Adaptive Monotone Multigrid Methods for some Non–Smooth Optimization Problems

Ralf Kornhuber
Weierstraß–Institut für Angewandte Analysis und Stochastik Berlin
Mohrenstraße 39, D–10117 Berlin, Fed. Rep. of Germany

Abstract. We consider the fast solution of non–smooth optimization problems as resulting for example from the approximation of elliptic free boundary problems of obstacle or Stefan type. Combining well–known concepts of successive subspace correction methods with convex analysis, we derive a new class of multigrid methods which are globally convergent and have logarithmic bounds of the asymptotic convergence rates. The theoretical considerations are illustrated by numerical experiments.

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

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

1 Introduction

The weak solution of an elliptic selfadjoint boundary value problem is obtained by minimizing the corresponding quadratic energy functional $\mathcal{J}$. We will consider the more general problem

$$\mathcal{J}(u) + \phi(u) = \min$$

with ϕ denoting a convex functional which is piecewise quadratic but not differentiable. Such non–smooth optimization problems are modeling physical phenomena involving a change of phase. Obstacle problems or time–discretized two–phase Stefan problems are typical examples (see e.g. [6, 7, 14] for further applications).

The continuous problem is discretized by piecewise linear finite elements with respect to a sequence of triangulations resulting from the successive adaptive refinement of a given initial mesh. A corresponding adaptive algorithm has been described in [20]. In this paper we will concentrate on the efficient solution of the nonlinear discrete problems arising on each refinement level.

The most delicate question in constructing a multigrid method for a nonlinear problem is how to represent the nonlinearity on the coarse grids. This process usually involves some kind of linearization. Unfortunately, the computed corrections may exceed the region in which the actual linearization is valid. This problem is often remedied by *a posteriori* damping of the coarse grid correction [12]. The appropriate selection of damping parameters is a non–trivial task [13, 15]. The basic idea of *monotone multigrid methods* to be presented here is first to find out a neighborhood of the actual iterate in which the actual linearization is valid and then to constrain the coarse grid correction to this neighborhood. In this way, we ensure monotonically decreasing energy in course of the iteration. It turns out that such kind of *local linearization* is equivalent to the damping of the inaccessible nonlinear coarse grid correction. Suitable damping parameters are implicitly incorporated in the constraints. This approach provides globally convergent methods and we can prove asymptotic multigrid convergence rates. In comparison with previous multigrid algorithms, monotone multigrid methods turned out to be superior both from a theoretical and from a numerical point of view [18, 19].

As proofs of the basic convergence results have been already presented elsewhere [18, 19], we will try to give an algorithmically oriented presentation here. In this way, we hope to simplify further generalizations of the underlying ideas and the implementation in existing multigrid codes. A more detailed description will be contained in a forthcoming work [21].

2 Discretization of the Continuous Problem

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

$$u \in H: \quad \mathcal{J}(u) + \phi(u) \leq \mathcal{J}(v) + \phi(v), \quad v \in H, \tag{2.1}$$

on a closed subspace $H \subset H^1(\Omega)$. For simplicity, we select $H = H_0^1(\Omega)$ corresponding to homogeneous Dirichlet boundary conditions. Other boundary conditions of Neumann or mixed type and the case of three space dimensions can be treated in a similar way [3, 4].

The quadratic functional $\mathcal{J}$,

$$\mathcal{J}(v) = \tfrac{1}{2} a(v,v) - \ell(v), \tag{2.2}$$

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 convex functional ϕ of the form

$$\phi(v) = \int_\Omega \Phi(v(x))\, dx, \tag{2.3}$$

is generated by a scalar function $\Phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$. We assume that Φ is convex and piecewise quadratic,

$$\Phi(z) = \tfrac{1}{2} b_i z^2 - f_i z + c_i, \quad \theta_i \leq z \leq \theta_{i+1}, \tag{2.4}$$

on a partition

$$-\infty \leq \theta_0 < \theta_1 < \ldots < \theta_N < \theta_{N+1} \leq +\infty$$

of the closed interval $K \subset \mathbb{R}$ bounded by θ_0, θ_{N+1} and that $\Phi(z) = \infty$ holds, if $z \notin K$. To make sure that $0 \in K$, we assume $\theta_0 \leq 0 \leq \theta_{N+1}$. The convexity implies that Φ is continuous on K but the derivative Φ' may be discontinuous at the transition points θ_i, $i = 1, \ldots, N$.

From the assumptions on Φ, the functional ϕ is convex, lower semi–continuous, and proper (i.e. $\phi(v) > -\infty$ and $\phi \not\equiv +\infty$). In particular, ϕ is finite and continuous on the closed convex set $\mathcal{K} \subset H$,

$$\mathcal{K} = \{v \in H \mid v(x) \in K, \text{ a.e. in } \Omega\} \neq \emptyset.$$

Hence, it follows from well–known results [10] that the optimization problem (2.1) has a unique solution $u \in H$ and can be equivalently rewritten as the elliptic variational inequality of the second kind

$$u \in H: \quad a(u, v-u) + \phi(v) - \phi(u) \geq \ell(v-u)\,, \quad v \in H. \tag{2.5}$$

Let $\mathcal{T}_j$ be a consistent partition of Ω in triangles with minimal diameter $h_j = \mathcal{O}(2^{-j})$. The interior nodes and edges of $\mathcal{T}_j$ are denoted by $\mathcal{N}_j$ and $\mathcal{E}_j$, respectively. The finite element space $\mathcal{S}_j \subset H$ contains all continuous functions $v \in H$ which are linear on each triangle $t \in \mathcal{T}_j$. $\mathcal{S}_j$ is spanned by the nodal basis $\Lambda_j = \{\lambda_p^{(j)} \mid p \in \mathcal{N}_j\}$. Replacing H by the finite dimensional approximation $\mathcal{S}_j$ and the functional ϕ by its $\mathcal{S}_j$–interpolate ϕ_j,

$$\phi_j(v) = \sum_{p \in \mathcal{N}_j} \Phi(v(p)) \int_\Omega \lambda_p^{(j)}(x)\, dx, \quad v \in \mathcal{S}_j, \tag{2.6}$$

we obtain the *discrete optimization problem*

$$u_j \in \mathcal{S}_j : \quad \mathcal{J}(u_j) + \phi_j(u_j) \le \mathcal{J}(v) + \phi_j(v), \quad v \in \mathcal{S}_j. \tag{2.7}$$

Observe that the discrete energy $\mathcal{J} + \phi_j$ is finite and continuous on the closed convex set $\mathcal{K}_j \subset \mathcal{S}_j$,

$$\mathcal{K}_j = \{v \in \mathcal{S}_j \mid v(p) \in K,\ p \in \mathcal{N}_j\} \neq \emptyset.$$

It is easily seen that the discrete functional ϕ_j still is convex, lower semi–continuous, and proper. Hence, the discrete problem (2.7) has a unique solution $u_j \in \mathcal{S}_j$ which is characterized by the variational inequality

$$u_j \in \mathcal{S}_j : \quad a(u_j, v - u_j) + \phi_j(v) - \phi_j(u_j) \ge \ell(v - u_j), \quad v \in \mathcal{S}_j. \tag{2.8}$$

The convergence of the discretization (2.7) follows from general results as condensed by Glowinski [10] and error estimates have been derived for example by Brezzi et al. [5] or Elliot [9].

