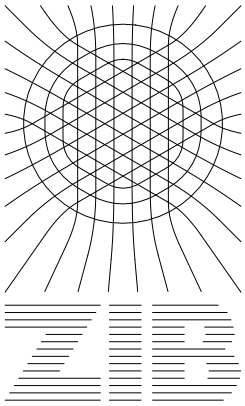
Konrad-Zuse-Zentrum für Informationstechnik Berlin



Ralf Kornhuber*

On Robust Multigrid Methods for Non–Smooth Variational Problems

* Mathematisches Institut A, Universität Stuttgart
Pfaffenwaldring 57, D–70569 Stuttgart, Fed. Rep. of Germany
e-mail: kornhuber@mathematik.uni-stuttgart.de

Preprint SC 96–38 (Dezember 1996)

Ralf Kornhuber

On Robust Multigrid Methods for Non–Smooth Variational Problems

Abstract We consider the fast solution of large, piecewise smooth minimization problems as resulting from the approximation of elliptic free boundary problems. The most delicate question in constructing a multigrid method for a nonlinear, non–smooth problem is how to represent the nonlinearity on the coarse grids. This process usually involves some kind of linearization. The basic idea of monotone multigrid methods to be presented here is first to select a neighborhood of the actual smoothed iterate in which a linearization is possible and then to constrain the coarse grid correction to this neighborhood. Such a local linearization allows to control the local corrections at each coarse grid node in such a way that the energy functional is monotonically decreasing. This approach leads to globally convergent schemes which are robust with respect to local singularities of the given problem. The numerical performance is illustrated by approximating the well-known Barenblatt solution of the porous medium equation.

Key words: free boundary problems, finite element methods, multigrid methods

AMS (MOS) subject classifications: 65N30, 65N55, 35J85

1 Introduction

Let Ω be a polygonal domain in the Euclidean space $\mathbb{R}^2$. We will consider the minimization problem

$$u \in H: \quad \mathcal{J}(u) + \phi(u) \leq \mathcal{J}(v) + \phi(v), \quad \forall v \in H, \tag{1.1}$$

on a closed subspace $H \subset H^1(\Omega)$. For simplicity, we select $H = H_0^1(\Omega)$. Other boundary conditions of Neumann or mixed type and the case of three space dimensions can be treated in a similar way [6, 7]. The quadratic functional $\mathcal{J}$,

$$\mathcal{J}(v) = \tfrac{1}{2} a(v, v) - \ell(v), \tag{1.2}$$

is induced by a continuous, symmetric, and H–elliptic bilinear form $a(\cdot, \cdot)$ and a bounded, linear functional ℓ. H is equipped with the energy norm $\| \cdot \| = a(\cdot, \cdot)^{1/2}$. The functional ϕ,

$$\phi(v) = \int_\Omega \Phi(v(x))\, dx, \tag{1.3}$$

is generated by a *piecewise smooth* function $\Phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$. If Φ is convex, takes finite values on some closed interval and satisfies certain growth conditions [18], then the functional ϕ is convex, lower semi–continuous and proper (i.e. $\phi(v) > -\infty$ and $\phi \not\equiv +\infty$). For such functionals ϕ, the minimization problem (1.1) has a unique solution $u \in H$ and can be equivalently rewritten as the elliptic variational inequality of the second kind

$$u \in H: \quad a(u, v - u) + \phi(v) - \phi(u) \geq \ell(v - u)\,, \quad \forall v \in H. \tag{1.4}$$

To fix the ideas, we will concentrate on the special case

$$\Phi(z) = \tfrac{m}{m+1} z^{1+\frac{1}{m}},\ z \geq 0, \quad \Phi(z) = +\infty,\ z < 0, \tag{1.5}$$

arising from an implicit time discretization of the porous medium equation as explained below.

Let $\mathcal{T}_j$ be a regular partition of Ω in triangles with minimal diameter of order 2^{-j}. The interior nodes and edges of $\mathcal{T}_j$ are denoted by $\mathcal{N}_j$ and $\mathcal{E}_j$, respectively. Discretizing H by continuous, piecewise linear finite elements $\mathcal{S}_j$, we obtain the *discrete minimization problem*

$$u_j \in \mathcal{S}_j: \quad \mathcal{J}(u_j) + \phi_j(u_j) \leq \mathcal{J}(v) + \phi_j(v), \quad \forall v \in \mathcal{S}_j. \tag{1.6}$$

Observe that the functional ϕ is approximated by its $\mathcal{S}_j$–interpolate ϕ_j,

$$\phi_j(v) = \sum_{p \in \mathcal{N}_j} \Phi(v(p)) \int_\Omega \lambda_p^{(j)}(x)\, dx, \quad v \in \mathcal{S}_j. \tag{1.7}$$

It is easily seen that the discrete functional ϕ_j is still convex, lower semi–continuous, and proper. Hence, the discrete problem (1.6) has a unique solution $u_j \in \mathcal{S}_j$ which is characterized by the variational inequality

$$u_j \in \mathcal{S}_j : \quad a(u_j, v - u_j) + \phi_j(v) - \phi_j(u_j) \geq \ell(v - u_j), \quad \forall v \in \mathcal{S}_j. \quad (1.8)$$

The convergence of the discretization (1.6) is a consequence of general results as condensed e.g. by Glowinski [10].

The main issue of this paper is the fast solution of the discrete minimization problem (1.6). A meanwhile classical way to construct a multigrid method for nonlinear variational problems is to use a Newton–type linearization as an outer iteration and then to replace the solution of the resulting linear system by a number of multigrid steps. Global convergence of such *Newton–multigrid* methods can be obtained by a suitable damping of the multigrid correction, provided that the nonlinearity is smooth enough. We refer to Bank and Rose [1] and to recent results of Deuflhard and Weiser [8] in these proceedings. In contrast to Newton–multigrid, *nonlinear multigrid* methods are based on a hierarchy of nonlinear problems (see e.g. Hackbusch [11]). Introducing a suitable damping of the corrections from each refinement level, Hackbusch and Reusken [12] proved global convergence in sufficiently smooth cases.

However, if the above Newton–multigrid or nonlinear multigrid methods are applied to problems of the form (1.8) with a strongly varying or even non–differentiable nonlinearity ϕ_j, then the damping parameters may become very small or even zero so that the resulting convergence rates are unacceptable. A common remedy is to replace the given non-smooth problem by a more regular one. Unfortunately, such *regularization* leads to a coupling of the discretization error with the convergence of the iterative solver. High convergence speed may have to be paid with low accuracy.

