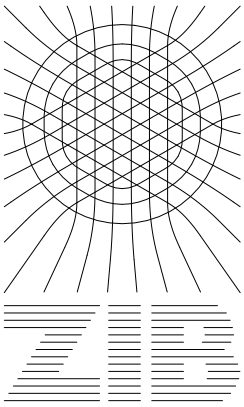
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Heilbronner Str. 10, D-10711 Berlin - Wilmersdorf



Ralf Kornhuber

Monotone Multigrid Methods for Elliptic Variational Inequalities II

Preprint SC 93–19 (März 1994)

Ralf Kornhuber

Monotone Multigrid Methods for Elliptic Variational Inequalities II

Abstract. We derive fast solvers for discrete elliptic variational inequalities of the second kind as resulting from the approximation by piecewise linear finite elements. Following the first part of this paper, monotone multigrid methods are considered as extended underrelaxations. Again, the coarse grid corrections are localized by suitable constraints, which in this case are fixed by fine grid smoothing. We consider the standard monotone multigrid method induced by the multilevel nodal basis and a truncated version. Global convergence results and asymptotic estimates for the convergence rates are given. The numerical results indicate a significant improvement in efficiency compared with previous multigrid approaches.

Key words: convex optimization, adaptive finite element methods, multigrid methods

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

Chapter 1

Introduction

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

$$u \in H_0^1(\Omega): \qquad \mathcal{J}(u) + \phi(u) \leq \mathcal{J}(v) + \phi(v), \quad v \in H_0^1(\Omega), \tag{1.1}$$

where 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 $H_0^1(\Omega)$–elliptic bilinear form $a(\cdot,\cdot)$ and a linear functional $\ell \in H^{-1}(\Omega)$. The convex functional ϕ of the form

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

is generated by a scalar convex function Φ. Denoting $z_- = \min\{z, 0\}$ and $z_+ = \max\{z, 0\}$ for $z \in \mathbb{R}$, then Φ is taken to be the piecewise quadratic convex function

$$\Phi(z) = \tfrac{1}{2} a_1 (z-\theta_0)_-^2 - s_1 (z-\theta_0)_- + \tfrac{1}{2} a_2 (z-\theta_0)_+^2 + s_2 (z-\theta_0)_+, \quad z \in \mathbb{R}, \tag{1.4}$$

with fixed $\theta_0 \in \mathbb{R}$ and non–negative constants $a_1, a_2, s_1, s_2 \in \mathbb{R}$. More general boundary conditions can be treated in the usual way.

It is well–known (c.f. Glowinski [8]) that (1.1) can be equivalently rewritten as the elliptic variational inequality of the second kind

$$u \in H_0^1(\Omega): \qquad a(u, v-u) + \phi(v) - \phi(u) \geq \ell(v-u)\,, \quad v \in H_0^1(\Omega), \tag{1.5}$$

and admits a unique solution $u \in H_0^1(\Omega)$. Note that (1.1) becomes a lower (or upper) obstacle problem, if s_1 (or s_2) tends to infinity.

Non–smooth optimization problems of the form (1.1) arise in a large scale of applications, ranging from friction problems or non–linear materials in elasticity to the spatial problems resulting from the implicit time–discretization of two–phase Stefan problems. Roughly speaking, the underlying physical situation is smooth in the different phases $u < \theta_0$ and $u > \theta_0$, respectively, but changes in a discontinuous way as u passes the threshold θ_0. We refer to Duvaut and Lions [4], Glowinski [8] and Elliot and Ockendon [7] for numerous examples and further information.

