NEW NONLINEAR MULTIGRID ANALYSIS*

Dexuan Xie
Courant Institute of Mathematical Sciences
New York University
251 Mercer St. New York, NY 10012

SUMMARY

The nonlinear multigrid is an efficient algorithm for solving the system of nonlinear equations arising from the numerical discretization of nonlinear elliptic boundary problems [7],[9]. In this paper, we present a new nonlinear multigrid analysis as an extension of the linear multigrid theory presented by Bramble, et al. in [5], [6], and [17]. In particular, we prove the convergence of the nonlinear *V-cycle* method for a class of mildly nonlinear second order elliptic boundary value problems which do not have full elliptic regularity.

INTRODUCTION

Multigrid methods have been used extensively to solve linear systems of equations which arise in the numerical discretization of linear partial differential equations. We call such multigrid methods "linear multigrid methods" in this paper. With the development of the linear multigrid methods, the multigrid technique also has been applied to the numerical solution of nonlinear boundary value problems. Two important algorithms have been proposed so far. One is Newton-multigrid iteration, in which a linear multigrid method is used to solve the linear system that arises from a Newton iterative method [4]. The other one is the nonlinear multigrid method, which is an extension of the linear multigrid method to the nonlinear case [9]. In literature, it is also referred to as the Full Approximation Scheme (FAS) by Brandt in [7]. The convergence of the nonlinear multigrid method was first studied by Hackbusch in [9] and later by Reusken in [11] and [12]. Hackbusch's nonlinear multigrid theory is based on his linear multigrid theory, while Reusken's analysis is based on the linear multigrid analysis in [3].

Recently, Bramble, et al. have established a new linear multigrid theory [5] [6] [17] that has generalized the work in [3] and [9] in another way. Using this new multigrid theory, they have proved the convergence of linear multigrid methods with non-nested spaces or non-inherited quadratic forms, even with weak or no regularity assumptions. The purpose of this paper is to extend this new linear multigrid theory to the nonlinear case.

In this paper, we present the framework of our new multigrid theory. In particular, we prove a basic convergence theorem for the nonlinear *V-cycle* scheme based on two abstract conditions, which are referred to as the "smoothing assumption" and the "approximation assumption".

*This work was supported in part by the National Science Foundation through award number DMS-9105437 at the University of Houston.

We then apply it to show the convergence of the nonlinear *V-cycle* method with the damped-Jacobi-Newton smoother for a class of mildly nonlinear second order elliptic boundary value problems which do not have full elliptic regularity. Moreover, our new approach makes it possible to analyze the nonlinear multigrid method in more complicated cases, such as, non-nested spaces, non-inherited quadratic forms, numerical integration, and with weak or no regularity assumptions. We have shown the convergence of the nonlinear *V-cycle* method disturbed by numerical quadratures in [14]. We intend to study other cases in subsequent work.

In comparison to the linear multigrid method, the nonlinear multigrid method has two additional parameters. In practice, their choice is an important issue. We investigate this issue numerically through a model problem in this paper. We note that this model problem, in part, aids in the understanding of the solution procedures used in the code *UHBD* [10].

The outline of the remainder of the paper is as follows. In Section 2, we introduce the basic idea of our nonlinear multigrid analysis. In Section 3, we present a general convergence theorem of the nonlinear *V-cycle* method based on two abstract assumptions, the smoothing assumption and the approximation assumption. In Section 4, we apply the theory of Section 3 to show the convergence of the nonlinear multigrid method for a class of mildly nonlinear elliptic boundary value problems. In Section 5, we present numerical experiments with the nonlinear multigrid method focusing on its two auxiliary parameters.

THE NONLINEAR MULTIGRID METHOD

We consider a nonlinear variational problem coming from a nonlinear elliptic boundary value problem with domain Ω as follows: Find $u \in H$, such that

$$a(u, v) = 0 \qquad \forall v \in H, \tag{1}$$

where $H = H(\Omega)$ is an abstract Hilbert space with inner product $(\cdot, \cdot)$, and $a(\cdot, \cdot)$ is nonlinear only with respect to the first variable.

We assume that $a(u, v)$ is H-bounded, that is, there exists a constant C, such that

$$|a(u, v)| \leq C(1 + \|u\|)\|v\| \quad \forall u, v \in H,$$

where $\|u\| = \sqrt{(u, u)}$. Using the Riesz representation theorem [1], we then write (1) as

$$g(u) = 0, \tag{2}$$

where $g : H \to H$ is the nonlinear operator such that

$$a(u, v) = (g(u), v) \qquad \forall v \in H.$$

We make another assumption on g below:

A1) g is Frechet-differentiable on H, and the derivative of g at u, denoted by $Dg(u)$, is a symmetric, positive definite, bounded linear operator from H to itself.

From A1) it follows that Equation (2) has the unique solution u^* [16].

Let $\mathcal{U} \subseteq H$ be a neighborhood of u^* and $\mathcal{F}$ be the image of $\mathcal{U}$ under g. Since g satisfies the above assumptions, the implicit function theorem [1] implies that $g : \mathcal{U} \to \mathcal{F}$ is a homeomorphism. Thus, for any $f \in \mathcal{F}$, there exists unique $u \in \mathcal{U}$, such that the following equation holds:

$$g(u) = f. \tag{3}$$

Hence, we may consider equation (3) in the following.

Let u^{old} be an approximate solution of (3). The update u^{new} of u^{old} is defined by

$$u^{new} = u^{old} + q,$$

with q being a correction term satisfying the following correction equation of u^{old}:

$$g(q + u^{old}) = f. \tag{4}$$

If q is an exact solution of (4), then a direct method for solving (3) is derived. But solving (4) is as difficult as solving (3), so we often construct an approximate operator R of g^{-1} to simplify the computational work.

In the linear case, the correction equation (4) is often written as

$$g(q) = f - g(u^{old}), \tag{5}$$

and the term $f - g(u^{old})$ is often referred to as the residual of u^{old}. Clearly, if the operator R is defined by a linear iterative algorithm, then the linear iteration can be written as follows:

$$u^{new} = u^{old} + R[f - g(u^{old})]. \tag{6}$$

A key factor in the new linear multigrid theory in [5], [6] and [17] is the introduction of the operator R that characterizes the linear multigrid method, so the linear multigrid method can be expressed in form (6).

However, when g is nonlinear, the correction equation (4) cannot be written as (5). Noting the important role of the residual term in the context of the multigrid method, we introduce an "approximate" correction equation of (4) as follows:

$$g(s\hat{q} + \tilde{u}) = \tilde{f} + s[f - g(u^{old})], \tag{7}$$

where $\tilde{f} = g(\tilde{u})$, s is a given positive number and $\tilde{u}$ a given vector. Both s and $\tilde{u}$ are extra parameters, compared to the linear multigrid method, and they are chosen so that $\hat{q}$ approximates the solution q of (4) in some sense. Hence, the nonlinear multigrid method can be expressed by

$$u^{new} = u^{old} + \left[R(\tilde{f} + s[f - g(u^{old})]) - \tilde{u}\right]/s, \tag{8}$$

provided that the operator R is defined by the nonlinear multigrid iterative algorithm for solving $g(u) = f$. This is the main idea of our nonlinear multigrid analysis.