3 Multilevel Relaxations

Assume that $\mathcal{T}_j$ is resulting from j refinements of a given, intentionally coarse triangulation $\mathcal{T}_0$ of Ω. In this way, we obtain a sequence of triangulations $\mathcal{T}_0, \ldots, \mathcal{T}_j$ and of corresponding nested finite element spaces $\mathcal{S}_0 \subset \ldots \subset \mathcal{S}_j$. To avoid additional technicalities, we assume for the moment that each triangulation is uniformly refined, i.e. that each triangle $t \in \mathcal{T}_{k-1}$ is subdivided into four congruent subtriangles to obtain the next triangulation $\mathcal{T}_k$.

Collecting the nodal basis functions $\Lambda_k = \{\lambda_{p_i}^{(k)} \mid i = 1, \ldots, n_k\}$ 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 optimization 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 [24], Xu [26], or Yserentant [27]). 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 (2.7). To be precise, we introduce the splitting

$$\mathcal{S}_j = \sum_{l=1}^{m} V_l, \tag{3.1}$$

of $\mathcal{S}_j$ in the one–dimensional subspaces $V_l = \operatorname{span}\{\lambda_l\}$, $l = 1, \ldots, m$. For a given ν–th iterate $u_j^\nu \in \mathcal{K}_j$ one step of a *nonlinear* multilevel relaxation now reads as follows.

Algorithm 3.1 (Nonlinear Multilevel Relaxation)
initialize: $w_0 := u_j^\nu$
for $l = 1$ *step 1 until* m *do*

$$\begin{aligned} &\bar{v}_l \in V_l: \quad \mathcal{J}(w_{l-1} + \bar{v}_l) + \phi_j(w_{l-1} + \bar{v}_l) \leq \\ &\qquad\qquad \leq \mathcal{J}(w_{l-1} + v) + \phi_j(w_{l-1} + v), \quad v \in V_l \\ &w_l := w_{l-1} + \omega_l \bar{v}_l, \quad \omega_l \in [0,1] \end{aligned} \tag{3.2}$$

new iterate: $u_j^{\nu+1} := w_m$

Observe that we have introduced certain damping parameters ω_l which will be useful later on. Assuming

$$\omega_l = 1, \quad l = 1, \ldots, n_j, \tag{3.3}$$

the leading *fine grid corrections* in direction of $\lambda_l \in \Lambda_j$ can be regarded as one step of the well–known single grid relaxation [10]. The corresponding iteration operator is denoted by $\mathcal{M}_j$ and $\bar{u}_j^\nu := \mathcal{M}_j(u_j^\nu)$ is called smoothed iterate. Note that we have $\bar{u}_j^0 \in \mathcal{K}_j$ for an arbitrary 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 contributions of the error.

The following convergence proof will be based on the global convergence of the leading single grid relaxation and on the monotonicity

$$\mathcal{J}(w_l) + \phi_j(w_l) \leq \mathcal{J}(w_{l-1}) + \phi_j(w_{l-1}), \quad l = 1, \ldots, m, \tag{3.4}$$

of the local corrections.

Theorem 3.1 *For any initial iterate $u_j^0 \in \mathcal{S}_j$ and any sequence of damping parameters with the property (3.3) the sequence of iterates $(u_j^\nu)_{\nu \geq 0}$ produced by Algorithm 3.1 converges to the solution u_j of the discrete problem (2.7).*

Proof: We will use the abbreviation $\bar{\mathcal{J}} = \mathcal{J} + \phi_j$. The sequence of iterates u_j^ν, $\nu \geq 0$, is bounded because the monotonicity (3.4) yields

$$\bar{\mathcal{J}}(u_j^\nu) \leq \bar{\mathcal{J}}(\bar{u}_j^0) < \infty, \ \ \nu \geq 1,$$

and we have $\bar{\mathcal{J}}(u^\nu) \to \infty$ for any unbounded sequence $(u^\nu)_{\nu \geq 0} \subset \mathcal{S}_j$.

Let $u_j^{\nu_k}$, $k \geq 0$, be an arbitrary, convergent subsequence of u_j^ν with the limit $u^* \in \mathcal{S}_j$,

$$u_j^{\nu_k} \to u^*, \quad k \to \infty. \tag{3.5}$$

Such a subsequence exists, because u_j^ν is bounded and $\mathcal{S}_j$ has finite dimension. Observe that $u^* \in \mathcal{K}_j$, because $(u_j^{\nu_k})_{k \geq 1} \subset \mathcal{K}_j$ and $\mathcal{K}_j$ is a closed subset of $\mathcal{S}_j$. In order to prove $u^* = u_j$, we will show that u^* is a fixed point of the single grid relaxation $\mathcal{M}_j$. It is easily checked that $\mathcal{M}_j$ is continuous so that

$$\mathcal{M}_j(u_j^{\nu_k}) \to \mathcal{M}_j(u^*), \quad k \to \infty. \tag{3.6}$$

As each step of the multilevel relaxation starts with the single grid relaxation $\mathcal{M}_j$, the local monotonicity (3.4) implies

$$\bar{\mathcal{J}}(u_j^{\nu_k+1}) \leq \bar{\mathcal{J}}(\mathcal{M}_j(u_j^{\nu_k})) \leq \bar{\mathcal{J}}(u_j^{\nu_k}).$$

From (3.4) we also have

$$\bar{\mathcal{J}}(u_j^{\nu_{k+1}}) \leq \bar{\mathcal{J}}(u_j^{\nu_k+1}).$$

In virtue of the convergence (3.5), (3.6), and the continuity of $\bar{\mathcal{J}}$ on $\mathcal{K}_j$, this leads to

$$\bar{\mathcal{J}}(\mathcal{M}_j(u^*)) = \bar{\mathcal{J}}(u^*). \tag{3.7}$$

It is easily seen that (3.7) holds, if and only if all local corrections of the single grid relaxation applied to u^* are zero, i.e. $\mathcal{M}_j(u^*) = u^*$. The single grid relaxation is globally convergent so that u_j is the only fixed point of $\mathcal{M}_j$, giving $u^* = u_j$.