Let $\mathcal{T}_j$ be a given partition of Ω in triangles $t \in \mathcal{T}_j$ with minimal diameter of order 2^{-j}. The set of interior nodes is called $\mathcal{N}_j$. Discretizing (1.1) by

continuous, piecewise linear finite elements $\mathcal{S}_j \subset H_0^1(\Omega)$, we obtain the finite dimensional problem

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

Observe that the functional ϕ is approximated by $\mathcal{S}_j$–interpolation of the integrand $\Phi(v)$, giving

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

where $\Lambda_j = \{\lambda_p^{(j)}, \; p \in \mathcal{N}_j\}$ stands for the nodal basis in $\mathcal{S}_j$. Of course, (1.6) is uniquely solvable and can be reformulated as the variational inequality

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

For convergence results we refer to Elliot [6].

In this paper we will derive fast solvers for the discrete problem (1.6). Classical relaxation methods based on the successive optimization of the energy $\mathcal{J} + \phi_j$ in the direction of the nodal basis are discussed to some extend by Glowinski [8]. To overcome the well–known drawbacks of such single–grid relaxations, Hoppe and Kornhuber [15] have derived a multigrid algorithm, which was applied successfully to various practical problems [13, 16]. As a basic construction principle, the different phases must not be coupled by the coarse grid correction. Using advanced relaxation strategies of Hackbusch and Reusken [11, 12], Hoppe [14] recently derived a globally dampened version displaying a considerable improvement in asymptotic efficiency rates.

The construction of the previous multigrid methods was based on the full approximation scheme so that the possible implementation as a multigrid V–cycle was clear from the very beginning. However, suitable conditions for convergence were less obvious. Following the first part of this paper [18], we will derive monotone multigrid methods by extending the set of (high–frequent) search directions Λ_j by additional (intentionally low–frequent) search directions. As a consequence, our construction *starts* with a globally convergent method, which *then* is modified in such a way that the efficient implementation as a multigrid V–cycle becomes possible while the global convergence is retained. It is the main advantage of our approach that such modifications can be studied in an elementary way.

The corresponding theoretical framework will be derived in the next section. We formally introduce *extended relaxation methods* and describe so–called *quasioptimal approximations*, preserving the global convergence and asymptotically optimal convergence rates.

The actual construction of quasioptimal approximations takes place in Section 3. The reasoning is guided by the basic observation that the standard

V–cycle for linear problems relies on simple representations of linear operators and linear functionals on the coarse grid spaces. For *nonlinear* problems such (approximate) representations can be expected only *locally*. Consequently, the coarse-grid corrections of our monotone multigrid methods are obtained from certain obstacle problems, which are fixed by the preceding fine grid smoothing. In this way, the coupling of different phases *is not excluded*. Following the first part of this paper [18], we consider a *standard monotone multigrid method* and a *truncated* variant, relying on the multilevel nodal basis and its adaptation to the actual guess of the free boundary, respectively. Both methods can be regarded as permanent extensions of the classical multigrid method and of the corresponding algorithms presented in [18]. By construction, we obtain global convergence and the asymptotic convergence rates are bounded by $1 - \mathcal{O}(j^{-3})$.

In our numerical experiments reported in the final section, we basically found the same behavior as for obstacle problems (c.f. [18]). In particular, for good initial iterates as obtained by nested iteration, the overall convergence is dominated by the optimal asymptotic convergence rates, which are inherited from the related linear case. Compared to previous multigrid methods, this leads to a significant improvement in asymptotic efficiency.

Of course, our approach is not restricted to the special problem (1.6). We chose the very simple functional ϕ (and the related functionals ϕ_j) in order to keep the exposition as clear as possible. However, the basic convergence results to be presented extend without change to any functional ϕ_j of the form (1.7) with Φ replaced by arbitrary scalar, convex functionals Φ_p, $p \in \mathcal{N}_j$. For example, the restriction of the optimization (1.1) to a convex subset $\mathcal{K} \subset H_0^1(\Omega)$ of obstacle–type would cause no changes of the theoretical results and only minor modifications of the multigrid algorithms. If not explicitly otherwise stated, all our algorithmic considerations and convergence results are independent of the space dimension.