In this paper, we describe a *local damping strategy* that reflects the local singularities of the subdifferential $\partial\phi_j$ but does not involve any kind of regularization. The resulting algorithms can be regarded as an extension of recent *monotone multigrid methods* [16, 17, 18] from piecewise quadratic functionals ϕ_j to the piecewise smooth case. The basic idea is first to find out a neighborhood of the actual smoothed iterate in which a linearization of $\partial\phi_j$ can be controlled by pointwise Lipschitz constants and then to constrain the coarse grid correction to this neighborhood. Using suitable restrictions of the Lipschitz constants, we are able to compute local damping parameters for all local corrections at all different coarse grid nodes on all levels. In this way, our approach combines maximal flexibility of the coarse grid correction with global convergence.

We will present a standard and a truncated variant of the method and state their basic properties. Detailed convergence proofs featuring asymptotic

bounds for the convergence rates and a more sophisticated fine grid smoothing will be published elsewhere [19]. In order to illustrate the numerical properties of our monotone multigrid algorithms, we approximate the well-known Barenblatt solution of the porous medium equation. Using nested iteration, we observed a similar efficiency as in the linear selfadjoint case. Moreover, our experiments indicate a considerable robustness with respect to the smoothness of the nonlinearity ϕ.

2 A Standard Monotone Multigrid Method with Local Damping

Assume that $\mathcal{T}_j$ is resulting from j refinements of a given, intentionally coarse triangulation $\mathcal{T}_0$ of Ω. Though all algorithms and convergence results to be presented in this paper can easily be extended to locally refined grids, we assume for simplicity that the triangulations are uniformly refined. More precisely, each triangle $t \in \mathcal{T}_k$ is subdivided into four congruent subtriangles to obtain the next triangulation $\mathcal{T}_{k+1}$.

In this way, we obtain a sequence of triangulations $\mathcal{T}_0, \ldots, \mathcal{T}_j$ and of corresponding nested finite element spaces $\mathcal{S}_0 \subset \ldots \subset \mathcal{S}_j$. Each of the spaces $\mathcal{S}_k$ is spanned by the nodal basis

$$\Lambda_k = \{\lambda_p^{(k)} \mid p \in \mathcal{N}_j\}, \quad k = 0, \ldots, j,$$

with the n_k basis functions $\lambda_p^{(k)} \in \mathcal{S}_k$ defined by $\lambda_p^{(k)}(q) = \delta_{pq}$, $\forall p, q \in \mathcal{N}_k$, (Kronecker delta). Collecting the nodal basis functions from all refinement levels, we define the multilevel nodal basis $\Lambda_{\mathcal{S}}$,

$$\Lambda_{\mathcal{S}} = \left(\lambda_{p_1}^{(j)}, \ldots, \lambda_{p_{n_j}}^{(j)}, \lambda_{p_1}^{(j-1)}, \ldots, \lambda_{p_{n_{j-1}}}^{(j-1)}, \ldots, \lambda_{p_1}^{(0)}, \ldots, \lambda_{p_{n_0}}^{(0)}\right)$$

which is ordered from fine to coarse. We frequently write $\Lambda_{\mathcal{S}} = (\lambda_1, \ldots, \lambda_m)$ with $m = n_j + \ldots + n_0$.

In the special case of an elliptic selfadjoint problem (i.e. $\phi \equiv 0$) one step of a classical multigrid V–cycle with Gauß–Seidel smoother can be regarded as the successive minimization of the energy functional $\mathcal{J}$ in the direction of the multilevel nodal basis functions $\lambda_l \in \Lambda_{\mathcal{S}}$ (cf. e.g. McCormick [21], Xu [23], or Yserentant [24]). We will use a straightforward extension of this *multilevel relaxation* as the starting point for the construction of monotone multigrid methods for the non–smooth optimization problem (1.6). For this reason, we introduce the splitting

$$\mathcal{S}_j = \sum_{l=1}^{m} V_l, \tag{2.1}$$

of $\mathcal{S}_j$ in the one–dimensional subspaces $V_l = \operatorname{span}\{\lambda_l\}$, $l = 1, \ldots, m$. Then, for a given ν–th iterate $u_j^\nu \in \mathcal{S}_j$, we set $w_0 = u_j^\nu$ and compute a sequence of intermediate iterates w_l from the m local subproblems

$$\begin{aligned} \bar{v}_l \in V_l: \quad & \mathcal{J}(w_{l-1} + \bar{v}_l) + \phi_j(w_{l-1} + \bar{v}_l) \leq \\ & \leq \mathcal{J}(w_{l-1} + v) + \phi_j(w_{l-1} + v), \quad \forall v \in V_l, \end{aligned} \tag{2.2}$$

setting $w_l = w_{l-1} + \bar{v}_l$. Finally, the next iterate is given by $u_j^{\nu+1} = w_m$.

The leading *fine grid corrections* in the direction of $\lambda_l \in \Lambda_j$ can be regarded as one step of a *nonlinear* Gauß–Seidel relaxation which is well–known to be globally convergent [10]. Observe that the resulting *smoothed iterate* $\bar{u}_j^\nu := w_{n_j}$ satisfies $\bar{u}_j^\nu(p) \geq 0$, $\forall p \in \mathcal{N}_j$, or, equivalently, $\phi_j(\bar{u}_j^\nu) < \infty$ for all $\nu \geq 1$ and any initial iterate $u_j^0 \in \mathcal{S}_j$. The subsequent *coarse grid corrections* of the smoothed iterate $\bar{u}_j^\nu$ in the directions $\lambda_l \in \Lambda_{\mathcal{S}} \setminus \Lambda_j$ are intended to reduce the low frequency components of the error.

Let us now consider a more general iteration of the form

$$u_j^{\nu+1} = u_j^\nu + \sum_{l=1}^{m} v_l \tag{2.3}$$

where the corrections $v_l \in V_l$ may be regarded as approximate solutions of the local problems (2.2). A proof of the following convergence result can be taken almost literally from [18].

Theorem 2.1 *Assume that the local fine grid corrections are the unique solutions of (2.2),*

$$v_l = \bar{v}_l, \quad l = 1, \ldots, n_j, \tag{2.4}$$

and that the local coarse grid corrections satisfy the monotonicity condition

$$\mathcal{J}(w_{l-1}+v_l)+\phi_j(w_{l-1}+v_l) \leq \mathcal{J}(w_{l-1})+\phi_j(w_{l-1}), \quad l = n_j+1, \ldots, m. \tag{2.5}$$

Then, for any initial iterate $u_j^0 \in \mathcal{S}_j$, the sequence of iterates $(u_j^\nu)_{\nu\geq 0}$ defined in (2.3) converges to the solution u_j of the discrete problem (1.6).

In the special case $\phi_j \equiv 0$ the computation of all the optimal corrections $\bar{v}_l$ can be implemented as a V–cycle: Representing the bilinear form $a(\cdot,\cdot)$ on the coarse grid spaces $\mathcal{S}_k$ by their values on Λ_k, one can update the residual and evaluate the local corrections without visiting the fine grid. This provides optimal numerical complexity, i.e. $\mathcal{O}(n_j)$ operations, for each iteration step.