As $(u_j^{\nu_k})_{k\geq 0}$ was an arbitrary convergent subsequence, the whole sequence u_j^ν must converge to u_j. This completes the proof. ∎

In the special case $\phi_j \equiv 0$ Algorithm 3.1 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. To find a related implementation for the nonlinear case, we will now consider the local subproblems for the local corrections

$$\bar{v}_l = \bar{z}_l \lambda_l \in V_l$$

in more detail. Using subdifferential calculus [8], (3.2) can be rewritten as the following *scalar inclusion* for the unknown coefficient $\bar{z}_l \in \mathbb{R}$

$$0 \in a(\lambda_l, \lambda_l)\bar{z}_l - (\ell(\lambda_l) - a(w_{l-1}, \lambda_l)) + \partial\phi_j(w_{l-1} + \bar{z}_l\lambda_l)(\lambda_l). \tag{3.8}$$

Observe that $\partial\phi_j(w_{l-1} + z\lambda_l)(\lambda_l)$ is a piecewise linear function in z because the scalar function Φ which generates ϕ_j is piecewise quadratic. Hence, after some tedious calculations, the fine grid corrections in direction of $\lambda_l \in \Lambda_j$ are available in closed form [19, 20].

Let us consider the coarse grid correction in direction of some fixed $\lambda_l \in \Lambda_k$, $k < j$. It is clear that $\bar{z}_l$ can not be computed without evaluating the intermediate iterate w_{l-1} at all nodes $p \in \text{int supp } \lambda_l$, because the subdifferential $\partial\phi_j(w_{l-1} + z\lambda_l)(\lambda_l)$ is *nonlinear* with respect to the argument $w_{l-1} + z\lambda_l$. This leads to (at least) one additional prolongation for each local coarse grid correction. As a consequence, the number of operations for one complete iteration step is no longer linearly bounded but grows like $\mathcal{O}(n_j \log(n_j))$.

To preserve the optimal numerical complexity of the classical V–cycle, we will now approximate the exact coarse grid corrections $\bar{v}_l$ by a *local linearization* of the subproblems (3.8) in a neighborhood of the smoothed iterate $\bar{u}_j^\nu$. For this reason, we define the *discrete phases* $\mathcal{N}_j^i(\bar{u}_j^\nu) \subset \mathcal{N}_j$ of $\bar{u}_j^\nu$ by

$$\mathcal{N}_j^i(\bar{u}_j^\nu) = \{p \in \mathcal{N}_j \mid \bar{u}_j^\nu(p) \in (\theta_i, \theta_{i+1})\}, \quad , i = 0, \ldots, N.$$

At the remaining *critical nodes* $\mathcal{N}_j^\bullet(\bar{u}_j^\nu)$,

$$\mathcal{N}_j^\bullet(\bar{u}_j^\nu) = \mathcal{N}_j \setminus \bigcup_{i=0}^{N} \mathcal{N}_j^i(\bar{u}_j^\nu), \tag{3.9}$$

$\bar{u}_j^\nu$ has values in the set $\{\theta_1, \ldots, \theta_N\}$ of transition points.
Now the key observation is that $\phi_j(w)$ *is a quadratic functional as long as the discrete phases of* w *remain invariant.* Such a neighborhood of $\bar{u}_j^\nu$ is given by the closed convex subset $\mathcal{K}_{\bar{u}_j^\nu} \subset \mathcal{S}_j$,

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

where the obstacles $\underline{\varphi}_{\bar{u}_j^\nu}, \overline{\varphi}_{\bar{u}_j^\nu} \in \mathcal{S}_j$ are defined by

$$\begin{aligned} \underline{\varphi}_{\bar{u}_j^\nu}(p) = \theta_i,\ \ \overline{\varphi}_{\bar{u}_j^\nu}(p) = \theta_{i+1}, \quad &\text{if} \quad p \in \mathcal{N}_j^i(\bar{u}_j^\nu), \\ \underline{\varphi}_{\bar{u}_j^\nu}(p) = \overline{\varphi}_{\bar{u}_j^\nu}(p) = \bar{u}_j^\nu(p), \quad &\text{if} \quad p \in \mathcal{N}_j^\bullet(\bar{u}_j^\nu). \end{aligned} \tag{3.10}$$

Recall that $\mathcal{K}_{\bar{u}_j^\nu}$ *is fixed by the fine grid correction.*
By construction, the functional ϕ_j on $\mathcal{K}_{\bar{u}_j^\nu}$ can be rewritten as

$$\phi_j(w) = \tfrac{1}{2} b_{\bar{u}_j^\nu}(w, w) - f_{\bar{u}_j^\nu}(w) + \text{const.},\ \ w \in \mathcal{K}_{\bar{u}_j^\nu}, \tag{3.11}$$

with the symmetric positive semidefinite bilinear form $b_{\bar{u}_j^\nu}(v, w)$,

$$b_{\bar{u}_j^\nu}(v, w) = \sum_{i=0}^{N} \sum_{p \in \mathcal{N}_j^i(\bar{u}_j^\nu)} b_i v(p) w(p) \int_\Omega \lambda_p^{(j)}(x)\, dx, \tag{3.12}$$

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

$$f_{\bar{u}_j^\nu}(v) = \sum_{i=0}^{N} \sum_{p \in \mathcal{N}_j^i(\bar{u}_j^\nu)} f_i v(p) \int_\Omega \lambda_p^{(j)}(x)\, dx. \tag{3.13}$$

To take advantage of the simple representation of bilinear forms and linear operators on the coarse spaces $\mathcal{S}_k$, $k < j$, we want to constrain the local corrections in such a way that all the intermediate iterates w_l remain in $\mathcal{K}_{\bar{u}_j^\nu}$. Equivalently, *the coarse grid corrections must not cause a change of phase.*
Hence, the local subproblems (3.2) in Algorithm 3.1 are replaced by the quadratic obstacle problems

$$\begin{aligned} \bar{v}_l \in \mathcal{D}_l^* :\ \ &\mathcal{J}(w_{l-1} + v_l^*) + \phi_j(w_{l-1} + v_l^*) \le \\ &\le \mathcal{J}(w_{l-1} + v) + \phi_j(w_{l-1} + v),\ \ v \in \mathcal{D}_l^*, \end{aligned} \tag{3.14}$$

with constraints

$$\mathcal{D}_l^* = \{v \in V_l \mid w_{l-1} + v \in \mathcal{K}_{\bar{u}_j^\nu}\} \subset V_l. \tag{3.15}$$

In the light of (3.11), the energy functional on $\mathcal{D}_l^*$ has the representation

$$\mathcal{J}(w_{l-1} + v) + \phi_j(w_{l-1} + v) = \tfrac{1}{2} a_{\bar{u}_j^\nu}(v, v) - r_{\bar{u}_j^\nu}(w_{l-1})(v) + \text{const.} \tag{3.16}$$

where we have set $r_{\bar{u}_j^\nu}(w_{l-1}) = \ell_{\bar{u}_j^\nu} - a_{\bar{u}_j^\nu}(w_{l-1}, \cdot)$ and

$$a_{\bar{u}_j^\nu}(\cdot, \cdot) = a(\cdot, \cdot) + b_{\bar{u}_j}(\cdot, \cdot), \quad \ell_{\bar{u}_j^\nu} = \ell + f_{\bar{u}_j^\nu}. \tag{3.17}$$

