad(*)$$

To prove (b_1) take $s\in]1,m[$ and p, q, r as above with $p\in]s,m[$, $\frac{1}{r}=1-\frac{1}{s}$. Then the functional $z\to(g\circ y-g\circ v-(g'\circ w)(y-v),z)_{L^2}$ is an element of $(L^r)'$, and

$$\|g\circ y-g\circ v-(g'\circ w)(y-v)\|_{L^s}\leqq c_B\,\|y-v\|\quad\text{with}\quad c_B:=\Gamma'(\mathrm{conv}(B);p)\,I(q)\,\mathrm{diam}(B).$$

Taking $w\in B$ fixed and $t>1$ such that $ts<m$ we have that for all $y, v\in B$:

$$\|g\circ y-g\circ v\|_{L^s}\leqq\|g\circ y-g\circ v-(g'\circ w)(y-v)\|_{L^s}+\|g'\circ w\|_{L^{st}}\,\|y-v\|_{L^{st'}}$$
$$\leqq(c_B+\|g'\circ w\|_{L^{st}}I(st'))\,\|y-v\|.$$

This proves (b_1).

Assume B is some open ball, $w\in B$ fixed. Define the continuous linear operator $\hat{b}_w\colon H\to H'$ by $\hat{b}_w(h)(e)=a(h,e)+(b(g'\circ w)\,h,e)_{L^2}$. Using (*) with $y=w+h\in B$, $v=w$, $z\in H$ *with* $\|z\|=1$ we get:

$$|n(w+h)(z)-n(w)(z)-\hat{b}_w(h)(z)|$$
$$=|(b(g\circ(w+h)-g\circ w-(g'\circ w)\,h),z)_{L^2}|\leqq b_+\,\tfrac{1}{2}\Gamma'(B;p)\,I(q)\,I(r)\,\|h\|^2.$$

Now (b_2) easily follows.

For (b_3) and (b_4) note that with $v, w \in B$, $u, z \in H$:

$$\begin{aligned}|(Dn(v)-Dn(w))(u)(z)| &= |(b(g'\circ v - g'\circ w)\, u, z)_{L^2}| \\ &\leqq b_+ \|g'\circ v - g'\circ w\|_{L^p} \|u\|_{L^q} \|z\|_{L^r} \leqq b_+ \Gamma'(B;p)\, I(q)\, \|v-w\|\, \|u\|\, \|z\|_{L^r}. \quad (**)\end{aligned}$$

Now (b_3) follows using $\|z\|_{L^r} \leqq I(r)\|z\|$.

If $m>2$ we can take $p \in]2, m]$ and $r \leqq 2$, and (b_4) now follows from (**). □

Remark 8.5. With respect to the Lemmas 8.3 and 8.4 we note that a restriction on the growth of the function g is necessary to get a reasonable operator n. This is shown by the following example.

Let $\tilde{\Omega}(r) := \{(x,y) \in \mathbb{R}\times\mathbb{R} \mid x^2+y^2<r\}$ $(r>0)$.
For $\varepsilon \in]0, \frac{1}{2}[$: $u(x,y) := \log^{\varepsilon}((x^2+y^2)^{-1})$ on $\tilde{\Omega}(\frac{1}{4})$, $g(t) := \exp(t^{\varepsilon^{-1}})$ for $t \in [0, \infty[$ and g is defined on $]-\infty, 0[$ such that $g \in C^1(\mathbb{R})$ and $g' \geqq 0$.

Also $v(x,y) := 1$ on $\tilde{\Omega}(\frac{1}{4})$ and v is suitably extended such that $v \geqq 0$ and $v \in H_0^1(\tilde{\Omega}(1))$. It is easy to check that $u \in H^1(\tilde{\Omega}(\frac{1}{4}))$; now extend u such that $u \geqq 0$ and $u \in H_0^1(\tilde{\Omega}(1))$. Now we have (taking $b \equiv 1$):

$$\begin{aligned}|n(u)(v)| &= |a(u,v) + (g\circ u, v)_{L^2}| \geqq |(g\circ u, v)_{L^2}| - a_+ \|u\|\,\|v\| \\ &= \int_{\tilde{\Omega}(1)}\int g(u(x,y))\, v(x,y)\, dx\, dy - a_+\|u\|\,\|v\| \geqq \int_{\tilde{\Omega}(\frac{1}{4})}\int \frac{1}{x^2+y^2}\, dx\, dy - a_+\|u\|\,\|v\| \\ &= \int_0^{2\pi}\int_0^{\frac{1}{2}} \frac{1}{r}\, \mathrm{d}r\, d\varphi - a_+\|u\|\,\|v\| = +\infty.\end{aligned}$$

So for this (too rapidly growing) function g we have that n does not map H into H'. This also implies (use 8.4.(a_1)) that g does not have property $LBA(m)$ for any $m>1$ (cf. 8.3).

Remark 8.6. Using Lemma 8.4 it is clear that if g' has property $LBA(m)$ for some $m>1$ then the operator n is such that Assumption 3.1 is fulfilled.