For the nonlinearity defined in (1.5), the *fine grid corrections* in the direction of some fixed $\lambda_l \in \Lambda_j$ can be computed iteratively by a simple bisection method. A more sophisticated analysis shows that condition (2.4) can be relaxed in such a way that the exact solution of (2.2) can be replaced by one step of a suitable descent method. Details will be contained in a forthcoming paper [19].

From now on, we assume that the smoothed iterate $\bar{u}_j^\nu$ is available. Let us consider the *coarse grid correction* in the direction of some $\lambda_l \in \Lambda_k$, $k < j$. Using subdifferential calculus [9], the local correction $\bar{v}_l = \bar{z}_l\lambda_l \in V_l$ can be equivalently computed from the following *scalar inclusion* for the unknown coefficient $\bar{z}_l \in \mathbb{R}$

$$0 \in a(\lambda_l, \lambda_l)\bar{z}_l - (\ell(\lambda_l) - a(w_{l-1}, \lambda_l)) + \partial\phi_j(w_{l-1} + \bar{z}_l\lambda_l)(\lambda_l). \tag{2.6}$$

For all $w \in \mathcal{S}_j$ with non–negative values and all $v \in \mathcal{S}_j$ the subdifferential $\partial\phi_j$ takes the form

$$\partial\phi_j(w)(v) = \sum_{p\in\mathcal{N}_j} \partial\Phi(w(p))v(p) \int_\Omega \lambda_p^{(j)}(x)\, dx$$

where the subdifferential $\partial\Phi$ of the scalar function Φ (cf.(1.5)) is given by

$$\partial\Phi(z) = \sqrt[m]{z}, \quad z > 0, \qquad \partial\Phi(0) = [0, -\infty).$$

It is clear that the solution of (2.6) requires (at least) one evaluation of $\partial\phi_j(w_{l-1} + z\lambda_l)(\lambda_l)$ which in turn leads to (at least) one additional prolongation, because $\partial\phi_j$ is non–linear. As a consequence, the number of operations for one complete step of the nonlinear multilevel relaxation grows (at least) like $\mathcal{O}(n_j \log(n_j))$.

This motivates the *local linearization* of the subproblems (2.6) in a neighborhood of the smoothed iterate $\bar{u}_j^\nu$. Let us first introduce the *regular nodes*

$$\mathcal{N}_j^\circ(\bar{u}_j^\nu) = \{p \in \mathcal{N}_j \mid \bar{u}_j^\nu(p) > 0\} \subset \mathcal{N}_j. \tag{2.7}$$

Φ is differentiable in $\bar{u}_j^\nu(p)$, if and only if $p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu)$. Using Taylor's expansion

$$\Phi'(w(p)) \approx \Phi'(\bar{u}_j^\nu(p)) + \Phi''(\bar{u}_j^\nu(p))(w(p) - \bar{u}_j^\nu(p)),$$

we introduce the formal linearization

$$\partial\phi_j(w)(v) \approx b_{\bar{u}_j^\nu}(w, v) - f_{\bar{u}_j^\nu}(v) \tag{2.8}$$

involving the symmetric positive semidefinite bilinear form $b_{\bar{u}_j^\nu}(\cdot,\cdot)$,

$$b_{\bar{u}_j^\nu}(w, v) = \sum_{p\in\mathcal{N}_j^\circ(\bar{u}_j^\nu)} \Phi''(\bar{u}_j^\nu(p))w(p)v(p) \int_\Omega \lambda_p^{(j)}(x)\, dx, \tag{2.9}$$

and the linear functional $f_{\bar{u}_j^\nu}$,

$$f_{\bar{u}_j^\nu}(v) = -\sum_{p\in\mathcal{N}_j^\circ(\bar{u}_j^\nu)} \Big(\Phi'(\bar{u}_j^\nu(p)) - \Phi''(\bar{u}_j^\nu(p))\bar{u}_j^\nu(p)\Big)v(p) \int_\Omega \lambda_p^{(j)}(x)\, dx. \tag{2.10}$$

Observe that no contribution is taken from the remaining *critical nodes*

$$\mathcal{N}_j^\bullet(\bar{u}_j^\nu) = \mathcal{N}_j \setminus \mathcal{N}_j^\circ(\bar{u}_j^\nu) \tag{2.11}$$

where the subdifferential $\partial\Phi$ is set–valued. Of course, the linearization (2.8) can only make sense in a suitable neighborhood $\mathcal{K}_{\bar{u}_j^\nu} \subset \mathcal{S}_j$,

$$\mathcal{K}_{\bar{u}_j^\nu} = \{w \in \mathcal{S}_j \mid \underline{\varphi}_{\bar{u}_j^\nu}(p) \le w(p) \le \overline{\varphi}_{\bar{u}_j^\nu}(p),\ p \in \mathcal{N}_j\}.$$

The obstacles $\underline{\varphi}_{\bar{u}_j^\nu}, \overline{\varphi}_{\bar{u}_j^\nu} \in \mathcal{S}_j$ are selected in such a way that the *pointwise Lipschitz conditions*

$$|\Phi''(z_1) - \Phi''(z_2)| \leq L_p|z_1 - z_2|, \quad \forall z_1, z_2 \in [\underline{\varphi}_{\bar{u}_j^\nu}(p), \overline{\varphi}_{\bar{u}_j^\nu}(p)], \tag{2.12}$$

hold for all $p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu)$ while

$$\underline{\varphi}_{\bar{u}_j^\nu}(p) = \overline{\varphi}_{\bar{u}_j^\nu}(p) = 0 \tag{2.13}$$

is set at the critical nodes $p \in \mathcal{N}_j^\bullet(\bar{u}_j^\nu)$. In order to control the linearized coarse grid correction by *local Lipschitz constants*, we will require $w_l \in \mathcal{K}_{\bar{u}_j^\nu}$ for all $l = n_j + 1, \ldots, m$. As a consequence, the values $\bar{u}_j^\nu(p)$ at the critial nodes $p \in \mathcal{N}_j^\bullet(\bar{u}_j^\nu)$ remain invariant under coarse grid correction.