Chapter 2

Extended Relaxation Methods

Let $(M^\nu)_{\nu\geq 0}$ be a given sequence of finite subsets $M^\nu \subset \mathcal{S}_j$, $\nu \geq 0$, with the property

$$\Lambda_j = \{\lambda_p^{(j)} \mid p \in \mathcal{N}_j\} \subset M^\nu, \quad \nu \geq 0. \tag{2.1}$$

Recall that $\lambda_p^{(j)}$, $p \in \mathcal{N}_j$, denote the nodal basis functions of the given finite element space $\mathcal{S}_j$. Each set $M^\nu = \{\mu_1^\nu, \ldots, \mu_{m^\nu}^\nu\}$ is ordered in a suitable way and we assume that all functions μ_l^ν are non–negative, i.e. that

$$0 \leq \mu_l^\nu(p), \qquad p \in \mathcal{N}_j, \tag{2.2}$$

holds for all $\mu_l^\nu \in M^\nu$, $\nu \geq 0$. The elements of $M_c^\nu = M^\nu \setminus \Lambda_j$ are intended to play the role of *coarse grid functions* with large support, in contrast to the *fine grid functions* contained in Λ_j.

The *extended relaxation method* induced by $(M^\nu)_{\nu\geq 0}$ is resulting from the successive minimization of the energy $\mathcal{J}+\phi_j$ in the search directions $\mu_l^\nu \in M^\nu$. More precisely, we introduce the splitting

$$\mathcal{S}_j = \sum_{l=1}^{m^\nu} V_l^\nu, \quad V_l^\nu = \operatorname{span}\{\mu_l^\nu\}, \quad \nu \geq 0, \tag{2.3}$$

of $\mathcal{S}_j$ in the one–dimensional subspaces $V_l^\nu \subset \mathcal{S}_j$. Then, for a given iterate $u_j^\nu \in \mathcal{S}_j$, we compute a sequence of intermediate iterates w_l^ν, $l = 0, \ldots, m^\nu$, from the m^ν local subproblems

$$\begin{aligned} v_l^* \in V_l^\nu : \quad & \mathcal{J}(w_{l-1} + v_l^*) + \phi_j(w_{l-1} + v_l^*) \leq \\ & \leq \mathcal{J}(w_{l-1} + v) + \phi_j(w_{l-1} + v), \quad v \in V_l^\nu, \end{aligned} \tag{2.4}$$

setting $w_0^\nu = u_j^\nu$ and $w_l^\nu = w_{l-1}^\nu + v_l^*$, $l = 1, \ldots, m^\nu$. The next iterate is given by $u_j^{\nu+1} = w_{m^\nu}^\nu$.

Of course (2.4) is just the nonlinear multiplicative Schwarz method induced by the splitting (2.3). Observe that M^ν may change in each iteration step, so that the corresponding splitting can be iteratively adapted to the actual discrete free boundary. By construction, the extended relaxation (2.4) is monotone in the sense that

$$\mathcal{J}(w_l^\nu) + \phi_j(w_l^\nu) \leq \mathcal{J}(w_{l-1}^\nu) + \phi_j(w_{l-1}^\nu). \tag{2.5}$$

For notational convenience, the index ν will be frequently suppressed in the sequel.

Before investigating the convergence of extended relaxation methods, we will consider the (approximate) solution of the local subproblems (2.4). It is easily seen that (2.4) admits a unique solution and can be equivalently rewritten as the following variational inequality

$$\begin{aligned} v_l^* \in V_l: \quad a(v_l^*, v - v_l^*) + \phi_j(w_{l-1} + v) - \phi_j(w_{l-1} + v_l^*) \geq \\ \geq \ell(v - v_l^*) - a(w_{l-1}, v - v_l^*), \quad v \in V_l. \end{aligned} \tag{2.6}$$