In the linear case, we can simply set $\tilde{u} = \tilde{f} = 0$ and $s = 1$. Thus, (8) reduces to (6). In this sense, the nonlinear multigrid method defined by (8) is an extension of the linear multigrid method.

To define a nonlinear multigrid operator, we need some further notation given below.

Let H be a finite element space with grid size h. Suppose that we have subspaces M_k with inner product $(\cdot,\cdot)_k$ satisfying

$$M_1 \subset M_2 \subset \cdots \subset M_l = H.$$

Set $g_l = g$, and define the nonlinear operator $g_k : M_k \to M_k$ by

$$(g_k(u), v)_k = a(u, v), \qquad \forall v \in M_k, \quad k = 1, 2, \cdots, l-1. \tag{9}$$

We define a projector $Q_k : M_{k+1} \to M_k$ by

$$(Q_k u, v)_k = (u, v)_{k+1}, \qquad \forall v \in M_k.$$

Obviously, g_k satisfies Assumption A1), so there exist $\mathcal{U}_k$ and $\mathcal{F}_k$ such that g_k is a homeomorphism between them. Hence, for $f_k \in \mathcal{F}_k$, we may consider the following equation

$$g_k(u) = f_k, \tag{10}$$

and its solution is denoted by u_k^*.

The smoothing process on M_k is denoted by the operator

$$S_k^m(\cdot; f_k) : M_k \to M_k \tag{11}$$

satisfying $u_k^* = S_k^m(u_k^*; f_k)$. We assume that S_k^m is Frechet-differentiable on M_k. Here m indicates that S_k^m may be defined by m steps of a nonlinear relaxation iteration (e.g., the damped-Jacobi-Newton or the Gauss-Seidel-Newton [13]). Without confusion, we denote $S_k^m(u; f_k)$ as $S_k^m(u)$.

Denote $\Xi_k = \{\zeta \mid \zeta = \tilde{f}_k + s_k[f_k - g_k(u_k)]$ for all $f_k \in M_k\}$. Here $\tilde{u}_k, s_k$ and u_k are fixed, and $\tilde{f}_k = g_k(\tilde{u}_k)$. We define the nonlinear multigrid operator B_k on Ξ_k inductively in the following algorithm:

Algorithm 1 *Given positive integers m_1, m_2 and p.*

0) $B_1 = g_1^{-1}$.

For each $\zeta_k \in \Xi_k$ with $k > 1$, there exists an $f_k \in M_k$ such that $\zeta_k = \tilde{f}_k + s_k[f_k - g_k(u_k)]$. We define $B_k(\zeta_k)$ in terms of B_{k-1} as follows:

1) Pre-smoothing : $v_1 = S_k^{m_1}(u_k; f_k)$.

2) Coarse grid correction: $v_2 = v_1 + \dfrac{q_p - \tilde{u}_{k-1}}{s_{k-1}}$,

where q_p is defined by (12).

$$q_i = q_{i-1} + \Big[B_{k-1}(\tilde{f}_{k-1} + s_{k-1}[f_{k-1} - g_{k-1}(q_{i-1})]) - \tilde{u}_{k-1}\Big] / s_{k-1}, \tag{12}$$

for $i = 1, 2, \cdots, p$. Here $q_0 = \tilde{u}_{k-1}$, and

$$f_{k-1} = \tilde{f}_{k-1} + s_{k-1} Q_{k-1}[f_k - g_k(v_1)]. \tag{13}$$

3) Post-smoothing :

$$B_k(\zeta_k) = s_k[S_k^{m_2}(v_2; f_k) - u_k] + \tilde{u}_k. \tag{14}$$

We note that Algorithm 1 using $u_k = \tilde{u}_k = 0$, $s_k = 1$, and $p = 1$ reduces to the linear multigrid algorithm described in [5], [6] and [17] provided that g is linear.

THE CONVERGENCE ANALYSIS

In our nonlinear multigrid analysis, we need a new inner product $b_k(u,v)$ defined by

$$b_k(u,v) = (Dg_k(u_k^*)u, v)_k, \quad \forall u, v \in M_k.$$

From Assumption A1) we see that $b_k(u,v)$ is symmetric, positive definite.

With this new inner product, we define an orthogonal operator $P_k : M_{k+1} \to M_k$ by

$$b_k(P_k u, v) = b_{k+1}(u,v) \qquad \forall v \in M_k.$$

From the definitions of Q_k and P_k an important equality follows:

$$Q_{k-1}Dg_k(u_k^*) = Dg_{k-1}(u_{k-1}^*)P_{k-1}, \quad k = 1, 2, \cdots, l. \tag{15}$$

Using the nonlinear multigrid operator B_k, we define the nonlinear multigrid method as follows:

$$u_k^{j+1} = \psi_k(u_k^j) \qquad j = 0, 1, 2, \cdots, \tag{16}$$

with the operator $\psi_k : M_k \to M_k$ being defined by

$$\psi_k(u_k) = u_k + \left[B_k(\tilde{f}_k + s_k[f_k - g_k(u_k)]) - \tilde{u}_k\right]/s_k. \tag{17}$$

Noting that $g_k(\tilde{u}_k) = \tilde{f}_k$ and $S_k^{m_i}(\tilde{u}_k; \tilde{f}_k) = \tilde{u}_k$ for $i = 1, 2$, we can show by induction that

$$B_k(\tilde{f}_k) = \tilde{u}_k. \tag{18}$$

Thus, the scheme (16) is consistent in the sense that u_k^* is a fixed point of the sequence $\{u_k^j\}$.

A fundamental recurrence relation with respect to the nonlinear multigrid operators B_k is given in the following theorem.

Theorem 1 *The fundamental recurrence relation for the nonlinear multigrid operators B_k, defined by Algorithm 1, is*

$$\begin{aligned} I - DB_k(\tilde{f}_k)Dg_k(u_k^*) &= DS_k^{m_2}(u_k^*)\{I - [I - (I - DB_{k-1}(\tilde{f}_{k-1})Dg_{k-1}(\tilde{u}_{k-1}))^p] \\ &\quad Dg_{k-1}(\tilde{u}_{k-1})^{-1}Dg_{k-1}(u_{k-1}^*)P_{k-1}\}DS_k^{m_1}(u_k^*), \end{aligned} \tag{19}$$

where $k = 1, 2, \cdots, l$, and u_k^* is a solution of $g_k(u_k) = f_k$ on M_k.