Now, using the local linearization

$$a_{\bar{u}_j^\nu}(\cdot,\cdot) = a(\cdot,\cdot) + b_{\bar{u}_j}(\cdot,\cdot), \quad \ell_{\bar{u}_j^\nu} = \ell + f_{\bar{u}_j^\nu}$$

and the corresponding quadratic energy functional

$$\mathcal{J}_{\bar{u}_j^\nu}(v) = \tfrac{1}{2} a_{\bar{u}_j^\nu}(v,v) - \ell_{\bar{u}_j^\nu}(v)$$

the local subproblems (2.2) are approximated by the quadratic obstacle problems

$$v_l = z_l \lambda_l \in \mathcal{D}_l : \ \ \mathcal{J}_{\bar{u}_j^\nu}(w_{l-1} + v_l) \leq \mathcal{J}_{\bar{u}_j^\nu}(w_{l-1} + v), \ \ \forall v \in \mathcal{D}_l. \tag{2.14}$$

To make sure that $w_l = w_{l-1} + v_l \in \mathcal{K}_{\bar{u}_j^\nu}$ holds true, the set of constraints

$$\mathcal{D}_l = \{v \in V_l \mid \underline{\psi}_l \leq v \leq \overline{\psi}_l\} \subset V_l$$

must satisfy

$$0 \in \mathcal{D}_l \subset \mathcal{D}_l^* = \{v \in V_l \mid w_{l-1} + v \in \mathcal{K}_{\bar{u}_j^\nu}\}. \tag{2.15}$$

Hence, the obstacles $\underline{\psi}_l$, $\overline{\psi}_l$ have to be chosen in such a way that

$$\underline{\varphi}_{\bar{u}_j^\nu} - w_{l-1} \leq \underline{\psi}_l \leq 0 \leq \overline{\psi}_l \leq \overline{\varphi}_{\bar{u}_j^\nu} - w_{l-1}. \tag{2.16}$$

We cannot simply take $\underline{\psi}_l = \underline{\varphi}_{\bar{u}_j^\nu} - w_{l-1}$ and $\overline{\psi}_l = \overline{\varphi}_{\bar{u}_j^\nu} - w_{l-1}$, because then the evaluation of w_{l-1} would be necessary to check whether some v is contained in $\mathcal{D}_l$ or not. This would still require an extra prolongation. Suitable *monotone approximations* $\underline{\psi}_l$, $\overline{\psi}_l \in V_l$ of $\underline{\varphi}_{\bar{u}_j^\nu} - w_{l-1}$, $\overline{\varphi}_{\bar{u}_j^\nu} - w_{l-1}$ can be computed by *monotone restriction* of the initial defect obstacles $\underline{\varphi}_{\bar{u}_j^\nu} - \bar{u}_j^\nu$ and $\underline{\varphi}_{\bar{u}_j^\nu} - \bar{u}_j^\nu$, respectively [16, 18].

Unfortunately, the approximate local corrections v_l resulting from (2.14) may violate the monotonicity condition (2.5). Hence, we introduce an additional *local damping parameter* ω_l. The proof of the following proposition relies on the fact that $v_l \neq 0$ holds only if ϕ_j is locally smooth in the direction of λ_l.

Proposition 2.1 *Let $v_l = z_l\lambda_l$ be the solution of (2.14). Assume that the damping parameter satisfies $\omega_l \in [0,1]$ and*

$$\omega_l|z_l| \le 2\left\{\frac{|\ell_{\bar{u}_j^\nu}(\lambda_l) - a_{\bar{u}_j^\nu}(w_{l-1},\lambda_l)| - L_l\|\bar{u}_j^\nu - w_{l-1}\|_{\infty,l}^2}{a_{\bar{u}_j^\nu}(\lambda_l,\lambda_l) + L_l(\|\bar{u}_j^\nu - w_{l-1}\|_{\infty,l} + \omega_l|z_l|)}\right\}_+ \tag{2.17}$$

with

$$L_l = \sum_{p\in\mathcal{N}_j^\circ(\bar{u}_j^\nu)} L_p\lambda_l(p)\int_\Omega \lambda_p^{(j)}(x)\,dx \tag{2.18}$$

and the local maximum norm

$$\|v\|_{\infty,l} = \max_{p\in\mathcal{N}_j\cap \text{int supp }\lambda_l} |v(p)|. \tag{2.19}$$

Then the damped correction $\omega_l v_l$ satisfies the local monotonicity condition (2.5).

The approximation $\omega_l v_l$ can be regarded as resulting from a damped simplified Newton step applied to the original local problem (2.2). Indeed, the approximate correction z_l can be written as

$$z_l = \omega_{\mathcal{D}_l}\frac{\ell_{\bar{u}_j^\nu}(\lambda_l) - a_{\bar{u}_j^\nu}(w_{l-1},\lambda_l)}{a_{\bar{u}_j^\nu}(\lambda_l,\lambda_l)}$$

where the factor $\omega_{\mathcal{D}_l} \in [0,1]$ reflects the condition $v_l = z_l\lambda_l \in \mathcal{D}_l$. Inserting this representation in (2.17), we obtain the upper bound

$$\omega_{\mathcal{D}_l}\omega_l \le 2\frac{a_{\bar{u}_j^\nu}(\lambda_l,\lambda_l)}{a_{\bar{u}_j^\nu}(\lambda_l,\lambda_l) + L_l(\|\bar{u}_j^\nu - w_{l-1}\|_{\infty,l} + \omega_l|z_l|)}\cdot \cdot\left\{\frac{|\ell_{\bar{u}_j^\nu}(\lambda_l) - a_{\bar{u}_j^\nu}(w_{l-1},\lambda_l)| - L_l\|\bar{u}_j^\nu - w_{l-1}\|_{\infty,l}^2}{|\ell_{\bar{u}_j^\nu}(\lambda_l) - a_{\bar{u}_j^\nu}(w_{l-1},\lambda_l)|}\right\}_+$$

for the damping parameter $\omega_{\mathcal{D}_l}\omega_l$ of the fully linearized equation. The first factor on the right hand side is measuring the effect of linearizing $\partial\phi_j$ while the second factor is the price we have to pay for taking the derivative at $\bar{u}_j^\nu$ and not at w_{l-1}. In our numerical experiments, we found that the damping parameters ω_l tend to one as $\bar{u}_j^\nu$ approaches u_j. A theoretical justification will be given elsewhere [19].

In the light of (2.17), we compute the damping parameter $\omega_l \in [0,1]$ by projecting the solutions of the quadratic equation

$$L_l\omega^2|z_l|^2 + \omega|z_l|(a_l + L_lB_l) - 2(|r_l| - L_lB_l^2) = 0 \tag{2.20}$$

to the interval $[0,1]$. Observe that (2.20) only involves the correction z_l, the residual $r_l = \ell_{\bar{u}_j^\nu}(\lambda_l) - a_{\bar{u}_j^\nu}(w_l, \lambda_l)$, the diagonal element $a_l = a_{\bar{u}_j^\nu}(\lambda_l, \lambda_l)$, the local Lipschitz constant L_l and an upper bound B_l,

$$B_l \geq \|\bar{u}_j^\nu - w_{l-1}\|_{\infty,l},$$

for the maximal correction in the support of λ_l. All these parameters are available *without any extra prolongations*. In particular, the Lipschitz constants $L_l = L(\lambda_l)$ are just values of the linear functional $L \in \mathcal{S}_j'$,

$$L(v) = \sum_{p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu)} L_p v(p) \int_\Omega \lambda_p^{(j)}(x)\, dx,$$

which can be restricted like a residual. Upper bounds B_l for the maximal local corrections follow inductively from the triangle inequality. To be precise, we introduce the identification

$$\lambda_{l_{ik}} = \lambda_{p_i}^{(k)}, \quad i = 1, \ldots, n_k,\ k = j, \ldots, 0,$$

of the supporting point p_i and the level k of $\lambda_l \in \Lambda_{\mathcal{S}}$. On the finest grid, we clearly have $B_{l_{ij}} = 0$, $\forall i = 1, \ldots, n_j$. Given values $B_{l_{ik+1}}$ on some level $k+1 \leq j$ are restricted to the next coarser grid by simply taking

$$B_{l_{ik}} := \max_{\{s | s=i \text{ or } e=(p_s,p_i) \in \mathcal{E}_{k+1}\}} B_{l_{sk+1}}, \quad i = 1, \ldots, n_k. \tag{2.21}$$