This formulation avoids the derivative of the convex functional ϕ_j, which does not exist in the classical sense. However, using subdifferential calculus (c.f. Ekeland and Temam [5] or Clarke [2]), we can reformulate (2.4) as the differential inclusion

$$v_l^* \in V_l: \quad 0 \in a(v_l^*, v) + a(w_{l-1}, v) - \ell(v) + \partial\phi_j(w_{l-1} + v_l^*)(v), \quad v \in V_l. \tag{2.7}$$

Here, the subset $\partial\phi_j(w) \subset \mathcal{S}_j'$ denotes the set of subgradients of ϕ_j at $w \in \mathcal{S}$. Denoting $v_l^* = z_l^* \mu_l$, the inclusion (2.7) can be rewritten as the scalar differential inclusion

$$z_l^* \in \mathbb{R}: \qquad 0 \in a_{ll} z_l^* - r_l + \partial\Phi_l(z_l^*), \tag{2.8}$$

where we have used the definitions

$$a_{ll} = a(\mu_l, \mu_l), \qquad r_l = \ell(\mu_l) - a(w_{l-1}, \mu_l)$$

and $\partial\Phi_l(z) \subset \mathbb{R}$ denotes the subdifferential of the scalar convex function

$$\Phi_l(z) = \phi_j(w_{l-1} + z\mu_l), \quad z \in \mathbb{R}.$$

Recall that $\Phi_l = \Phi_l^\nu$ is depending on ν. Using the abbreviation $|p| = \int_\Omega \lambda_p^{(j)}(x)\, dx$ and the representation (1.7), we obtain

$$\Phi_l(z) = \sum_{p \in \mathcal{N}_j} \Phi(w_{l-1}(p) + z\mu_l(p))|p|, \quad z \in \mathbb{R}. \tag{2.9}$$

Exploiting (2.2), the subdifferential $\partial\Phi_l$ is a scalar, maximal monotone multifunction consisting of a weighted sum of translated subdifferentials of the given scalar, convex function Φ,

$$\partial\Phi_l(z) = \sum_{p \in \mathcal{N}_j} \mu_l(p)\, \partial\Phi(w_{l-1}(p) + z\mu_l(p))|p|, \quad z \in \mathbb{R}. \tag{2.10}$$

Note that the subdifferential $\partial\Phi$ is the maximal monotone multifunction

$$\partial\Phi(z) = \begin{cases} a_1(z - \theta_0) - s_1 & \text{if} \quad z < \theta_0 \\ [-s_1, s_2] & \text{if} \quad z = \theta_0. \\ a_2(z - \theta_0) + s_2 & \text{if} \quad z > \theta_0 \end{cases} \tag{2.11}$$

For fine grid functions $\mu_l = \lambda_{p_l}^{(j)} \in \Lambda_j$, the sum in (2.10) is reducing to

$$\partial\Phi_l(z) = \partial\Phi(w_{l-1}(p_l) + z\lambda_{p_l}^{(j)})|p_l|, \quad \lambda_{p_l}^{(j)} \in \Lambda_j.$$

Hence, the subdifferentials $\partial\Phi_l$ corresponding to coarse grid functions $\mu_l \in M_c$ are the sum of their fine grid counterparts. In multigrid terminology this means that the evaluation of the subdifferentials on coarse grids can be performed by canonical weighted restriction.

For fine grid functions $\mu_l = \lambda_{p_l}^{(j)} \in \Lambda_j$, the local problems (2.8) can be easily solved by means of

$$z_l^* = \theta_0 - w_{l-1}(p_l) + \begin{cases} (r_{p_l} + s_1)/(a_{p_l} + a_1), & r_{p_l} < -s_1 \\ 0, & r_{p_l} \in [-s_1, s_2] \\ (r_{p_l} - s_2)/(a_{p_l} + a_2), & r_{p_l} > s_2 \end{cases}, \qquad (2.12)$$

denoting

$$a_{p_l} = a_{ll}/|p_l|, \quad r_{p_l} = (r_l - a_{ll}(\theta_0 - w_{l-1}(p_l)))/|p_l|.$$