Proof. Using (14), we immediately get the following equality:

$$u_k + \left[B_k(\tilde{f}_k + s_k[f_k - g_k(u_k)]) - \tilde{u}_k\right]/s_k = S_k^{m_2}(S_k^{m_1}(u_k) + \frac{q_p(u_k) - \tilde{u}_{k-1}}{s_{k-1}}), \forall u_k \in M_k. \tag{20}$$

The expression (13) of $f_{k-1}(u)$ follows

$$f_{k-1}(u_k^*) = \tilde{f}_{k-1}. \tag{21}$$

Then, by the induction and (18), we can show that

$$q_i(u_k^*) = \tilde{u}_{k-1}, \text{ for } \ i = 0, 1, 2, \cdots, p. \tag{22}$$

Thus, differentiating with respect to u_k at u_k^* on both sides of the equality (20), and using (22), we get

$$I - DB_k(\tilde{f}_k)Dg_k(u_k^*) = DS_k^{m_2}(u_k^*)[DS_k^{m_1}(u_k^*) + Dq_p(u_k^*)/s_{k-1}]. \tag{23}$$

Here the operations are based on the calculus in Hilbert space [1].

Using (21) and (22), we see that

$$Dq_i(u_k^*) = [I - DB_{k-1}(\tilde{f}_{k-1})Dg_{k-1}(\tilde{u}_{k-1})]Dq_{i-1}(u_k^*) + DB_{k-1}(\tilde{f}_{k-1})Df_{k-1}(u_k^*).$$

In addition, with (13) and (15),

$$Df_{k-1}(u_k^*) = -s_{k-1}Q_{k-1}Dg_k(u_k^*)DS_k^m(u_k^*) = -s_{k-1}Dg_{k-1}(u_{k-1}^*)P_{k-1}DS_k^m(u_k^*). \tag{24}$$

Hence,

$$\begin{aligned} Dq_p(u_k^*) &= \{I + [I - DB_{k-1}(\tilde{f}_{k-1})Dg_{k-1}(\tilde{u}_{k-1})] + \cdots \\ &+ [I - DB_{k-1}(\tilde{f}_{k-1})Dg_{k-1}(\tilde{u}_{k-1})]^{p-1}\}DB_{k-1}(\tilde{f}_{k-1})Df_{k-1}(u_k^*) \\ &= [I - (I - DB_{k-1}(\tilde{f}_{k-1})Dg_{k-1}(\tilde{u}_{k-1}))^p]Dg_{k-1}(\tilde{u}_{k-1})^{-1}Df_{k-1}(u_k^*) \\ &= -s_{k-1}[I - (I - DB_{k-1}(\tilde{f}_{k-1})Dg_{k-1}(\tilde{u}_{k-1}))^p]Dg_{k-1}(\tilde{u}_{k-1})^{-1}Dg_{k-1}(u_{k-1}^*)P_{k-1}DS_k^m(u_k^*). \end{aligned} \tag{25}$$

Therefore, the equality (19) follows by substituting (25) into (23). □

The schemes (16) with $p = 1$ and 2 are often used in practice. We refer to them as the *V-cycle* and the *W-cycle* methods, respectively. In this paper, we only consider the convergence of the nonlinear *V-cycle* method. The discussion of the other cases is similar.

Setting $p = 1$ in (19), we immediately get a fundamental recursion relation of the *V-cycle*:

$$\begin{aligned} & I - DB_k(\tilde{f}_k)Dg_k(u_k^*) \\ = \ & DS_k^{m_2}(u_k^*)[I - DB_{k-1}(\tilde{f}_{k-1})Dg_{k-1}(u_{k-1}^*)P_{k-1}]DS_k^{m_1}(u_k^*). \end{aligned} \tag{26}$$

From the definition of $b_k(\cdot,\cdot)$, it follows that the inequality $b_k(u,u) \le b_{k-1}(u,u)$ may not hold for some $u \in M_{k-1}$. Thus, operator $I - DB_k(\tilde{f}_k)Dg_k(u_k^*)$ may be negative with respect to the inner product $b_k(\cdot,\cdot)$. To show the convergence of the *V-cycle*, it is sufficient to prove that there exists a constant η_k in $[0,1)$, independent of h_k, such that

$$|b_k([I - DB_k(\tilde{f}_k)Dg_k(u_k^*)]u, u)| \le \eta_k b_k(u,u), \quad \forall u \in M_k. \tag{27}$$

The following two basic assumptions are made to show (27):

$$|b_k((I - P_{k-1})u, u)| \ \le \ C_\beta^2(\frac{\|Dg_k(u_k^*)u\|_k^2}{\lambda_k})^\beta b_k(u,u)^{1-\beta}, \quad \forall u \in M_k, \tag{28}$$

$$\frac{\|Dg_k(u_k^*)u\|_k^2}{\lambda_k} \ \le \ C_S b_k([I - DS_k^1(u_k^*)]u, u), \qquad \forall u \in M_k, \tag{29}$$

where λ_k is the largest eigenvalue of $Dg_k(u_k^*)$, and $0 < \beta < 1$. (28) and (29) are referred to as "the regularity and approximation assumption" and "the smoothing assumption", respectively.

The following theorem provides an estimation for a value of the parameter η_k.

Theorem 2 *Let B_k be defined by Algorithm 1 with $p = 1$ and $m_1 = m_2 = m$. Assume that*

a) Assumptions (28) and (29) hold.

b) The smoothing process S_k^m is formed by m steps of the nonlinear relaxation method S_k, such that $DS_k(u_k^*)$ is symmetric and non-negative with respect to inner product $b_k(\cdot,\cdot)$, and

$$DS_k^m(u_k^*) = [DS_k(u_k^*)]^m.$$

c) The auxiliary vector $\tilde{u}_1 = u_1^*$.
Then there exist two constants, independent of h_k,

$$\eta_{k,1} = \frac{\mathcal{M}(k)}{m^\beta + \mathcal{M}(k)} \quad \textit{and } \eta_{k,2} = 1 - \left(1 + \frac{C_\beta^2 C_S^\beta}{(2m)^\beta}\right)^k,$$

such that

$$\eta_{k,2} b_k(u,u) \le b_k([I - DB_k(\tilde{f}_k)Dg_k(u_k^*)]u,u) \le \eta_{k,1} b_k(u,u), \quad \forall u \in M_k. \tag{30}$$

Furthermore, if m is sufficiently large, then the estimate (27) holds with

$$\eta_k = \max\{|\eta_{k,1}|, |\eta_{k,2}|\} < 1.$$

Here $\mathcal{M}(k)$ is a positive constant related to C_β, C_S, m, β and k. Its detail expression can be found in Theorem 1 of [5].