This procedure defines a restriction operator $M_{k+1}^k : \mathbb{R}^{n_{k+1}} \to \mathbb{R}^{n_k}$. Note that the values $B_{l_{sk+1}}$ appearing in (2.21) are associated with the $k+1$–neighbors $p_s \in \mathcal{N}_{k+1}$ of the node $p_i \in \mathcal{N}_k$. Finally, each local correction $z_{l_{ik}}$ at p_i on the new level k is followed by an update of the bounds $B_{l_{sk}}$ at all k–neighbors p_s of p_i according to

$$B_{l_{sk}} := B_{l_{sk}} + |z_{l_{ik}}|. \tag{2.22}$$

The resulting approximate multilevel relaxation (2.3) based on exact fine grid corrections (cf. (2.2)) and on damped approximate coarse grid corrections (cf. (2.14), (2.20)) can be regarded as an extension of a *standard monotone multigrid method* [16, 17, 18] to piecewise smooth nonlinearities ϕ_j. Hence, we call it *standard monotone multigrid method with local damping*. Other variants with symmetric smoothers, W–cycles or post smoothing can be obtained in a similar way. The convergence follows immediately from Theorem 2.1 and Proposition 2.1.

Theorem 2.2 *The standard monotone multigrid method with local damping is globally convergent.*

In contrast to the nonlinear multilevel relaxation (2.2), the locally linearized coarse grid correction can be implemented as a classical V–cycle.

Algorithm 2.1 Standard Monotone Multigrid Method with Local Damping

given iterate: u_j^ν

nonlinear fine grid smoothing: $\bar{u}_j^\nu := \mathcal{M}_j(u_j^\nu)$

local linearization: *neighborhood* $\mathcal{K}_{\bar{u}_j^\nu}$

pointwise Lipschitz constants L_p, $p \in \mathcal{N}_j$

$a_{\bar{u}_j^\nu} := a + b_{\bar{u}_j^\nu}, \quad \ell_{\bar{u}_j^\nu} := \ell + f_{\bar{u}_j^\nu}$

coarse grid correction:

initialize:

bilinear form and residual: $a^{(j)} := a_{\bar{u}_j^\nu}, \quad r^{(j)} := \ell_{\bar{u}_j^\nu} - a_{\bar{u}_j^\nu}(\bar{u}_j^\nu, \cdot)$

defect obstacles: $\underline{\psi}^{(j)} := \underline{\varphi}^\nu_{\bar{u}_j^\nu} - \bar{u}_j^\nu, \quad \overline{\psi}^{(j)} := \overline{\varphi}^\nu_{\bar{u}_j^\nu} - \bar{u}_j^\nu$

local Lipschitz constants $L^{(j)} := L$

maximal corrections $B^{(j)} := 0$

global correction: $v_j^\nu := 0$

for $k = j-1$ *step* -1 *until* 0 *do*

restrictions:

stiffness matrix: $a^{(k)} := a^{(k+1)}|_{\mathcal{S}_k \times \mathcal{S}_k}$

residual and Lipschitz constants: $r^{(k)} := r^{(k+1)}|_{\mathcal{S}_k}$, $L^{(k)} := L^{(k+1)}|_{\mathcal{S}_k}$

maximal corrections: $B^{(k)} = M_{k+1}^k B^{(k+1)}$

defect obstacles: $\underline{\psi}^{(k)} := \underline{R}_{k+1}^k \underline{\psi}^{(k+1)}, \quad \overline{\psi}^{(k)} := \overline{R}_{k+1}^k \overline{\psi}^{(k+1)}$

coarse grid smoothing: $v^{(k)} := \bar{\mathcal{M}}_k(a^{(k)}, r^{(k)}, \underline{\psi}^{(k)}, \overline{\psi}^{(k)}, L^{(k)}, B^{(k)})(0)$

update:

residual: $r^{(k)} := r^{(k)} - a^{(k)}(v^{(k)}, \cdot)$

defect obstacles: $\underline{\psi}^{(k)} := \underline{\psi}^{(k)} - v^{(k)}, \quad \overline{\psi}^{(k)} := \overline{\psi}^{(k)} - v^{(k)}$

for $k = 0$ *step* 1 *until* $j-1$ *do*

canonical prolongation: $v_j^\nu := v_j^\nu + v^{(k)}$

new iterate: $u_j^{\nu+1} := \bar{u}_j^\nu + v_j^\nu$

In Algorithm 2.1, $\mathcal{M}_j$ stands for one step of the nonlinear Gauß–Seidel relaxation. The *exact* solution of the corresponding one–dimensional problems (2.2) may be quiet costly motivating more economic variants [19]. For the nonlinearity Φ defined in (1.5), the construction of a neighborhood $\mathcal{K}_{\bar{u}_j^\nu}$ and of pointwise Lipschitz constants L_p is straightforward, because $|\Phi'''(z)|$ exists and is monotonically decreasing for $z > 0$. In this special case, we simply set

$$\underline{\varphi}_{\bar{u}_j^\nu}(p) = \bar{u}_j^\nu(p)/2, \qquad \overline{\varphi}_{\bar{u}_j^\nu}(p) = +\infty, \qquad L_p = |\Phi'''(\underline{\varphi}_{\bar{u}_j^\nu}(p))|$$

for all $p \in \mathcal{N}^\circ(\bar{u}_j^\nu)$. Of course, the local linearization may be more difficult (and expensive) in other cases. The bilinear form $a^{(k)}(\cdot,\cdot)$ and the actual

residual $r^{(k)}$ can be applied directly to the elements of the subspaces $\mathcal{S}_k$ which gives the canonical restriction. Suitable *monotone restrictions* $\overline{R}_{k+1}^k$ and $\underline{R}_{k+1}^k$ of the defect obstacles can be found in [16, 18] and the maximal restrictions M_{k+1}^k are defined in (2.21). The evaluation of the correction $v^{(k)}$,

$$v^{(k)} = \sum_{i=1}^{n_k} \omega_{l_{ik}} v_{l_{ik}},$$

can be regarded as one step of a projected Gauß–Seidel method with successive underrelaxation. This includes the successive computation of the damping parameters ω_l (cf. (2.20)) and the inductive update of the maximal corrections $\underline{B}_{l_{ik}}$ (cf. (2.22)). For given bilinear form a, right–hand side r, obstacles $\underline{\psi}$, $\overline{\psi}$, local Lipschitz constants L and maximal corrections B the corresponding iteration operator is denoted by $\bar{\mathcal{M}}_k(a, r, \underline{\psi}, \overline{\psi}, L, B)$. The canonical prolongation is defined by the interpolation of $v_j^\nu \in \mathcal{S}_k \subset \mathcal{S}_{k+1}$.