The set $\mathcal{D}_l^*$ clearly contains all $v \in V_l$ satisfying

$$\underline{\varphi}_{\bar{u}_j^\nu} - w_{l-1} \le v \le \overline{\varphi}_{\bar{u}_j^\nu} - w_{l-1}. \tag{3.18}$$

Hence, we still have to evaluate the intermediate iterate $w_{l-1} \in \mathcal{S}_j$ to check wether some v is contained in $\mathcal{D}_l^*$ or not. For this reason, we approximate $\mathcal{D}_l^*$ by replacing the fine grid defect obstacles $w_{l-1} - \underline{\varphi}_{\bar{u}_j^\nu}$, $w_{l-1} - \overline{\varphi}_{\bar{u}_j^\nu}$ appearing in (3.18) by coarse grid approximations $\underline{\psi}_l$, $\overline{\psi}_l \in V_l$. To make sure that the resulting subset

$$\mathcal{D}_l = \{v \in V_l \mid \underline{\psi}_l \le v \le \overline{\psi}_l\} \subset V_l \tag{3.19}$$

satisfies $0 \in \mathcal{D}_l \subset \mathcal{D}_l^*$, we require

$$\underline{\varphi}_{\bar{u}_j^\nu} - w_{l-1} \le \underline{\psi}_l \le 0 \le \overline{\psi}_l \le \overline{\varphi}_{\bar{u}_j^\nu} - w_{l-1}. \tag{3.20}$$

Let us postpone the construction of such *monotone approximations* $\underline{\psi}_l$, $\overline{\psi}_l$ to the next section. We now summarize one complete step of our linearized multilevel relaxation.

Algorithm 3.2 (Linearized Multilevel Relaxation)
fine grid smoothing: $\bar{u}_j^\nu := \mathcal{M}_j(u_j^\nu)$
local linearization: $a_{\bar{u}_j^\nu} := a + b_{\bar{u}_j^\nu}$, $\ell_{\bar{u}_j^\nu} := \ell + f_{\bar{u}_j^\nu}$
coarse grid correction:
initialization: $w_{n_j} := \bar{u}_j^\nu$
for $l = n_j + 1$ *step 1 until* m *do*
update $\mathcal{D}_l$

$$v_l \in \mathcal{D}_l : \ \tfrac{1}{2} a_{\bar{u}_j^\nu}(v_l, v_l) - r_{\bar{u}_j^\nu}(w_{l-1})(v_l) \le \tfrac{1}{2} a_{\bar{u}_j^\nu}(v, v) - r_{\bar{u}_j^\nu}(w_{l-1})(v), \ \ v \in \mathcal{D}_l \tag{3.21}$$

$w_l := w_{l-1} + v_l$
new iterate: $u_j^{\nu+1} := w_m$

It is the main result of this section that local linearization can be regarded as local damping.

Lemma 3.1 *For a given intermediate iterate $w_{l-1} \in \mathcal{S}_j$ the local corrections $\bar{v}_l$ and v_l resulting from the subproblems (3.2) and (3.21), respectively, are related by*

$$v_l = \omega_l \bar{v}_l \tag{3.22}$$

with some $\omega_l \in [0,1]$.

Proof: If $\bar{v}_l \in \mathcal{D}_l$, then the inclusion $\mathcal{D}_l \subset \mathcal{D}_l^*$ yields $\bar{v}_l = v_l$. As $\bar{v}_l \in V_l$ and $\underline{\psi}_l$, $\overline{\psi}_l \in V_l$, we only have to consider the remaining cases $\bar{v}_l < \underline{\psi}_l$ and $\overline{\psi}_l < \bar{v}_l$. In the first case, (3.20) gives $\bar{v}_l < v_l = \underline{\psi}_l \leq 0$. The second case can be treated in a similar way. ■

Lemma 3.1 implies that Algorithm 3.2 is a special case of Algorithm 3.1. In particular, it is globally convergent. By keeping the local coarse grid corrections v_l in $\mathcal{D}_l$, the damping parameters ω_l are *implicitly* selected in such a way that the local linearization (3.11) remains valid. A similar approach can be used, if the functional ϕ_j is not piecewise linear but piecewise smooth. This will be the subject of a forthcoming paper.

4 Standard Monotone Multigrid Methods

To complete the construction of a monotone multigrid method, we now derive suitable local obstacles $\underline{\psi}_l$ and $\overline{\psi}_l$, $l = n_j + 1, \ldots, m$. For symmetry reasons, it is sufficient to consider only the upper obstacles $\overline{\psi}_l$. The construction relies on suitable successive restrictions of the initial defect obstacle $\overline{\varphi}_{\bar{u}_j^\nu} - \bar{u}_j^\nu$.

To identify the supporting points and the levels of $\lambda_l \in \Lambda_{\mathcal{S}}$, we will use the notation

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

Then the correction

$$v^{(k)} = v_{p_1}^{(k)} + \ldots + v_{p_{n_k}}^{(k)}$$

is the sum of all local corrections $v_{l_{ik}} = v_{p_i}^{(k)}$ in direction of the basis functions $\lambda_{l_{ik}} = \lambda_{p_i}^{(k)}$ on level k. The following lemma is easily proved by induction.

Lemma 4.1 *Assume that the mappings $R_{k+1}^k : \mathcal{S}_{k+1} \to \mathcal{S}_k$, $k = j-1, \ldots, 0$, are monotone in the sense that*

$$0 \le R_{k+1}^k v(p) \le v(p), \quad p \in \mathcal{N}_{k+1}, \tag{4.1}$$

holds for all non–negative $v \in \mathcal{S}_{k+1}$. Then, for a given smoothed iterate $\bar{u}_j^\nu$ and the initial defect obstacle $\overline{\psi}^{(j)} = \overline{\varphi}_{\bar{u}_j^\nu} - \bar{u}_j^\nu \ge 0$, the recursive restriction

$$\overline{\psi}^{(k)} = R_{k+1}^k(\overline{\psi}^{(k+1)} - v^{(k+1)}), \quad k = j-1, \ldots, 0, \tag{4.2}$$

provides local upper obstacles $\overline{\psi}_l \in V_l$ with the property (3.20) by the definition

$$\overline{\psi}_{l_{ik}} = \overline{\psi}^{(k)}(p_i)\lambda_{p_i}^{(k)}, \quad i = 1, \ldots, n_k. \tag{4.3}$$