The situation is more difficult if $\mu_l \in M_c$. The main reason is that the number of critical values of $\partial\Phi_l$, where $\partial\Phi_l$ is set–valued, is growing with the number of nodes $p \in \mathcal{N}_j \cap$ int supp μ_l. Recall that supp μ_l is assumed to be large for $\mu_l \in M_c$. This motivates the approximation of $\partial\Phi_l$ by scalar multifunctions $\partial\Psi_l$ for $\mu_l \in M_c$. In abuse of our preceding notation, the multifunctions $\partial\Psi_l$ do not need to be subdifferentials.

Assume that $\partial\Psi_l$ is maximal monotone on $D_l \subset \mathbb{R}$, $D_l \neq \emptyset$. Then D_l must be a (possibly degenerated) interval. If D_l is bounded from above, say sup $D_l = z_0$, then sup $\partial\Psi_l(z)$ tends to ∞ as z tends to z_0. Hence, we formally set $\partial\Psi_l(z) = \infty$ for all $z \notin D_l$, $z \geq z_0$. In the same way, we extend $\partial\Psi_l$ by $-\infty$, if D_l is bounded from below.

A maximal monotone multifunction $\partial\Psi_l$ is called a *monotone approximation* of $\partial\Phi_l$, if

$$\begin{aligned} \sup\, \partial\Psi_l(z) &\geq \sup\, \partial\Phi_l(z), \qquad z \geq 0, \\ \inf\, \partial\Psi_l(z) &\leq \inf\, \partial\Phi_l(z), \qquad z \leq 0. \end{aligned} \qquad (2.13)$$

In particular, we have $\partial\Phi_l(0) \subset \partial\Psi_l(0)$. This motivates the trivial choice $\partial\Psi_l = \partial\Psi_\infty$, with $\partial\Psi_\infty(0) = (-\infty, \infty)$ defined on $D_\infty = \{0\}$. As a further example, consider the finite differences $\partial\Psi_l(z) = (\Phi_l((q+1)z) - \Phi_l(z))/(qz)$, with some fixed $q \neq 0$, providing a monotone approximation for $z \neq 0$. Other variants of practical interest will be described in the next section.

The approximations $\partial\Psi_l$, $\mu_l \in M_c$, give rise to the approximate subproblems

$$z_l \in \mathbb{R}: \qquad 0 \in a_{ll}z_l - r_l + \partial\Psi_l(z_l), \qquad \mu_l \in M_c. \qquad (2.14)$$

The resulting approximate coarse grid corrections are given by $v_l = z_l\mu_l$.

We will need the following location principle, which can be shown by standard arguments from convex analysis.

Lemma 2.1 *Assume that F is a scalar, strongly maximal monotone multifunction on $D_F \subset \mathbb{R}$, which is extended to $\mathbb{R} \setminus D_F$ as described above. Let $[z_0, z_1] \subset \mathbb{R}$ and*

$$\inf F(z_0) \leq 0 \leq \sup F(z_1).$$

Then there is a unique $\xi \in [z_0, z_1]$, such that $0 \in F(\xi)$.

If $\partial\Psi_l$ is a monotone approximation, then Lemma 2.1 applied to

$$F(z) = a_{ll} z - r_l + \partial\Psi_l(z), \quad z \in \mathbb{R}, \tag{2.15}$$

shows that the approximate subproblem (2.14) admits a unique solution z_l. We now generalize a related result from the first part of this paper [18].