Proof. With $b_k(DS_k^m(u_k^*)u, v) = b_k(u, DS_k^m(u_k^*)v)$, (26) and the definition of P_{k-1}, we have

$$\begin{aligned}&b_k([I - DB_k(\tilde{f}_k)Dg_k(u_k^*)]u,u) = b_k((I - P_{k-1})DS_k^m(u_k^*)u, DS_k^m(u_k^*)u)\\ &+b_{k-1}([I - DB_{k-1}(\tilde{f}_{k-1})Dg_{k-1}(u_{k-1}^*)]P_{k-1}DS_k^m(u_k^*)u, P_{k-1}DS_k^m(u_k^*)u).\end{aligned}$$

We now show (30) by induction on k. For $k = 1$, we have $B_1 = g_1^{-1}$ and $\tilde{u}_1 = u_1^*$. Thus,

$$|b_1([I - DB_1(\tilde{f}_1)Dg_1(u_1^*)]u,u)| = 0.$$

Suppose (30) holds for $k - 1$. We first prove the right hand side of (30). By induction,

$$\begin{aligned}&b_k([I - DB_k(\tilde{f}_k)Dg_k(u_k^*)]u,u)\\ \le\ & b_k((I - P_{k-1})DS_k^m(u_k^*)u, DS_k^m(u_k^*)u) + \eta_{k-1,1} b_{k-1}(P_{k-1}DS_k^m(u_k^*)u, P_{k-1}DS_k^m(u_k^*)u)\\ =\ & b_k((I - P_{k-1})DS_k^m(u_k^*)u, DS_k^m(u_k^*)u) + \eta_{k-1,1} b_k(P_{k-1}DS_k^m(u_k^*)u, DS_k^m(u_k^*)u)\\ =\ & (1 - \eta_{k-1,1}) b_k((I - P_{k-1})DS_k^m(u_k^*)u, DS_k^m(u_k^*)u) + \eta_{k-1,1} b_k(DS_k^m(u_k^*)u, DS_k^m(u_k^*)u).\end{aligned}$$

By (28), (29) and the generalized arithmetic mean inequality,

$$
\begin{aligned}
& b_k((I-P_{k-1})DS_k^m(u_k^*)u, DS_k^m(u_k^*)u) \\
\le\ & C_\beta^2\Big(\frac{\|Dg_k(u^*)DS_k^m(u_k^*)u\|_k^2}{\lambda_k}\Big)^\beta b_k(DS_k^m(u_k^*)u, DS_k^m(u_k^*)u)^{1-\beta} \\
\le\ & C_\beta^2[\beta r_k\frac{\|Dg_k(u_k^*)DS_k^m(u_k^*)u\|_k^2}{\lambda_k} + (1-\beta)r_k^{-\frac{\beta}{1-\beta}} b_k(DS_k^m(u_k^*)u, DS_k^m(u_k^*)u)] \\
\le\ & C_\beta^2[\beta r_k C_S b_k((I-DS_k(u_k^*))DS_k^{2m}(u_k^*)u,u) + (1-\beta)r_k^{-\frac{\beta}{1-\beta}} b_k(DS_k^m(u_k^*)u, DS_k^m(u_k^*)u)] \\
\le\ & C_\beta^2[\beta r_k\frac{C_S}{2m} b_k((I-DS_k^{2m}(u_k^*))u,u) + (1-\beta)r_k^{-\frac{\beta}{1-\beta}} b_k(DS_k^m(u_k^*)u, DS_k^m(u_k^*)u)].
\end{aligned}
$$

Combining the above inequalities gives

$$
\begin{aligned}
& b_k([I-DB_k(\tilde f_k)Dg_k(u_k^*)]u,u) \\
\le\ & [(1-\eta_{k-1,1})C_\beta^2(1-\beta)r_k^{-\frac{\beta}{1-\beta}} + \eta_{k-1,1}]b_k(DS_k^{2m}(u_k^*)u,u) \\
& +(1-\eta_{k-1,1})C_\beta^2 C_S\frac{\beta}{2m}r_k b_k([I-DS_k^{2m}(u_k^*)]u,u).
\end{aligned}
$$

Now, with the same proof as that in the proof of Theorem 1 of [5], we have that

$$(1-\eta_{k-1,1})C_\beta^2(1-\beta)r_k^{-\frac{\beta}{1-\beta}} + \eta_{k-1,1} \le \eta_{k,1}$$

and

$$(1-\eta_{k-1,1})C_\beta^2 C_S\frac{\beta}{2m}r_k \le \eta_{k,1}.$$

This completes the proof of the right hand side of (30).

We next prove the left hand side of (30). From the spectral properties of $DS_k(u_k^*)$, it follows

$$b_k(DS_k^m(u_k^*)u, DS_k^m(u_k^*)u) \le b_k(u,u),\ k=1,2,\cdots,l. \tag{31}$$

Combining (31) and assumptions (28) and (29) gives

$$
\begin{aligned}
& -b_k((I-P_{k-1})DS_k^m(u_k^*)u, DS_k^m(u_k^*)u) \\
\le\ & \frac{C_\beta^2 C_S^\beta}{(2m)^\beta}\Big[b_k((I-DS_k^{2m}(u_k^*))u,u)\Big]^\beta b_k(DS_k^m(u_k^*)u, DS_k^m(u_k^*)u)^{1-\beta} \\
\le\ & \frac{C_\beta^2 C_S^\beta}{(2m)^\beta}[b_k(u,u) - b_k(DS_k^m(u_k^*)u, DS_k^m(u_k^*)u)]^\beta b_k(u,u)^{1-\beta} \le \frac{C_\beta^2 C_S^\beta}{(2m)^\beta} b_k(u,u),
\end{aligned}
$$

where we have used the following inequality (which is similar to (3.16) in [5]):

$$b_k([I-DS_k(u_k^*))DS_k^{2m}(u_k^*)]u,u) \le \frac{1}{2m}b_k([I-DS_k^{2m}(u_k^*)]u,u).$$

Let $\tau_k = \left(1+\frac{C_\beta^2 C_S^\beta}{(2m)^\beta}\right)^k$. By the induction assumption, we have

$$b_{k-1}([I-DB_{k-1}(\tilde f_{k-1})Dg_{k-1}(u_{k-1}^*)]u,u) > (1-\tau_{k-1})b_{k-1}(u,u),$$

which can be written as

$$-b_{k-1}([I - DB_{k-1}(\tilde{f}_{k-1})Dg_{k-1}(u^*_{k-1})]u, u) < -\eta_{k-1,2}b_{k-1}(u, u).$$