As we are interested in multigrid convergence rates, we want to exclude the trivial choice $R_{k+1}^k \equiv 0$ which would bring back the single grid relaxation. Hence, we will now derive monotone restrictions R_{k+1}^k satisfying

$$\min\{v(q) \mid q \in \mathcal{N}_{k+1} \cap \text{ int supp } \lambda_p^{(k)}\} \le R_{k+1}^k v(p), \quad p \in \mathcal{N}_k, \tag{4.4}$$

for non–negative $v \in \mathcal{S}_{k+1}$, instead of the weaker lower estimate in (4.1). It will turn out later on that such *quasioptimal* restrictions provide asymptotic multigrid convergence rates. Let us select a certain ordering of the edges $\mathcal{E}_k = \{e_1, \ldots, e_s\}$ with midpoints $p_e \in \mathcal{N}_{k+1}$, $e \in \mathcal{E}_k$. Then the restriction operator $R_{k+1}^k : \mathcal{S}_{k+1} \to \mathcal{S}_k$ is defined by

$$R_{k+1}^k v = I_{\mathcal{S}_k} \circ R_{e_s} \circ \ldots \circ R_{e_1} v, \quad v \in \mathcal{S}_{k+1}. \tag{4.5}$$

Here $I_{\mathcal{S}_k}$ denotes the $\mathcal{S}_k$–interpolation and the operators $R_e : \mathcal{S}_{k+1} \to \mathcal{S}_{k+1}$, $e \in \mathcal{E}_k$, are of the form

$$R_e v = v + v_1 \lambda_{p_1}^{(k+1)} + v_2 \lambda_{p_2}^{(k+1)}, \quad v \in \mathcal{S}_{k+1}, \tag{4.6}$$

with $p_1, p_2 \in \mathcal{N}_k$ denoting the vertices of $e = (p_1, p_2) \in \mathcal{E}_k$. The scalars $v_1, v_2 \in \mathbb{R}$ in (4.6) are chosen such that

$$R_e v(p) \leq v(p), \quad p = p_1, p_e, p_2.$$

In particular, we set $v_1 = 0$, if $v(p_1) \leq v(p_e)$ or $v(p_1) + v(p_2) \leq 2v(p_e)$. In the remaining case, v_1 is determined by

$$v_1 = \begin{cases} 2v(p_e) - v(p_1) - v(p_2), & \text{if} \quad v(p_2) \leq v(p_e) \leq v(p_1), \\ v(p_e) - v(p_1), & \text{if} \quad v(p_e) \leq v(p), \ p = p_1, p_2. \end{cases}$$

The value of v_2 is obtained in a symmetrical way.

The following proposition can be checked by elementary considerations.

Proposition 4.1 *For any fixed enumeration of $\mathcal{E}_k$ the definition (4.5) provides a quasioptimal upper restriction operator R_{k+1}^k in the sense of (4.4).*

We will now formulate Algorithm 3.2 as a multigrid V–cycle. For this reason, we rewrite the computation of the correction $v^{(k)}$ from all local subproblems (3.21) on a fixed level k as one step of a *projected Gauss–Seidel–method.* The corresponding iteration operator for a bilinear form $a = a(\cdot, \cdot)$, a right hand side r, and obstacles $\underline{\psi}$, $\overline{\psi}$ is denoted by $\bar{\mathcal{M}}_k(a, r, \underline{\psi}, \overline{\psi}) : \mathcal{S}_k \to \mathcal{S}_k$. Recall that $\mathcal{M}_j : \mathcal{S}_j \to \mathcal{S}_j$ stands for the *nonlinear single grid relaxation.* Lower and upper monotone restrictions will be denoted by $\underline{R}_{k+1}^k$ and $\overline{R}_{k+1}^k$, respectively.

Algorithm 4.1 (Standard Monotone Multigrid Method)
fine grid smoothing: $\bar{u}_j^\nu := \mathcal{M}_j(u_j^\nu)$
local linearization: $a_{\bar{u}_j^\nu} := a + b_{\bar{u}_j^\nu}, \quad \ell_{\bar{u}_j^\nu} := \ell + f_{\bar{u}_j^\nu}$
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: $\underline{\psi}^{(j)} := \underline{\varphi}_{\bar{u}_j^\nu} - \bar{u}_j^\nu, \quad \overline{\psi}^{(j)} := \overline{\varphi}_{\bar{u}_j^\nu} - \bar{u}_j^\nu$
    global correction: $v_j^\nu := 0$
  for $k = j - 1$ *step* -1 *until* 0 *do*
    canonical restrictions: $a^{(k)} := a^{(k+1)}|_{\mathcal{S}_k \times \mathcal{S}_k}, \quad r^{(k)} := r^{(k+1)}|_{\mathcal{S}_k}$
    quasioptimal restrictions: $\underline{\psi}^{(k)} := \underline{R}_{k+1}^k \underline{\psi}^{(k+1)}, \quad \overline{\psi}^{(k)} := \overline{R}_{k+1}^k \overline{\psi}^{(k+1)}$

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

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

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

Note that Algorithm 4.1 contains a slightly improved variant of Mandels method [23] for linear complementary problems as a special case. See [18] for details.

We will now briefly sketch that quasioptimal restrictions $\underline{R}_{k+1}^k$, $\overline{R}_{k+1}^k$ lead to *asymptotic multigrid convergence rates*. It can be shown that for large ν the discrete phases of the iterates u_j^ν are equal to the discrete phases $\mathcal{N}_j^i(u_j)$, $i = 0, \ldots, N$, of the exact finite element solution u_j, if the discrete problem (2.7) is non–degenerate in the sense that

$$p \in \mathcal{N}_j^\bullet(u_j) \Rightarrow \ell(\lambda_p^{(j)}) - a(u_j, \lambda_p^{(j)}) \in \text{ int } \partial\phi_j(u_j)(\lambda_p^{(j)}). \tag{4.7}$$

Let us assume for the moment that the discrete phases of u_j are known. Then it is easily checked that $u_j = u_j^\circ$ is the unique solution of the *reduced linear problem*

$$u_j^\circ \in \bar{\mathcal{S}}_j^\circ : \quad a_{u_j}(u_j^\circ, v) = \ell_{u_j}(v), \quad v \in \mathcal{S}_j^\circ, \tag{4.8}$$

where the bilinear form a_{u_j} and the linear functional ℓ_{u_j} are defined in analogy to (3.17), $\bar{\mathcal{S}}_j^\circ = \{v \in \mathcal{S}_j \mid v(p) = u_j(p),\ p \in \mathcal{N}_j^\bullet(u_j)\}$ and the reduced subspace $\mathcal{S}_j^\circ \subset \mathcal{S}_j$ is given by

$$\mathcal{S}_j^\circ = \{v \in \mathcal{S}_j \mid v(p) = 0,\ p \in \mathcal{N}_j^\bullet(u_j)\}. \tag{4.9}$$

Observe that the *reduced multilevel basis*

$$\Lambda_{\mathcal{S}}^\circ = \Lambda_{\mathcal{S}} \cap \mathcal{S}_j^\circ \tag{4.10}$$

generates a splitting of $\mathcal{S}_j^\circ$ in one–dimensional subspaces which in turn gives rise to a corresponding multigrid method for (4.8). It is not difficult to see that for non–degenerate problems the undampened version of Algorithm 3.1 is asymptotically reducing to this multigrid method. The linearized Algorithm 4.1 has the same property (i.e. it asymptotically coincides with the "optimal" undampened Algorithm 3.1), if quasioptimal restrictions $\overline{R}_{k+1}^k$, $\underline{R}_{k+1}^k$ (cf. (4.4)) are used.