Lemma 2.2 *Assume that $\partial\Psi_l$ is a monotone approximation of $\partial\Phi_l$. Then the corrections v_l^* and v_l, computed from (2.8) and (2.14), respectively, are related by*

$$v_l = \omega_l v_l^*, \quad \omega_l \in [0, 1]. \tag{2.16}$$

Proof. We will make use of the strongly maximal multifunction $F(z)$ defined in (2.15). Assume that the solution z_l^* of (2.8) is non–negative. Utilizing (2.13), we easily get

$$\inf F(0) \leq 0 \leq \sup F(z_l^*)$$

and Lemma 2.1 yields $0 \leq z_l \leq z_l^*$. In the remaining case, the assertion follows in a symmetrical way. ∎

An approximate scheme based on exact fine grid corrections v_l^*, $\mu_l \in \Lambda_j$, and dampened coarse grid corrections $v_l = \omega_l v_l^*$, $\omega_l \in [0, 1]$, $\mu_l \in M_c$, respectively, is called *extended underrelaxation*. Lemma 2.2 states that an extended underrelaxation is induced by a sequence of monotone approximations $(\partial\Psi_l^\nu)_{\nu \geq 0}$. Note that the classical single grid relaxation is recovered by the trivial choice $\partial\Psi_l = \partial\Psi_\infty$ for $\mu_l \in M_c$.

It follows from the convexity of $\mathcal{J} + \phi_j$ that extended underrelaxations preserve the monotonicity (2.5). The following Theorem is an immediate consequence of this property and the convergence of the fine grid relaxation.

Theorem 2.1 *An extended underrelaxation is globally convergent.*

We omit the proof, which can be almost literally taken from [18]. As a by–product, we obtain the convergence of the whole sequence of intermediate iterates w_l^ν,

$$w_l^\nu \to u_j, \quad \nu \to \infty. \tag{2.17}$$

We have described a general approach to construct convergent iterative schemes by selecting suitable search directions $(M^\nu)_{\nu\geq 0}$ and monotone approximations $(\Psi_l^\nu)_{\nu\geq 0}$. Note that only the representation (2.12) of the exact solution of the fine grid problems makes use of the actual choice of the scalar function Φ. As a consequence, Theorem 2.1 remains valid for all functionals ϕ_j of the form (1.7), which are represented by a family of arbitrary scalar, convex functions Φ_p, $p \in \mathcal{N}_j$.

In the remainder of this section, we will investigate the asymptotic behavior of extended underrelaxations. Denote

$$\mathcal{N}_j^\bullet(v) = \{p \in \mathcal{N}_j \mid v(p) = \theta_0\}, \quad v \in \mathcal{S}_j,$$

and $\mathcal{N}_l^\circ(v) = \mathcal{N}_j \setminus \mathcal{N}_j^\bullet(v)$. The critical points $p \in \mathcal{N}_j^\bullet(v)$ will take the role of the active points occurring in solution of obstacle problems. The discrete problem (1.6) is called *non–degenerate*, if

$$p \in \mathcal{N}_j^\bullet(u_j) \Rightarrow \ell(\lambda_p^{(j)}) - a(u_j, \lambda_p^{(j)}) \in \text{ int } \partial\phi_j(u_j)(\lambda_p^{(j)}). \tag{2.18}$$

This condition describes the stability of the critical nodes $\mathcal{N}_j^\bullet(u_j)$ with respect to small perturbations of u_j. The *discrete phases* $\mathcal{N}_j^-(v)$ and $\mathcal{N}_j^+(v)$ of a function $v \in \mathcal{S}_j$ consist of all nodes $p \in \mathcal{N}_j$ with $v(p) < \theta_0$ and $v(p) > \theta_0$, respectively.

We say that M^ν *is ordered from fine to coarse*, if $\mu_l = \lambda_{p_l}^{(j)}$ and $p_l \in$ int supp $\mu_{l'}^\nu$ implies $l < l'$ for all $\mu_l \in \Lambda_j$ and $\mu_{l'} \in M_c^\nu$. The sequence $(M^\nu)_{\nu\geq 0}$ is called *positive and bounded*, if there are positive constants c, C not depending on ν, such that

$$0 < c \leq \mu_l^\nu(p) \leq C, \qquad p \in \text{ int supp } \mu_l^\nu \cap \mathcal{N}_j, \;\; \mu_l^\nu \in M^\nu, \tag{2.19}$$

holds uniformly for $\nu \geq 0$. A positive, bounded sequence $(M^\nu)_{\nu\geq 0}$ is called *regular*, if $\mathcal{N}_j^\bullet(w_l^\nu) = \mathcal{N}_j^\bullet(u_j)$, $\nu \geq \nu_0$, implies that the sets M^ν also remain invariant for $\nu \geq \nu_0$.