Then, from the above inequalities, we obtain

$$\begin{aligned}
& -b_k([I - DB_k(\tilde{f}_k)Dg_k(u^*_k)]u, u) \\
= \; & -b_k((I - P_{k-1})DS^m_k(u^*_k)u, DS^m_k(u^*_k)u) \\
& -b_{k-1}([I - DB_{k-1}(\tilde{f}_{k-1})Dg_{k-1}(u^*_{k-1})]P_{k-1}DS^m_k(u^*_k)u, P_{k-1}DS^m_k(u^*_k)u) \\
\leq \; & -b_k((I - P_{k-1})DS^m_k(u^*_k)u, DS^m_k(u^*_k)u) - \eta_{k-1,2}b_k(P_{k-1}DS^m_k(u^*_k)u, DS^m_k(u^*_k)u) \\
= \; & -\tau_{k-1}b_k((I - P_{k-1})DS^m_k(u^*_k)u, DS^m_k(u^*_k)u) - \eta_{k-1,2}b_k(DS^m_k(u^*_k)u, DS^m_k(u^*_k)u) \\
\leq \; & \left(\tau_{k-1}\frac{C^2_\beta C^\beta_S}{(2m)^\beta} + \tau_{k-1} - 1\right) b_k(u, u) = (\tau_k - 1)b_k(u, u) = -\eta_{k,2}b_k(u, u).
\end{aligned}$$

The proof of the left hand side of (30) is completed. □

With Theorem 2, we now can obtain a convergence theorem of the nonlinear *V-cycle*.

Theorem 3 *Let $\{u^j_k\}$ be a sequence of iterative values of the nonlinear multigrid* V-cycle *algorithm, and let u^*_k be a solution of equation $g_k(u) = f_k$. If the assumptions in Theorem 2 hold, and m is sufficiently large, then there exists a constant σ_k with $0 < \sigma_k < 1$, independent of grid size h_k, and a neighborhood $O(u^*_k, \epsilon_k)$ of u^*_k, such that all $u^j_k \in O(u^*_k, \epsilon_k)$,*

$$\|u^{j+1}_k - u^*_k\|_{b,k} \leq \sigma_k\|u^j_k - u^*_k\|_{b,k} \quad j = 0, 1, 2, \cdots,$$

when the initial guess $u^0_k \in O(u^*_k, \epsilon_k)$. Here $\|\cdot\|_{b,k}$, the induced norm from $b_k(\cdot, \cdot)$, is defined by $\|u\|^2_{b,k} = b_k(u, u)$.

Proof. Clearly, from Theorem 2 it follows that

$$\|I - DB_k(\tilde{f}_k)Dg_k(u^*_k)\|_{b,k} = \sup_u \frac{|b_k([I - DB_k(\tilde{f}_k)Dg_k(u^*_k)]u, u)|}{b_k(u, u)} \leq \eta_k.$$

For a given positive number δ_k satisfying $\sigma_k = \delta_k + \eta_k < 1$, the differentiability of ψ_k at u^*_k gives that there exists a neighborhood of u^*_k, $O(u^*_k, \epsilon_k) = \{u_k : \|u_k - u^*_k\|_{b,k} \leq \epsilon_k\}$, such that

$$\|\psi_k(u_k) - \psi_k(u^*_k) - D\psi_k(u^*_k)(u_k - u^*_k)\|_{b,k} \leq \delta_k\|u_k - u^*_k\|_{b,k},$$

where $u_k \in O(u^*_k, \epsilon_k)$, ϵ_k is a positive number, and ψ_k is defined in (17). Thus

$$\begin{aligned}
& \|\psi_k(u_k) - u^*_k\|_{b,k} = \|\psi_k(u_k) - \psi_k(u^*_k)\|_{b,k} \\
& \leq \|\psi_k(u_k) - \psi_k(u^*_k) - D\psi_k(u^*_k)(u_k - u^*_k)\|_{b,k} + \|D\psi_k(u^*_k)(u_k - u^*_k)\|_{b,k} \\
& \leq (\delta_k + \|D\psi_k(u^*_k)\|_{b,k})\|u_k - u^*_k\|_{b,k} \leq \sigma_k\|u_k - u^*_k\|_{b,k}.
\end{aligned}$$

Hence, by induction, for any $u^0_k \in O(u^*_k, \epsilon_k)$, we can easily show that $u^j_k \in O(u^*_k, \epsilon_k)$, and

$$\|u^{j+1}_k - u^*_k\|_{b,k} \leq \sigma_k\|u^j_k - u^*_k\|_{b,k} \quad j = 0, 1, 2, \cdots.$$

□

In a nonlinear multigrid algorithm, the following equations have been used on M_k for $k < l$:

$$g_k(v) = \tilde{f}_k + s_k[f_k - g_k(u_k^j)], \tag{32}$$

and

$$g_k(v) = \tilde{f}_k + s_k Q_k[f_{k+1} - g_{k+1}(v_1)], \tag{33}$$

where u_k^j is the j-th iterate of the nonlinear multigrid method, and v_1 is the iterative value after the pre-smoothing step of the nonlinear multigrid algorithm. Hence, to ensure that a nonlinear multigrid algorithm is well-defined, we should show that the solution of either (32) or (33) lies in the neighborhood $O(u_k^*, \epsilon_k)$ given in Theorem 3.

Theorem 4 *Let $O(u_k^*, \epsilon_k)$ be a neighborhood of u_k^*. Assume that*

(a) There exists a constant $\mathcal{C}$ such that for all $u \in M_k$ $\|Dg_k^{-1}(u)\|_{b,k} \leq \mathcal{C}$.

(b) The auxiliary vector $\tilde{u}_k$ satisfies $\tilde{u}_k \in O(u_k^*, \epsilon_k/2)$.

(c) The auxiliary value s_k satisfies $s_k \leq \dfrac{\epsilon_k}{2\mathcal{C}r}$, when $r \neq 0$, otherwise, $s_k = 0$. Here

$$r = \max\{\|f_k - g_k(u_k^j)\|_{b,k}, \|Q_k[f_{k+1} - g_{k+1}(v_1)]\|_{b,k}\}$$

, and v_1 is the iterative value after the pre-smoothing.

Then, the solution of either (32) or (33) lies in the neighborhood $O(u_k^*, \epsilon_k)$.

Proof. We only show that the solution of (32) lies in $O(u_k^*, \epsilon_k)$. The proof for (33) is similar.