We finally remark that the definition (2.7) of the regular set $\mathcal{N}_j^\circ(\bar{u}_j^\nu)$ is a bit dangerous from the numerical point of view, because the pointwise Lipschitz constants L_p may become arbitrary large. For this reason, it is useful to select the regular nodes according to the strengthened condition

$$\mathcal{N}_j^\circ(\bar{u}_j^\nu) = \{p \in \mathcal{N}_j \mid \bar{u}_j^\nu(p) > 0 \text{ and } L_p < L_{\max}\} \tag{2.23}$$

with some given threshold $L_{\max} > 0$. Of course, this modification preserves the global convergence. Moreover, if $L_{\max}$ is large enough, then the local corrections which by (2.23) are excluded *a priori* would be excluded anyway *a posteriori* by local damping. This argument is supported by our numerical experiments, showing a considerable robustness of the convergence speed with respect to the actual choice of $L_{\max}$.

3 A Truncated Variant

The standard multigrid method relies on the condition that the coarse grid correction must not change the values of the smoothed iterate $\bar{u}_j^\nu$ at the critical nodes $p \in \mathcal{N}_j^\bullet(\bar{u}_j^\nu)$. Hence, only the $\lambda_l \in \Lambda_\mathcal{S} \setminus \Lambda_j$ with the property

$$\text{int supp } \lambda_l \cap \mathcal{N}_j^\bullet(\bar{u}_j^\nu) = \emptyset \tag{3.1}$$

are allowed to contribute to the coarse grid correction. In comparison with the linear selfadjoint case, this leads to a poor representation of the low frequency parts of the error. To improve the convergence rates by improved coarse grid transport, we will now modify all $\lambda_l \in \Lambda_\mathcal{S} \setminus \Lambda_j$ with the property (3.1) according to the actual guess of the free boundary.

Following [16, 17, 18], we define the modified basis functions

$$\tilde{\lambda}_p^{(k)} = T_{j,k}^\nu \lambda_p^{(k)}, \quad p \in \mathcal{N}_k, \tag{3.2}$$

by using the truncation operators $T_{j,k}^\nu$, $k = 0, \ldots, j$,

$$T_{j,k}^\nu = I_{\mathcal{S}_j^\nu} \circ \ldots \circ I_{\mathcal{S}_k^\nu}. \tag{3.3}$$

Here $I_{\mathcal{S}_k^\nu} : \mathcal{S}_j \to \mathcal{S}_k^\nu$ denotes the $\mathcal{S}_k^\nu$–interpolation, and the spaces $\mathcal{S}_k^\nu \subset \mathcal{S}_k$,

$$\mathcal{S}_k^\nu = \{v \in \mathcal{S}_k \mid v(p) = 0, \ p \in \mathcal{N}_k^\nu\} \subset \mathcal{S}_k, \tag{3.4}$$

are the reduced subspaces with respect to $\mathcal{N}_k^\nu = \mathcal{N}_k \cap \mathcal{N}_j^\bullet(\bar{u}_j^\nu)$, $k = 0, \ldots, j$. Similar subspaces of $\mathcal{S}_j$ a have been considered recently by other authors [2, 13, 20] in connection with the coarsening of a given mesh.

Replacing the multilevel nodal basis $\Lambda_\mathcal{S}$ by the actual truncation $\tilde{\Lambda}_\mathcal{S}^\nu$,

$$\tilde{\Lambda}_\mathcal{S}^\nu = \left(\lambda_{p_1}^{(j)}, \ldots, \lambda_{p_{n_j}}^{(j)}, \tilde{\lambda}_{p_1}^{(j-1)}, \ldots, \tilde{\lambda}_{p_{n_{j-1}}}^{(j-1)}, \ldots, \tilde{\lambda}_{p_1}^{(0)}, \ldots, \tilde{\lambda}_{p_{n_0}}^{(0)}\right), \quad \nu \geq 0,$$

we can now derive a globally convergent *truncated monotone multigrid method* by the same reasoning as described in the previous section. As all $\lambda_l \in \Lambda_\mathcal{S}$ with the property (3.1) are contained in $\tilde{\Lambda}_\mathcal{S}^\nu$, we can hope for a higher convergence speed than in the standard case. This heuristic argument is supported by the numerical results reported below. To ensure the monotonicity of the coarse grid correction in the sense of (2.5), we can compute appropriate local damping parameters $\tilde{\omega}_l$ from a straightforward analogue of (2.20). Then the global convergence of the resulting *truncated monotone multigrid method with local damping* follows from a suitable variant of Theorem 2.1 (see e.g. Theorem 2.2 in [18]).

Theorem 3.1 *The truncated monotone multigrid method with local damping is globally convergent.*

The algorithm can be implemented as a variant of the standard monotone multigrid method with local damping. More precisely, the restrictions and prolongations appearing in Algorithm 2.1 have to be modified as follows:

Modifications of Algorithm 2.1 Truncated Variant

modified restrictions of the bilinear form and of the residual:

treat all entries from the actual critical nodes $\mathcal{N}_j^\bullet(\bar{u}_j^\nu)$ *as zero*

modified quasioptimal restrictions of the upper (lower) defect obstacle:

treat all entries from the actual critical nodes $\mathcal{N}_j^\bullet(\bar{u}_j^\nu)$ *as* ∞ $(-\infty)$

modified prolongations of the corrections:

prolongate zero to all critical nodes

Observe that the restrictions M_{k+1}^k of the maximal corrections $B_{l_{ik+1}}$ remain unchanged.

Again, the definition (2.7) of the regular set $\mathcal{N}_j^\circ(\bar{u}_j^\nu)$ should be replaced by (2.23) in numerical computations. For truncated variants, the convergence speed might be even improved by excluding the nodes with very large pointwise Lipschitz constants from the coarse grid correction, because then the damping of the corrections associated with the resulting truncated basis functions could be considerably reduced. This behavior is different from the standard case where it does not make much difference whether such a node is considered as regular or not.