As a consequence, we can derive asymptotic estimates of the convergence rates of the nonlinear Algorithm 4.1 by investigating the corresponding reduced multigrid method for the linear problem (4.8). This can be done

by using recent results of Kornhuber and Yserentant [22], Oswald [25], and Griebel and Oswald [11].

Theorem 4.1 *The standard monotone multigrid method described in Algorithm 4.1 is globally convergent.*
Assume that the discrete problem (2.7) is non–degenerate in the sense of (4.7). Then the phases of the iterates $(u_j^\nu)_{\nu\geq 0}$ *converge to the phases of* u_j *and the error estimate*

$$\|u_j - u_j^{\nu+1}\| \leq (1 - c(j+1)^{-4})\|u_j - u_j^\nu\| \tag{4.11}$$

holds, if ν *is large enough. The positive constant* $c < 1$ *depends only on the ellipticity of* $a(\cdot,\cdot)$*, on the maximal coefficient* b_i*,* $i = 0, \ldots, N$*, of* Φ*, and on the initial triangulation* $\mathcal{T}_0$*.*

We emphasize that the estimate (4.11) describes the worst case. Absolutely no regularity assumptions on the continuous or discrete free boundary enter the constant c. In addition, we have considered the most simple variant of standard monotone multigrid methods. By repeating the (approximate) optimization in the direction of the basis functions $\lambda_p^{(k)}$ on each level $k = j, \cdots, 0$ in reversed order, we obtain a *standard monotone multigrid method with symmetric smoother*. For this variant, we get a $\mathcal{O}(j^2(\log j)^2)$ estimate. We can further improve this bound by imposing regularity conditions on $a(\cdot,\cdot)$ (providing $\mathcal{O}(j^2)$) or by using L^2–like projections instead of modified interpolation operators. In contrast to (4.11) the latter estimates also hold in the case of more than two space dimensions. However, we then need a certain regularity of the critical set $\mathcal{N}_j^\bullet(u_j)$. A detailed discussion can be found in [22, 25].

Let us now consider non–uniform refinement. In this situation, the canonical ordering of the multilevel nodal basis $\Lambda_\mathcal{S}$ would contradict our requirement that each multilevel relaxation should start with a fine grid relaxation step. Of course, one could rearrange $\Lambda_\mathcal{S}$ in a suitable way. For the implementation in an existing multigrid code it might be simpler to use the search directions $\Lambda = (\Lambda_j, \Lambda_\mathcal{S})$ instead of $\Lambda_\mathcal{S}$ or, equivalently, to start with a complete fine grid relaxation and then linearize all corrections in direction of $\lambda_l \in \Lambda_\mathcal{S}$. Both of these algorithms have the convergence properties stated in Theorem 4.1. The second algorithm will be used in our numerical experiment reported below.

5 Truncated Monotone Multigrid Methods

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

$$\text{int supp } \lambda_l \cap \mathcal{N}_j^\bullet(\bar{u}_j^\nu) \neq \emptyset \tag{5.1}$$

must not contribute to the coarse grid correction of the standard multigrid method. 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 (5.1) according to the actual guess of the free boundary. Again, it is sufficient to consider only uniform refinement. The non–uniform case can be treated in the same way as described above.

We define the modified basis functions

$$\tilde{\lambda}_p^{(k)} = T_{j,k}^\nu \lambda_p^{(k)}, \quad p \in \mathcal{N}_k, \tag{5.2}$$

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

$$T_{j,k}^\nu = I_{\mathcal{S}_j^\nu} \circ \ldots \circ I_{\mathcal{S}_k^\nu}. \tag{5.3}$$

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

$$\mathcal{S}_k^\nu = \{v \in \mathcal{S}_k \mid v(p) = 0,\ p \in \mathcal{N}_k^\nu\} \subset \mathcal{S}_k, \tag{5.4}$$

are the reduced subspaces with respect to $\mathcal{N}_k^\nu = \mathcal{N}_k \cap \mathcal{N}_j^\bullet(\bar{u}_j^\nu)$, $k = 0, \ldots, j$. 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. The resulting algorithm can be implemented as a variant of the standard monotone multigrid method. More precisely, in the neighborhood of the critical nodes $p \in \mathcal{N}_j^\bullet(\bar{u}_j^\nu)$ (cf. (3.9)) the restrictions and prolongations appearing in Algorithm 4.1 have to be modified as follows:

Modifications of Algorithm 4.1 (Truncated Monotone Multigrid Method)
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

Again, we can derive asymptotic estimates of the convergence rates by analysing the corresponding *reduced method* for the solution of the linear reduced problem (4.8). This time the reduced method is generated by the one–dimensional subspaces spanned by the truncation $\tilde{\Lambda}_{\mathcal{S}}^\circ \subset \mathcal{S}_j^\circ$ with respect to the exact critical set $\mathcal{N}_j^\bullet(u_j)$. Related algorithms have been recently considered by Hoppe and Kornhuber [16],Bank and Xu [1, 2], and Kornhuber and Yserentant [22].
Observe that $\tilde{\lambda}_p^{(k)} = \lambda_p^{(k)}$ holds for all $\lambda_p^{(k)} \in \mathcal{S}_j^\circ$, giving $\Lambda_{\mathcal{S}}^\circ \subset \tilde{\Lambda}_{\mathcal{S}}^\circ$. Hence, we can hope for improved asymptotic convergence rates of the truncated multigrid method as compared to the standard case. This is supported by the numerical results reported below. However, the theoretical analysis suffers from the fact that there is no strengthened Cauchy–Schwarz inequality for the spans of truncated basis functions $\tilde{\lambda}_p^{(k)} \notin \mathcal{S}_k$. Without any additional regularity this leads to even more pessimistic estimates than for the standard case.