Lemma 2.3 *Assume that the discrete problem (1.6) is non–degenerate. If $(M^\nu)_{\nu\geq 0}$ is positive, bounded and ordered from fine to coarse, then the phases of the intermediate iterates w_l^ν, $l = 1, \ldots, m^\nu$, resulting from an extended underrelaxation induced by $(M^\nu)_{\nu\geq 0}$, converge to the phases of u_j in the sense that*

$$\mathcal{N}_j^-(w_l^\nu) = \mathcal{N}_j^-(u_j), \;\; \mathcal{N}_j^\bullet(w_l^\nu) = \mathcal{N}_j^\bullet(u_j), \;\; \mathcal{N}_j^+(w_l^\nu) = \mathcal{N}_j^+(u_j) \tag{2.20}$$

holds for $\nu \geq \nu_0$, $l = 1, \ldots, m^\nu$, and some $\nu_0 \geq 0$.

Proof. It is easily seen that the convergence (2.17) of the whole sequence w_l^ν implies that there is a $\nu_1 \geq 0$ with the property

$$\mathcal{N}_j^-(u_j) \subset \mathcal{N}_j^-(w_l^\nu), \;\; \mathcal{N}_j^+(u_j) \subset \mathcal{N}_j^+(w_l^\nu), \qquad \nu \geq \nu_1. \tag{2.21}$$

Then, the assertion easily follows from the inclusion $\mathcal{N}_j^\bullet(u_j) \subset \mathcal{N}_j^\bullet(w_l^\nu)$ for large ν. This is what we are going to show now. As a first step, we derive the extended non–degeneracy condition

$$\ell(\mu_l^\nu) - a(u_j, \mu_l^\nu) \in I_l \subset \text{ int } \partial\phi_j(u_j)(\mu_l^\nu), \qquad \nu \geq 0, \tag{2.22}$$

for all $\mu_l^\nu \in M^\nu$ with the property int supp $\mu_l^\nu \cap \mathcal{N}_j^\bullet(u_j) \neq \emptyset$. The closed intervals $I_l \subset \mathbb{R}$ are defined by

$$I_l = \{z \in \mathbb{R} \,|\, |z - (\ell(\mu_l^\nu) - a(u_j, \mu_l^\nu))| \leq \varepsilon\}$$

and ε is independent of l or ν. Indeed, as a consequence of the non–degeneracy condition (2.18), we can find an $\varepsilon_j > 0$ such that (2.22) holds for all $\mu_l^\nu = \lambda_{p_l}^{(j)} \in \Lambda_j$. Taking the constant c from (2.19), it is easily checked that (2.22) is valid for all $\mu_l^\nu \in M^\nu$, if ε satisfies $0 < \varepsilon \leq c\, \varepsilon_j$.

Because $(M^\nu)_{\nu \geq 0}$ is bounded, the functionals $a(\cdot, \mu_l^\nu) \in \mathcal{S}_j'$ are uniformly bounded in l, ν. Hence, utilizing (2.22) and the convergence of w_l^ν, we can find a threshold $\nu_2 \geq \nu_1$ such that

$$\ell(\lambda_{p_l}^{(j)}) - a(w_l^\nu, \lambda_{p_l}^{(j)}) \in \text{ int } \partial\phi_j(u_j)(\lambda_{p_l}^{(j)}), \qquad \nu \geq \nu_2, \tag{2.23}$$

holds for all $p_l \in \mathcal{N}_j^\bullet(u_j)$. Consider some fixed $p_l \in \mathcal{N}_j^\bullet(u_j)$ and recall that w_l^ν is resulting from the fine grid correction associated with $\lambda_{p_l}^{(j)}$. This property can be rewritten as

$$\ell(\lambda_{p_l}^{(j)}) - a(w_l^\nu, \lambda_{p_l}^{(j)}) \in \partial\phi_j(w_l^\nu)(\lambda_{p_l}^{(j)}). \tag{2.24}$$