4 Numerical Experiments

We will now illustrate the numerical performance of our monotone multigrid methods in the case of the well–known porous medium equation (cf. e.g. [22])

$$\theta_t = \Delta\theta^m, \quad m > 1, \text{ in } x \in (-3,3) \times (-3,3), \ t > 0 \tag{4.1}$$

with homogeneous Dirichlet boundary conditions. An analytical solution of this model problem was given by Barenblatt [4]

$$\theta(x,t) = \frac{1}{(t+T_0)^{1/m}} \left\{ |a|^2 - \frac{m-1}{4m^2} \frac{|x|^2}{(t+T_0)^{1/m}} \right\}_+^{1/(m-1)} .$$

We select $a = 0.5$, $T_0 = 0.1$ and $m = 6$ in our computations. A very popular discretization scheme for (4.1) was proposed by Jäger and Kačur [14, 15]. See e.g. Bänsch [3] for numerical results. The basic idea is to use a Chernov–type formula with pointwise relaxation parameters $\mu_p \approx 1/(m\theta(p)^{m-1})$.

After the Kirchhoff–type transformation

$$U = \theta^m$$

we apply an implicit Euler discretization with uniform time step τ, to obtain spatial problems of the form (1.1) with Φ defined in (1.5) and

$$a(v,w) = \tau \int_\Omega \nabla v \cdot \nabla w \, dx, \quad \ell(v) = \int_\Omega U_\tau(\cdot, t_{k-1}) v \, dx$$

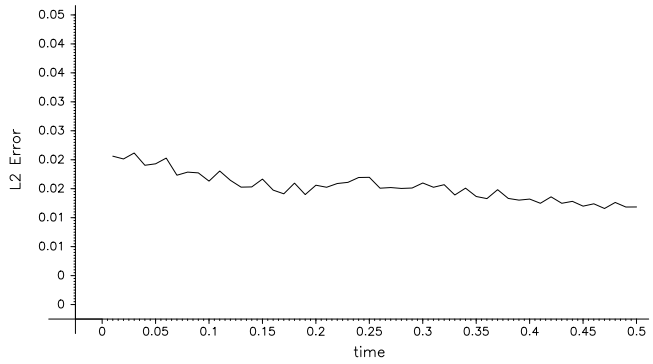


Figure 4.1: Spatial Discretization Error $\|\theta(\cdot, t_k) - \theta_k\|_{L^2(\Omega)}$

for the solution $U_k \approx U(\cdot, t_k)$ at each time step. Stationary solutions, e.g. with inflow and absorption, can be computed in the same way. Observe that the inverse transformation $\theta_k = \sqrt[m]{U_k}$ is ill–conditioned at the free boundary, reflecting the the degeneracy of (4.1).

We discretize the spatial problems on the subdomain $\Omega = (0,3) \times (0,3)$ using symmetry boundary conditions at $x = 0$ and $y = 0$. The initial triangulation $\mathcal{T}_0$ is obtained by subdividing Ω into four congruent triangles and the final triangulation $\mathcal{T}_j$ is resulting from $j = 7$ uniform refinement steps. Of course, adaptive techniques are very attractive for such kind of problems but this is not our subject here. The time step is $\tau = 0.01$ and we stopped the computation at $t_{50} = 0.5$. The implementation was carried out in the framework of the finite element toolbox KASKADE [5]. Figure 4.1 shows the spatial discretization error in $L^2(\Omega)$ over the time t. Observe the parabolic smoothing by the porous medium equation.

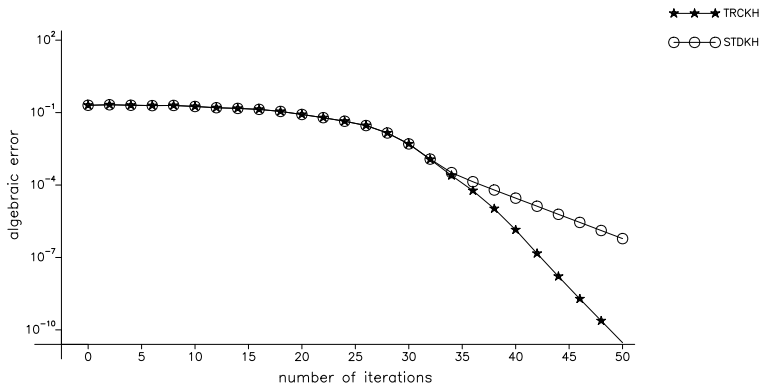


Figure 4.2: Iteration History: Initial Iterate $u_j^0 = 0$

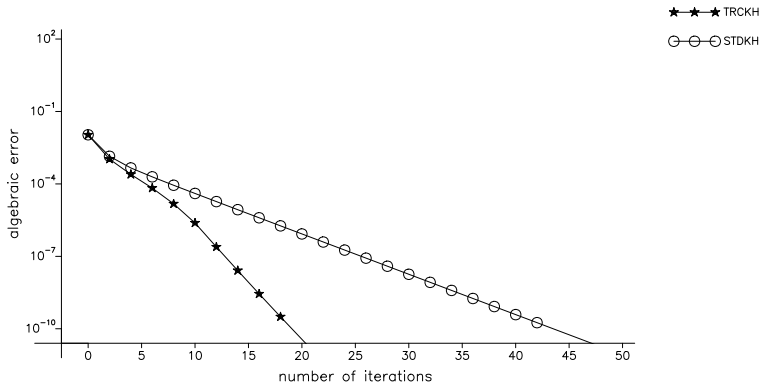


Figure 4.3: Iteration History: Nested Iteration

In order to compare the convergence properties of the standard monotone multigrid method (STDKH) and of the truncated version (TRCKH), we now consider the iterative solution of the discrete problem arising in the first time step on the final triangulation $\mathcal{T}_j$, $j = 7$. Throughout the following, we use the modified definition (2.23) of $\mathcal{N}_j^\circ(\bar{u}_j^\nu)$ with $L_{\max} = 10^6$. Starting with the initial iterate $u_j^0 = 0$, we obtain the algebraic errors $\|u_j - u_j^\nu\|$, $\nu = 0, \ldots, 50$, as shown in Figure 4.2. The overall convergence behavior can be divided into a *transient* phase, characterized by (severe) damping of the coarse grid correction and an *asymptotic* phase, where the local damping parameters ω_l are equal to one. As the iteration starts from the singularity $u_j^0 = 0$, it takes more than 25 steps until the asymptotic behavior is reached. In comparison with the standard method, the truncated variant then exhibits a tremendous improvement of the convergence rates, giving a numerical justification for the truncation of nodal basis functions. Note that the transient convergence properties of both algorithms are basically the same.