Theorem 5.1 *The truncated monotone multigrid method is globally convergent.*
Assume that the discrete problem (2.7) is non–degenerate in the sense of (4.7). Then the phases of the iterates $(u_j^\nu)_{\nu\geq 0}$ *converge to the phases of* u_j *and the error estimate*

$$\|u_j - u_j^{\nu+1}\| \leq (1 - c(j+1)^{-6})\|u_j - u_j^\nu\| \tag{5.5}$$

holds, if ν *is large enough. The positive constant* $c < 1$ *depends only on the ellipticity of* $a(\cdot,\cdot)$*, on the maximal coefficient* b_i*,* $i = 1, \ldots, N$*, of* Φ*, and on the initial triangulation* $\mathcal{T}_0$*.*

As in the standard case, we can derive various improvements of the worst–case result (5.5). For example, we get a $\mathcal{O}(j^3)$ estimate, if symmetric smoothers are used.

6 Numerical Experiments

We will now illustrate the numerical performance of monotone multigrid methods in the framework of an adaptive algorithm. In this case, the underlying hierarchy of triangulations is resulting from the adaptive refinement of an initial triangulation $\mathcal{T}_0$. The adaptive refinement strategy and stopping criteria for the iterative solver on each refinement level are based on a posteriori estimates of the *approximation error* $\|u - \tilde{u}_j\|$ and of the *algebraic error* $\|u_j - \tilde{u}_j\|$ of some given $\tilde{u}_j \in \mathcal{S}_j$. A detailed description is contained in [20].

We consider the following model problem involving a jump discontinuity of Stefan type together with an upper obstacle. We choose the bilinear form $a(\cdot,\cdot)$ and the functional ℓ according to

$$a(v,w) = \int_\Omega (\partial_1 v \partial_1 w + \partial_2 v \partial_2 w)\, dx, \quad \ell(v) = \int_\Omega f\, v\, dx$$

with a peak source

$$f(x_1,x_2) = 3000 x_1 x_2 (x_1 - 1)(x_2 - 1) exp\left(-10(0.5 - x_1)^2 (0.5 - x_2)^2\right),$$

and $\Omega = (0,1) \times (0,1)$. The scalar function Φ defined in (2.4) is given by the parameters $N = 1$, $\theta_0 = -\infty$, $\theta_1 = 0.5$, $\theta_2 = 0.75$, and $b_0 = 400$, $f_0 = 200$, $c_0 = 50$, $b_1 = 0$, $f_1 = -100$, $c_1 = -50$.

The initial triangulation $\mathcal{T}_0$ is obtained by subdividing Ω in four congruent triangles. We now apply the adaptive algorithm as described in [20], using the truncated monotone multigrid method as iterativer solver. On each refinement level j the discrete problem is solved up to an (estimated) accuracy of 0.5% in order to obtain the approximate finite element solution $\tilde{u}_j \in \mathcal{S}_j$. The whole adaptive algorithm stops as soon as the (estimated) approximation error $\|u - \tilde{u}_j\|$ is less than 5%.

This final accuracy is reached after 8 adaptive refinement steps, providing the triangulation $\mathcal{T}_8$ together with the approximate solution $\tilde{u}_8$ as depicted in Figure 6. Observe the occurrence of a "mushy" region where $\tilde{u}_8 \equiv \theta_1$ and of a contact zone where $\tilde{u}_8 \equiv \theta_2$. Both are reflected by the adaptively refined mesh.

The complete approximation history is given in Table 6.1. Recall that the refinement depth is the maximal number of successive refinements. The effectivity index is the ratio of the a posteriori estimation and of a sufficiently accurate approximation of the exact error (cf. e.g. [4, 20]).

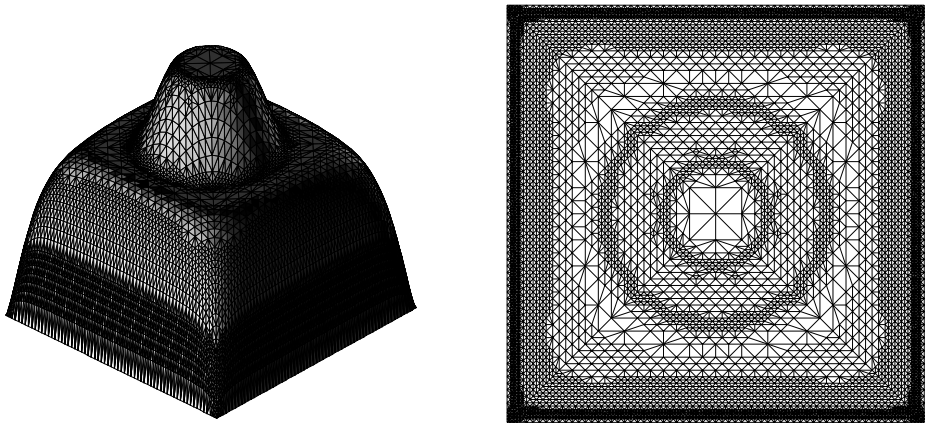

Figure 6.1: Final Approximation $\tilde{u}_8$ and Final Triangulation $\mathcal{T}_8$

Level	Depth	Nodes	Iterations	est. Approx. Error	Effectivity
0	0	1	2	7.6 %	0.22
1	1	5	2	16.8 %	1.21
2	2	25	2	14.7 %	1.24
3	3	77	2	13.0 %	1.42
4	4	277	3	10.4 %	1.69
5	5	733	3	8.33 %	1.90
6	6	2937	2	6.2 %	2.06
7	7	4413	2	5.6 %	2.06
8	7	7249	2	4.9 %	1.99

Table 6.1: Approximation History

From the moderate number of iterations on each refinement level it can be hardly perceived that we are dealing with a nonlinear problem. Only the severe underestimation of the error on the initial level indicates that it may be dangerous to start an adaptive algorithm from such a coarse mesh.

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 on the final triangulation $\mathcal{T}_8$. Starting with the initial iterate $u_8^0 = 0$, we obtain the algebraic errors $\|u_8 - u_8^\nu\|$, $\nu = 0, \ldots, 20$, as shown in Figure 6.

The overall convergence behavior can be divided into a *transient* phase, dominated by the search for the (discrete) free boundary, and an *asymptotic* phase, corresponding to the iterative solution of the reduced linear problem (4.8). As compared to the standard method STDKH, the truncated version TRCKH

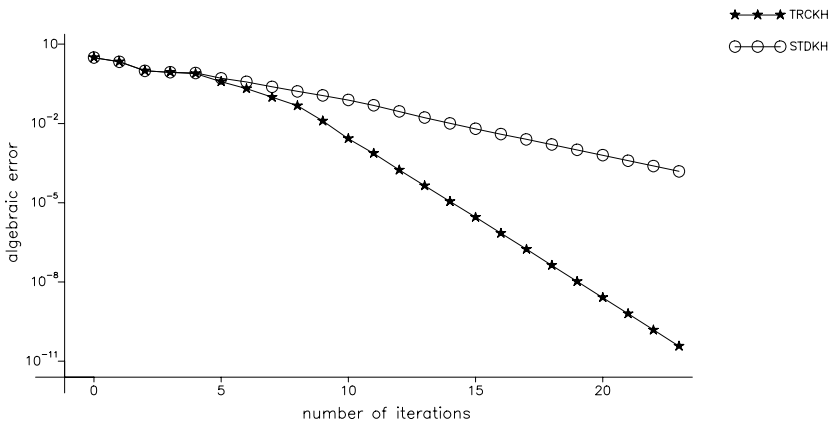

Figure 6.2: Iteration History: Initial Iterate $u_j^0 = 0$

exhibits a tremendous improvement of the asymptotic convergence rates, giving a numerical justification of the truncation of nodal basis functions. Note that the transient convergence properties remain basically the same.