Lemma 8.7. *Let Assumption* 3.1 *be fulfilled. Assumption* 4.7 *is fulfilled if* n *is such that:*

(1) b_{w*} *is two-regular, i.e.:* $\exists d_1$: $\forall m \in L^2(\Omega) \subset H'$: $b_{w*}^{-1} m \in H^2(\Omega)$ *and* $\|b_{w*}^{-1} m\|_2 \leqq d_1 \|m\|_{L^2}$.
(2) $\forall r \geqq 0\ \exists c_r$: $\forall v \in B(r)\ \forall u, z \in H$: $|(b_v - b_{w*})(u)(z)| \leqq c_r \|u\|\, \|z\|_{L^2}$.

Proof. Note that b_v is uniformly one-regular, i.e.:

$$\exists d_2 : \forall v \in H\ \forall m \in L^2(\Omega) : \|b_v^{-1} m\| \leqq d_2 \|m\|_{L^2}.$$

This is clear using 3.4(3):

$$\begin{aligned}\|b_v^{-1} m\|^2 &= (b_v^{-1} m, b_v^{-1} m) \leqq a_-^{-1} b_v(b_v^{-1} m, b_v^{-1} m) = a_-^{-1}(m, b_v^{-1} m)_{L^2} \\ &\leqq a_-^{-1} I(2)\, \|m\|_{L^2} \|b_v^{-1} m\|.\end{aligned}$$

Take $r>0$, $v \in B(r)$ and $m \in L^2(\Omega)$. Note that b_{w*}^{-1} maps $L^2(\Omega)$ into $H^2(\Omega)$ and $b_v - b_{w*}$ maps H into $L^2(\Omega) \subset H'$. Using $b_v^{-1} = b_{w*}^{-1} - b_{w*}^{-1}(b_v - b_{w*})\, b_v^{-1}$ we have that b_v^{-1} maps $L^2(\Omega)$ into $H^2(\Omega)$ and taking norms results in:

$$\|b_v^{-1} m\|_2 \leqq d_1 \|m\|_{L^2} + d_1 \|(b_v - b_{w*})\, b_v^{-1} m\|_{L^2} \leqq d_1(1 + c_r a_-^{-1} I(2))\, \|m\|_{L^2}. \quad \square$$

Remark 8.8. Note that (1) and (2) in Lemma 8.7 are fulfilled if g' has property $LBA(m)$ for some $m>2$. This is easily concluded from 8.4 (b_2) and (b_4) and the well-known fact (see e.g. [8] Ch. III) that $b_{w*}(=Dn(w^*))$ of the form as in 8.4(b_2) is two-regular (use that $g' \circ w^* \in C(\bar{\Omega})$).

Combining this with lemma 8.7 and the remarks 8.6 and 8.2 we have that the operator n satisfies the Assumptions 3.1 and 4.7 if g' has property $LBA(m)$ for some $m>2$. This last condition is fulfilled (Lemma 8.3) if $g \in C^2(\mathbb{R})$ and g'' has atmost polynomial growth.

Acknowledgements. The author is indebted to Professor A. v.d. Sluis and Dr. G.L.G. Sleijpen for helpful discussions.

References

1. Bank, R.E., Douglas, C.C.: Sharp estimates for multigrid rates of convergence with general smoothing and acceleration. SIAM J. Numer. Anal. **22**, 617–633 (1985)
2. Brandt, A.: Guide in Multigrid Development. In: Multigrid Methods (W. Hackbusch, U. Trottenberg, eds.) pp. 220–312. Proceedings, Köln-Porz, 1981. Lect. Notes Math., Vol. 960. Berlin, Heidelberg, New York: Springer 1982
3. Brezis, H.: Operateurs Maximaux Monotones et semi-groupes de contractions dans les espaces de Hilbert. Amsterdam, London: North-Holland 1973; New York: American Elsevier 1973
4. Brezis, H., Crandall, M.G., Pazy, A.: Perturbations of Nonlinear Maximal Monotone Sets in Banach space. Commun. Pure Appl. Math. **23**, 123–144, 1970
5. Ciarlet, P.G.: The finite element method for elliptic problems. Amsterdam, New York, Oxford: North-Holland 1978
6. Hackbusch, W.: Multigrid Methods and Applications. Berlin, Heidelberg, New York, Tokyo: Springer 1985
7. Hackbusch, W., Trottenberg, U. (eds.). Multigrid Methods. Proceedings, Köln-Porz, 1981. Lect. Notes Math., Vol. 960. Berlin, Heidelberg, New York: Springer 1982
8. Necas, J.: Les methodes directes en theorie des equations elliptiques. Paris, Masson et C^{ie} 1967
9. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. New York, London: Academic Press 1970
10. Bank, R.E., Rose, D.J.: Analysis of a Multilevel Iterative Method for Nonlinear Finite Element Equations. Math. Comput. **39**, 453–465 (1982)
11. Maitre, J.F., Musy, F.: Multigrid Methods: Convergence Theory in a Variational Framework. SIAM J. Numer. Anal. **21**, 657–671 (1984)
12. Deimling, K.: Nonlinear Functional Analysis. Berlin, Heidelberg, New York: Springer 1985

Received April 6, 1987 / September 23, 1987