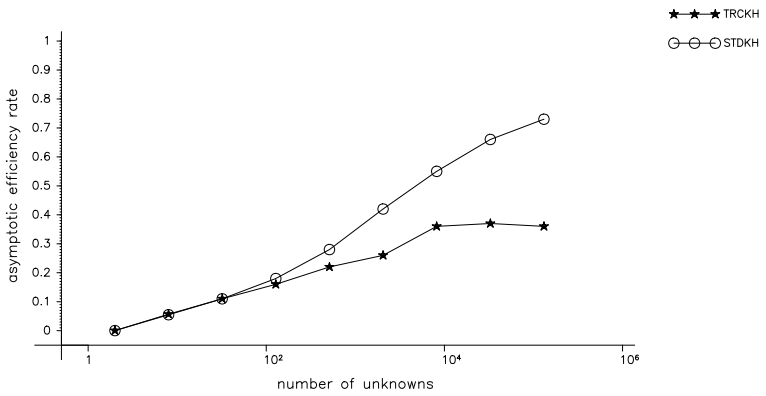


Figure 4.4: Asymptotic Efficiency Rates

Better initial iterates should be used in practical applications. Starting with the final iterate from the previous level, the transient phase is almost eliminated from the convergence history. This is illustrated in Figure 4.3. In particular, the fast asymptotic convergence of TRCKH now dominates the whole iteration process.

To study the convergence properties for increasing j, we introduce the asymptotic efficiency rates ρ_j,

$$\rho_j = \sqrt[\nu_0]{\delta_j^{\nu_0}/\delta_j^0}, \quad j = 0, \ldots, 8, \tag{4.2}$$

where δ_j^ν denotes the algebraic error after ν iteration steps. We choose ν_0 such that $\delta_j^{\nu_0} < 10.^{-12}$. The results are shown in Figure 4.4. Note that the

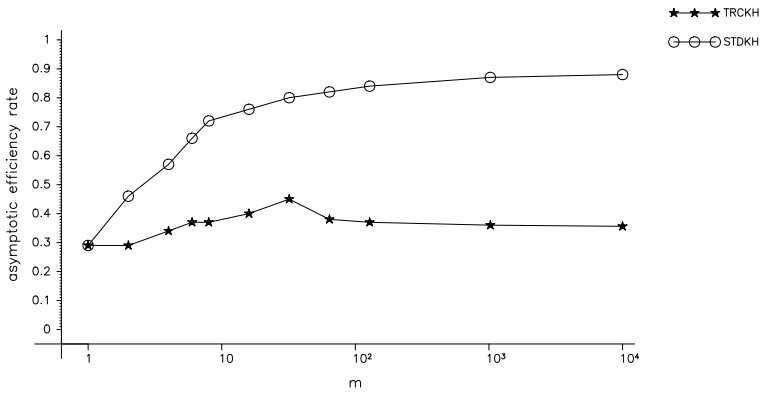


Figure 4.5: Robustness with respect to Φ_m

asymptotic efficiency rates for both multigrid methods seem to saturate with increasing j.

In our previous experiments we found similar convergence properties as in the piecewise linear case [16, 17, 18]. In order to illustrate the robustness of the methods with respect to the smoothness of the nonlinearity, we now vary the parameter m in the definition (1.5) of $\Phi = \Phi_m$. We still use the same bilinear form and right hand side (m=6) as before. The corresponding asymptotic efficiency rates on the refinement level $j = 7$ are depicted in Figure 4.5. Observe that the convergence speed of the truncated variant is almost independent of m, though Φ_m is linear on u_j for $m = 1$ and becomes even discontinuous as m tends to infinity.

Acknowledgements. The author gratefully acknowledges the hospitality of P. Deuflhard and his staff at the Konrad–Zuse–Zentrum Berlin during the preparation of this manuscript. The work was supported by a Konrad–Zuse–Fellowship.

Bibliography

[1] R.E. Bank and D.J. Rose. Analysis of a multilevel iterative method for nonlinear finite element equations. *Math. Comp.*, 39:453–665, 1982.

[2] R.E. Bank and J. Xu. An algorithm for coarsening unstructured meshes. *Numer. Math.*, 73:1–36, 1996.

[3] E. Bänsch. Numerical experiments with adaptivity for the porous medium equation. *Acta Math. Univ. Comenianae*, LXIV:157–172, 1995.

[4] G.I. Barenblatt. On self–similar solutions of compressible fluids in porous media. (In russian), *Prikl. Mat. Mech.*, 16:679–698, 1952.

[5] R. Beck, B. Erdmann, and R. Roitzsch. KASKADE Manual, Version 3.0. Technical Report TR95-4, Konrad–Zuse–Zentrum (ZIB), Berlin, 1995.

[6] F.A. Bornemann, B. Erdmann, and R. Kornhuber. Adaptive multilevel methods in three space dimensions. *Int. J. Numer. Meth. Engrg.*, 36:3187–3203, 1993.

[7] F.A. Bornemann, B. Erdmann, and R. Kornhuber. A posteriori error estimates for elliptic problems in two and three space dimensions. *SIAM J. Numer. Anal.*, 33, 1996.

[8] P. Deuflhard and M. Weiser. Local inexact Newton multilevel FEM for nonlinear elliptic problems. Preprint SC 96-30, Konrad–Zuse–Zentrum (ZIB), Berlin, 1996.

[9] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. North–Holland, Amsterdam, 1976.

[10] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Springer, New York, 1984.

[11] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer, Berlin, 1985.

[12] W. Hackbusch and A. Reusken. Analysis of a damped nonlinear multilevel method. *Numer. Math.*, 55:225–246, 1989.

[13] W. Hackbusch and S.A. Sauter. Composite finite elements for the approximation of PDEs on domains with complicated micro–structures. *Numer. Math.*, to appear.

[14] W. Jäger and J. Kačur. Solution of porous medium type systems by linear approximation schemes. *Numer. Math.*, 60:407–427, 1991.

[15] W. Jäger and J. Kačur. Solution of doubly nonlinear and degenerate parabolic problems by relaxation schemes. M^2AN, 29:605–627, 1995.

[16] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities I. *Numer. Math.*, 69:167 – 184, 1994.

[17] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities II. *Numer. Math.*, 72:481 – 499, 1996.

[18] R. Kornhuber. *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*. Wiley–Teubner, 1997.

[19] R. Kornhuber. Globally convergent multigrid methods for porous medium type problems. *To appear*.

[20] R. Kornhuber and H. Yserentant. Multilevel methods for elliptic problems on domains not resolved by the coarse grid. In D.E. Keyes and J. Xu, editors, *Proceedings of the 7th International Conference on Domain Decomposition Methods 1993*, pages 49–60, Providence, 1994. AMS.

[21] S.F. McCormick. *Multilevel Projection Methods for Partial Differential Equations*. SIAM, Philadalphia, 1992.

[22] J.L. Vazquez. An introduction to the mathematical theorie of the porous medium equation. In M.C. Delfour, editor, *Shape Optimization and Free Boundaries*, volume 380 of *Mathematical and Physical Sciences*, pages 347–389, Dordrecht, 1992. Kluwer.

[23] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34:581–613, 1992.

[24] H. Yserentant. Old and new convergence proofs for multigrid methods. *Acta Numerica*, pages 285–326, 1993.