Set $r_k = f_k - g_k(u_k^j)$, and $w = g_k^{-1}(\tilde{f}_k + s_k r_k)$. If $r_k = 0$, then $w = \tilde{u}_k \in O(u_k^*, \epsilon_k)$. If $r_k \neq 0$, with assumptions (a) to (c), we have

$$\begin{aligned}
\|w - u_k^*\|_{b,k} &= \|g_k^{-1}(\tilde{f}_k + s_k r_k) - u_k^*\|_{b,k} \\
&\leq \|g_k^{-1}(\tilde{f}_k + s_k r_k) - \tilde{u}_k\|_{b,k} + \|\tilde{u}_k - u_k^*\|_{b,k} \\
&= \|g_k^{-1}(\tilde{f}_k + s_k r_k) - g_k^{-1}(\tilde{f}_k)\|_{b,k} + \|\tilde{u}_k - u_k^*\|_{b,k} \\
&\leq s_k\|Dg_k^{-1}(u)\|_{b,k}\|r_k\|_{b,k} + \|\tilde{u}_k - u_k^*\|_{b,k} \\
&\leq s_k\mathcal{C}\|r_k\|_{b,k} + \|\tilde{u}_k - u_k^*\|_{b,k} \leq \epsilon_k/2 + \epsilon_k/2 = \epsilon_k,
\end{aligned}$$

i.e. $w \in O(u_k^*, \epsilon_k)$. We complete the proof of Theorem 4. □

AN APPLICATION

In this section, as an application of the theory in Section 3, we consider the convergence of the nonlinear *V-cycle* for solving the second order elliptic, mildly nonlinear boundary value problem

$$\begin{cases} -\nabla(\alpha \nabla u) + B(x,u) &= f(x), \quad \text{in } \Omega, \\ \qquad\qquad\qquad\qquad u &= 0, \qquad\quad \text{on } \partial\Omega, \end{cases} \tag{34}$$

where Ω is a bounded, Lipschitz, polyhedral domain in R^d, $\alpha \in W^{1,\infty}(\Omega)$, $\alpha \geq C_\alpha > 0$ a.e. on Ω, and $f \in L^2(\Omega)$.

Let D_2B denote the derivative of $B(\cdot,\cdot)$ with respect to the second variable. We make the following assumptions on D_2B in this section.

A2) $D_2B(x,u)$ is continuous in $\bar{\Omega}\times R$, and there exist constants C_1 and C_2 such that

$$0 < C_2 \le D_2B(x,u) \le C_1.$$

A3) $D_2B(x,u)$ satisfies a Lipschitz condition: there exists a constant L, independent of u and v, such that

$$|D_2B(x,u) - D_2B(x,v)| \le L|u-v|, \tag{35}$$

for all (x,u), (x,v) on a subset of $\bar{\Omega}\times R$.

Let $H = H_0^1(\Omega)$ be the Sobolev space [2]. The weak form of (34) is thus: Find $u \in H$, such that

$$a(u,v) = (f,v)_{L^2}, \qquad \forall v \in H \tag{36}$$

where

$$a(u,v) = \int_\Omega [\alpha \bigtriangledown u \bigtriangledown v + B(x,u)v]dx, \text{ and } (f,v)_{L^2} = \int_\Omega f(x)v(x)dx. \tag{37}$$

Let M_k be a set of piecewise linear functions with respect to a quasi-uniform triangulation $\mathcal{F}_k$ on Ω of size h_k in the usual sense [8]. We assume that there is a constant c , independent of k, such that $h_{k-1} \le ch_k$, and these triangulations should be nested in the sense that any triangle in $\mathcal{F}_{k-1}$ can be written as a union of triangles of $\mathcal{F}_k$.

The finite element discretization for (36) on each M_k is as follows: *Find* $u_k \in M_k$ *such that*

$$a(u_k,v) = (f,v)_{L^2}, \qquad \forall v \in M_k \ , \tag{38}$$

where $k = 1,2,\cdots,l$.

Based on Theorem 39.12 in [16], we assume that

A4) Equations (36) and (38) have unique solutions u^* and $\hat{u}_k$, respectively. For $u^* \in H^{1+\beta}(\Omega)$ with $\beta \in (0,1]$, there exists a constant c, independent of h_k, such that

$$\|u^* - \hat{u}_k\|_1 \le ch_k^\beta, \tag{39}$$

where $k = 1,2,\cdots,l$, and $\|\cdot\|_1$ is the usual norm in Sobolev space H^1 [2].

We solve equation (38) by the nonlinear multigrid *V-cycle* scheme with the smoother S_k^m defined by m steps of the damped-Jacobi-Newton iteration. To prove its convergence, using Theorem 3, we only need to verify Assumptions (29) and (28).

We first prove Assumption (29) for the smoother S_k^m below.

Let $\{\varphi_i\}_{i=1}^{n_k}$ be a natural nodal basis for M_k, where $n_k = dimM_k$. Apparently, we may consider the following equation on M_k: For $f_k \in M_k$, find $u_k \in M_k$ such that

$$(g_k(u_k),\varphi_\nu)_k = (f_k,\varphi_\nu)_k, \quad \nu = 1,2,\cdots,n_k,$$

with g_k being defined by

$$(g_k(u_k),v)_k = a(u_k,v) - (f,v)_{L^2}, \qquad \forall v \in M_k. \tag{40}$$

Let u_k^j be the j-th iterate of the damped-Jacobi-Newton iteration using a damping parameter θ, expressed as follows:

$$u_k^{j+1} = u_k^j + R_k(u_k^j)[f_k - g_k(u_k^j)],$$

where the linear operator $R_k(u) : M_k \to M_k$ is defined by

$$R_k(u)v = \theta \sum_{i=1}^{n_k} \left(\frac{\partial g_k(u)}{\partial u_{k,i}} \varphi_i, \varphi_i \right)_k^{-1} (v, \varphi_i)_k \varphi_i \quad \forall v \in M_k.$$

Since $S_k^1(u) = u + R_k(u)(f_k - g_k(u))$, and

$$DS_k^1(u_k^*) = I - R_k(u_k^*) Dg_k(u_k^*), \tag{41}$$

we have

$$DS_k^m(u_k^*) = [I - R_k(u_k^*) Dg_k(u_k^*)]^m = [DS_k^1(u_k^*)]^m.$$

Clearly, $DS_k^1(u_k^*)$ is symmetric, so Assumption b) of Theorem 2 holds. From (41) we see that the Jacobi-Newton iteration has a similar form as the damped-Jacobi method in [17]. Therefore, using the same argument as in [17], we can show that Assumption (29) is satisfied by the damped-Jacobi-Newton iteration.

We next verify Assumption (28). Let g be defined by

$$(g(u), v) = a(u, v) - (f, v)_{L^2}, \qquad \forall v \in H. \tag{42}$$

It is easy to show that $Dg(w)$, defined by

$$(Dg(w)u, v) = \int_\Omega [\alpha \bigtriangledown u \bigtriangledown v + D_2 B(x, w) uv] dx, \quad \forall v \in H,$$

is symmetric, positive definite on H.

Hence, from (40) it follows that $Dg_k(w)$ is a symmetric, positive definite operator on M_k. Thus, the bilinear form on $M_k \times M_k$

$$b_k(u, v) \equiv (Dg_k(w)u, v)_k, \quad \forall u, v \in M_k, \tag{43}$$

is symmetric, positive definite.