Replacing the artificial initial iterate zero by the interpolation from the previous level, the transient phase is eliminated from the convergence history. This is illustrated by Figure 6.

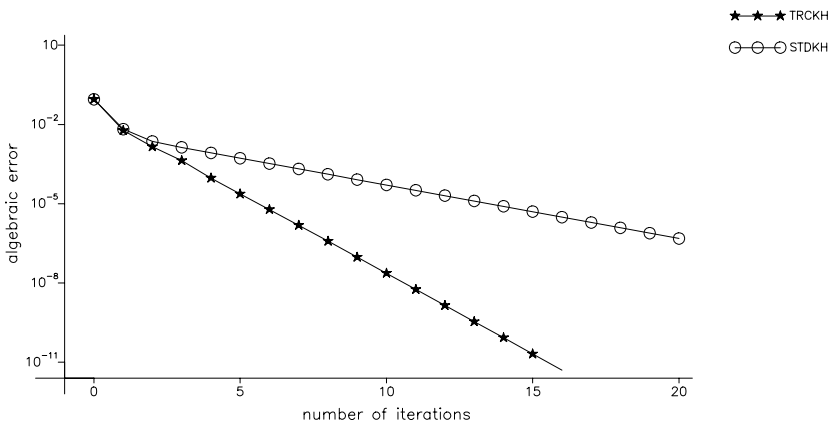

Figure 6.3: Iteration History: Interpolated Initial Iterate

To study the convergence properties for increasing j, we introduce the asymp-

totic efficiency rates ρ_j,

$$\rho_j = \sqrt[\nu_0]{\delta_j^{\nu_0}/\delta_j^0}, \quad j = 0, \ldots, 21, \tag{6.1}$$

where δ_j^ν denotes the algebraic error after ν iteration steps and the triangulations $\mathcal{T}_9, \ldots, \mathcal{T}_{21}$ are obtained by further adaptive refinement. We choose ν_0 such that $\delta_j^{\nu_0} < 10.^{-12}$. The results are shown in Figure 6.

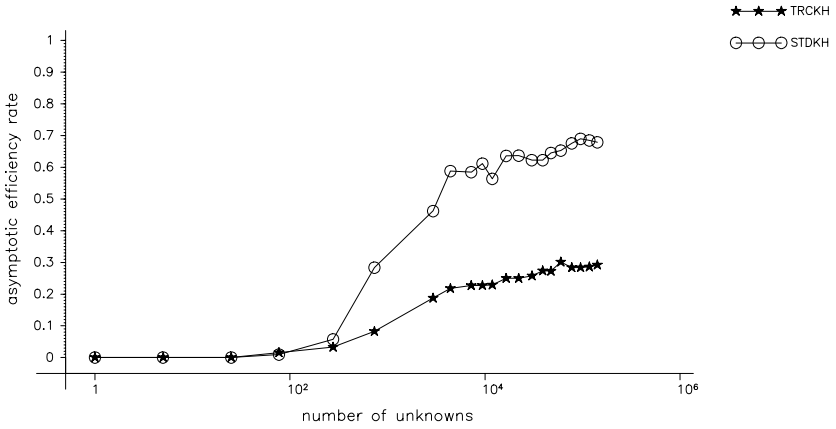

Figure 6.4: Asymptotic Efficiency Rates

The asymptotic efficiency rates for both multigrid methods seem to saturate with increasing j. This is better than predicted by the theoretical results. Even for the "bad" initial iterate zero, we observed uniform *global bounds* of the convergence rates. A theoretical verification of these experimental results will be the subject of future research.

Acknowledgements. The author thanks B. Erdmann from the Konrad–Zuse–Zentrum in Berlin for computational assistance.

Bibliography

[1] R.E. Bank and J. Xu. The hierarchical basis multigrid method and incomplete LU decomposition. In Keyes and Xu [17], pages 163–174.

[2] R.E. Bank and J. Xu. An algorithm for coarsening unstructured meshes. *Numer. Math.*, to appear.

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

[4] 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.*, to appear.

[5] F. Brezzi, W.W. Hager, and P.A. Raviart. Error estimates for the finite element solution of variational inequalities I. *Numer. Math.*, 28:431–443, 1977.

[6] J. Crank. *Free and Moving Boundary Problems.* Oxford University Press, Oxford, 1988.

[7] G. Duvaut and J.L. Lions. *Les inéquations en mécanique et en physique.* Dunaud, Paris, 1972.

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

[9] C.M. Elliot. Error analysis of the enthalpy method for the Stefan problem. *IMA J. Numer. Anal.*, 7:61–71, 1987.

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

[11] M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.*, submitted.

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

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

[14] K.-H. Hoffmann and J. Sprekels, editors. *Free Boundary Value Problems.*, Basel, 1990. Birkhäuser.

[15] R.H.W. Hoppe. A globally convergent multigrid algorithm for moving boundary problems of two-phase Stefan type. *IMA J. Numer. Anal.*, 13:235–253, 1993.

[16] R.H.W. Hoppe and R. Kornhuber. Adaptive multilevel methods for obstacle problems. *SIAM J. Numer. Anal.*, 31:301 – 323, 1994.

[17] D.E. Keyes and J. Xu, editors. *Proceedings of the 7th International Conference on Domain Decomposition Methods 1993.*, Providence, 1994. AMS.

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

[19] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities II. *Numer. Math.*, to appear.

[20] R. Kornhuber. A posteriori error estimates for elliptic variational inequalities. *Computers & Mathematics*, to appear.

[21] R. Kornhuber. *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems.* in preparation.

[22] R. Kornhuber and H. Yserentant. Multilevel methods for elliptic problems on domains not resolved by the coarse grid. In Keyes and Xu [17], pages 49–60.

[23] J. Mandel. A multilevel iterative method for symmetric, positive definite linear complementarity problems. *Appl. Math. Optimization*, 11:77–95, 1984.

[24] S.F. McCormick. *Multilevel Projection Methods for Partial Differential Equations.* Frontiers in Applied Mathematics, SIAM, Philadelphia, 1992

[25] P. Oswald. Stable subspace splittings for Sobolev spaces and domain decomposition algorithms. In Keyes and Xu [17], pages 87–98.

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

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