For simplicity, we let $A_k \equiv Dg_k(u_k^*)$, and define a family of norms as follows:

$$|||v|||_{r,k}^2 = (A_k^r v, v)_k, \qquad \forall v \in M_k,$$

where r is a positive number. In addition, we note that $|||v|||_{0,k}$ is equivalent to $||v||_{L^2}$ and $|||v|||_{1,k} = ||v||_{b,k}$.

We now can show that Assumption (28) holds in the following theorem. The proof of this theorem can be found in [15].

Theorem 5 *Let M_k be the space of continuous piecewise linear functions with respect to a quasi-uniform triangulation, and let u_k^* be the solution of equation $g_k(u) = f_k$ in M_k. Assume that (A1) to (A4) hold, and that the solution U of the variational problem*

$$b_k(U, v) = (F, v)_{L^2}, \qquad \forall v \in H \tag{44}$$

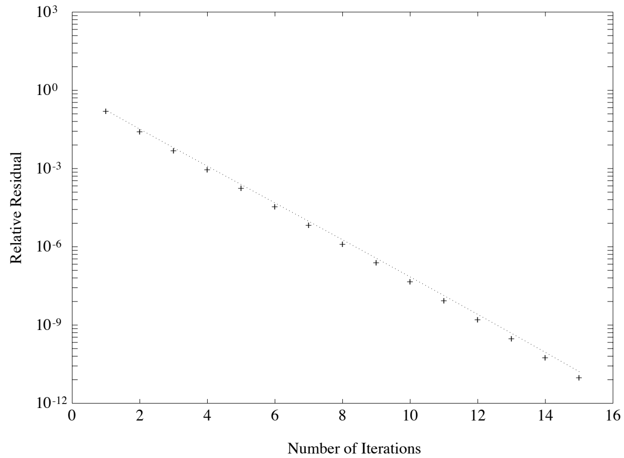

Figure 1: A comparison of a nonlinear *V-cycle* and a linear *V-cycle*. Here $\cdots$: the linear *V-cycle* method for solving (46) with $b = 0$, +++: the nonlinear *V-cycle* method for solving (46) with $a = b = 1$, and $h = \frac{1}{128}$.

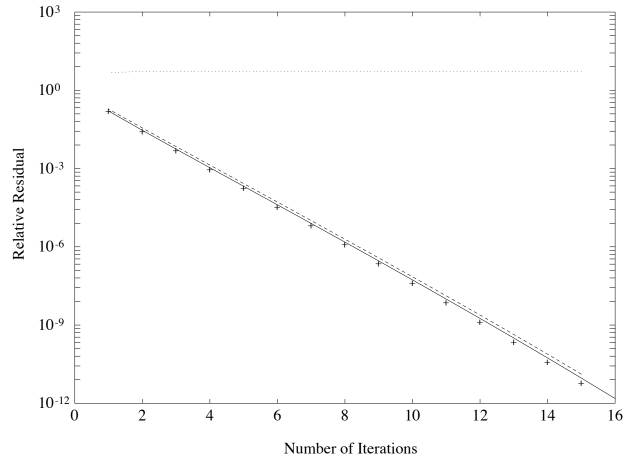

Figure 2: Dependency of the convergence rate of the nonlinear *V-cycle* on the auxiliary vector. Here $+++ : \tilde{u}_k = 0$, —: $\tilde{u}_k = Q_k u_{k+1}^{j,s}$, $---$: $\tilde{u}_k = S_k^{200}(0)$, $\cdots$: $\tilde{u}_k = 0.5$, $h = \frac{1}{128}$, and $a = b = 1$ in (46).

is in $H^{1+\beta}(\Omega)$ *for some* $\beta \in (0, 1]$, *and satisfies*

$$\|U\|_{H^{1+\beta}} \leq \mathcal{C}\|F\|_{H^{\beta-1}} \tag{45}$$

for some positive constant $\mathcal{C}$, *independent of* F. *Then, there exists a constant* C *such that*

$$|b_k((I - P_{k-1})u, u)| \leq C \left(\frac{\|Dg_k(u_k^*)u\|_k^2}{\lambda_k} \right)^{\frac{\beta}{2}} b_k(u, u)^{1-\frac{\beta}{2}}, \quad \forall u \in M_k,$$

where λ_k *is the largest eigenvalue of* $Dg_k(u_k^*)$.

NUMERICAL EXPERIMENTS

In this section, we present numerical experiments with the nonlinear multigrid method for solving the following model problem [10]:

$$\begin{cases} -(u_{xx} + u_{yy}) + b \sinh(au) & = f \text{ in } \Omega = (0,1) \times (0,1), \\ u & = 0 \text{ on } \partial\Omega, \end{cases} \tag{46}$$

where a and b are positive numbers. The right hand side term f of (46) is chosen such that $u = \sin \pi x \sin \pi y$ is the solution.

The discretization equation of (46) is defined by the five-point stencil with $h_k = 1/2^k$ $(1 \leq k \leq l)$. The smoothing process S_k^m consists of m steps of the Gauss-Seidel-Newton iteration.

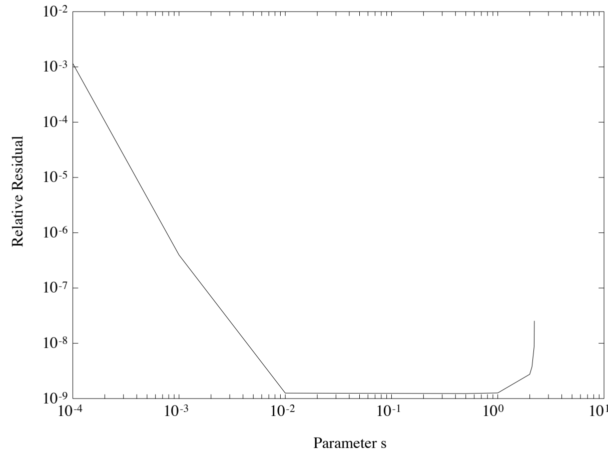

Figure 3: The relation of the relative residual of the nonlinear *V-cycle* with parameter s_k at the $12th$ *V-cycle* iteration. This figure shows that as s_k is around 1, the nonlinear *V-cycle* has an almost same convergent rate. Here $h = \frac{1}{64}$, and $a = b = 1$ in (46).

We set $m_1 = m_2 = m$ for all grid levels and the coarsest grid size $h_1 = \frac{1}{2}$ for all of our numerical examples. Besides, the full-weighting restriction operator Q_k, [9], was used, and only one step of the Gauss-Seidel-Newton iteration was applied to get the solution of the equation on the coarsest grid M_1. The initial guess $u_h^0 = 0$ and the relative residual stopping criterion were taken for all the numerical experiments, which were implemented on a $KSR1$ supercomputer with single precision, which is equal to the regular double precision.

We compared the performance of the nonlinear *V-cycle* with the linear *V-cycle* method. The linear *V-cycle* case was obtained from the nonlinear *V-cycle* program by setting $b = 0$ in (46). Thus, a Poisson equation was solved by the linear *V-cycle* method. From Figure 1 we see that the nonlinear multigrid method is as efficient as the linear multigrid method. We checked the dependency of the convergence rate of the nonlinear multigrid method on its two parameters $\tilde{u}$ and s_k. We used three different values of $\tilde{u}_k$ in the experiments.

1) $\tilde{u}_k = 0$ on all grid levels;

2) $\tilde{u}_k = S_k^m(0)$, i.e. $\tilde{u}_k$ is defined by m steps of the Gauss-Seidel-Newton iteration with zero initial guess. Clearly, by increasing m, we can make $\tilde{u}_k$ approach to the exact solution $g_k(u) = f_k$ as closely as desired.

3) $\tilde{u}_k = Q_k u_{k+1}^{j,s}$, where $u_{k+1}^{j,s}$ denotes the iterative value after the pre-smoothing step of the *V-cycle*. We call this type of $\tilde{u}_k$ Brandt's choice because it was first used by Brandt in [7]. Figure 2 shows that if $\tilde{u}_k$ is properly close to the solution of $g_k = f_k$, the convergence rate of the *V-cycle* will be almost the same. Otherwise, the nonlinear *V-cycle* may be divergent. For example, from this figure we see that the *V-cycle* with $\tilde{u}_k = 0.5$ was divergent.

For fixed $\tilde{u}_k = 0$, we also made experiments with different values of s_k. Figure 3 shows that it is satisfactory to let s_k be around 1.

Finally, we checked the influence of the a and b in (46) on the convergence of the nonlinear *V-cycle* method. The numerical results are reported in Tables 1 to 3. Here we used four different $\tilde{u}_k$, $h = \frac{1}{64}$ and $m_1 = m_2 = 1$ for all of these numerical experiments. We also used $a = 1.0$,

$b = 1.0$ and $a = 3.0$ in Table 1, Table 2 and Table 3, respectively. The notation — in the tables means that the *V-cycle* is divergent. From these tables we see that: 1) When $0 \le a < 3$ and $0 \le b < 10$, $\tilde{u}_k = 0$ is the simplest choice; 2) Brandt's choice worked for $0 \le a \le 6$ and $0 \le b \le 100$; and 3) the nonlinear *V-cycle* with $\tilde{u}_k = S_k^m(0)$ using large m can lead to convergence for a pair of a and b for which the nonlinear *V-cycle* with Brandt's choice is divergent.

Table 1: *The performance of the nonlinear* V-cycle *as the b in (46) becomes larger.*

	The Total number of Iterations			
b	$\tilde{u}_k = 0$	$\tilde{u}_k = Q_k u_{k+1}^{j,s}$	$\tilde{u}_k = S_k^1(0)$	$\tilde{u}_k = S_k^{10}(0)$
10	13	14	13	14
30	40	13	14	13
100	—	12	35	13

Table 2: *The performance of the nonlinear* V-cycle *as the a in (46) becomes larger.*

	The number of Iterations			
a	$\tilde{u}_k = 0$	$\tilde{u}_k = Q_k u_{k+1}^{j,s}$	$\tilde{u}_k = S_k^1(0)$	$\tilde{u}_k = S_k^{10}(0)$
0.001	14	14	14	14
2.0	13	14	14	14
3.0	32	14	14	15
6.0	—	12	—	30
7.0	—	—	—	20

Table 3: *The performance of the nonlinear* V-cycle *for solving (46) with large a and b.*

	The number of Iterations			
b	$\tilde{u}_k = 0$	$\tilde{u}_k = Q_k u_{k+1}^{j,s}$	$\tilde{u}_k = S_k^1(0)$	$\tilde{u}_k = S_k^{10}(0)$
0.01	14	14	14	14
1.0	32	14	14	15
20.0	—	12	—	16

ACKNOWLEDGMENTS

The author would like to thank his advisor Professor L. Ridgway Scott for valuable discussions and his continuous support. He also wants to thank Professor Olof Widlund and Professor Shuzi Zhou for their helpful comments and valuable suggestions.

REFERENCES

[1] Ralph. Abraham, *Manifolds, Tensor Analysis, and Applications.* Addison-Wesley Publishing Company, Inc. , 1983.

[2] R. A. ADAMS: *Sobolev Spaces,* Academic Press, New York, 1975.

[3] R. E. BANK, AND C. C. DOUGLAS: *Sharp estimates for multigrid rates of convergence with general smoothing and acceleration.* SIAM J. Numer. Anal. 22 (1985), 617-633.

[4] R. E. BANK, D. J. ROSE,: *Analysis of a Multilevel Iterative Method for Nonlinear Finite Element Equations,* Math. Comp. 39 (1982), 453-465.

[5] J. H. BRAMBLE AND J. E. PASCIAK: *New Convergence Estimates for Multigrid Algorithms.* Math. Comp. 49 (1987), No.180.

[6] J. H. BRAMBLE AND J. E. PASCIAK: *New Estimates for Multilevel Algorithms Including the V-cycle,* Math. Comp. 60 (1993), 447-471.

[7] A. BRANDT: *Guide in Multigrid Development.* Lect. Notes in Math., 960 (1982), Springer.

[8] P. C. CIARLET: *The Finite Element Methods for Elliptic Problems,* North-Holland, Amsterdam, New York, Oxford, 1978.

[9] W. HACKBUSCH: *Multigrid Methods and Applications.* Springer Heidelberg, 1985.

[10] M. HOLST AND F. SAIED: *Multigrid solution of the Poisson-Boltzmann equation,* Journal of Computational Chemistry, 14 (1993), 105-113.

[11] ARNOLD REUSKEN: *Convergence of the Multigrid Full Approximation Scheme for a Class of Elliptic Mildly Nonlinear Boundary Value Problems,* Numer. Math. 52 (1988), 251-277.

[12] ARNOLD REUSKEN: *Convergence of the Multilevel Full Approximation Scheme Including the* V-cycle*,* Numer. Math. 53 (1988), 663-686.

[13] ORTEGA J. M. AND W. C. RHEINHOLDT: *Iterative Solution of Nonlinear Equations in Several Variables,* Academic press, New York, 1970.

[14] L. R. SCOTT AND DEXUAN XIE, *Analysis of nonlinear multigrid methods with numerical integration*, to be submitted to Math. Comp., 1995.

[15] DEXUAN XIE, *New nonlinear multigrid analysis*, Chapter 6 of his Ph.D. Thesis, University of Houston, 1995.

[16] A. ŽENÍŠEK: *Nonlinear Elliptic and Evolution Problems and Their Finite Element Approximations,* Academic Press, 1990.

[17] JINCHAO XU: *Iterative Methods by Space Decomposition and Subspace Correction,* SIAM Review, Dec. 1992.