Using the representation $\partial\phi_j(w)(\lambda_p^{(j)}) = \partial\Phi(w(p))|p|$, $w \in \mathcal{S}_j$, and the monotonicity of $\partial\Phi$, it follows from (2.23) and (2.24) that $w_l^\nu(p_l) = u_j(p_l) = \theta_0$.

Hence, the fine grid correction makes sure that for large ν each critical point of u_j is a critical point of the corresponding intermediate iterate. We still have to show that these critical points are not affected by the coarse grid correction, i.e. that

$$\text{int supp } \mu_l^\nu \cap \mathcal{N}_j^\bullet(u_j) \neq \emptyset \Rightarrow v_l^* = v_l = 0, \qquad \nu \geq \nu_3, \tag{2.25}$$

holds for $\mu_l^\nu \in M_c^\nu$ and a suitable $\nu_3 \geq \nu_2$. Let $\mu_l^\nu \in M_c^\nu$ and int supp $\mu_l^\nu \cap \mathcal{N}_j^\bullet(u_j) \neq \emptyset$. As $(M^\nu)_{\nu \geq 0}$ is ordered from fine to coarse, we can assume inductively that the values of w_{l-1}^ν in $p \in$ int supp $\mu_l^\nu \cap \mathcal{N}_j^\bullet(u_j)$ were fixed to θ_0 by the preceding fine grid corrections and were not changed by possible

preceding coarse grid corrections. In this case, we can use (2.22) and the continuity of the derivative $\partial\Phi(z)$ in $z \neq \theta_0$ to find a $\nu_3 \geq \nu_2$ such that

$$\ell(\mu_l^\nu) - a(w_{l-1}^\nu, \mu_l^\nu) \in \partial\phi_j(w_{l-1}^\nu)(\mu_l^\nu), \quad \nu \geq \nu_3. \tag{2.26}$$

Using our 'scalar' notation (2.8), (2.26) can be rewritten as $r_l \in \partial\Phi_l(0)$, giving $z_l^* = 0$. This completes the proof. ∎

Once the correct phases

$$\mathcal{N}_j = \mathcal{N}_j^-(u_j) \cup \mathcal{N}_j^\bullet(u_j) \cup \mathcal{N}_j^+(u_j) \tag{2.27}$$

are known, we can define the bilinear form $b_{u_j}(v, w)$,

$$\begin{aligned} b_{u_j}(v,w) = \sum_{p \in \mathcal{N}_j^-(u_j)} & a_1 v(p) w(p) |p| \quad + \\ & + \sum_{p \in \mathcal{N}_j^+(u_j)} a_2 v(p) w(p) |p|, \quad v, w \in \mathcal{S}_j, \end{aligned} \tag{2.28}$$

and the functional $f_{u_j}(v)$,

$$\begin{aligned} f_{u_j}(v) = \sum_{p \in \mathcal{N}_j^-(u_j)} & (s_1 + a_1\theta_0) v(p) |p| \quad - \\ & - \sum_{p \in \mathcal{N}_j^+(u_j)} (s_2 - a_2\theta_0) v(p) |p|, \quad v \in \mathcal{S}_j. \end{aligned} \tag{2.29}$$

Denoting

$$a_{u_j}(v,w) = a(v,w) + b_{u_j}(v,w), \quad \ell_{u_j}(v) = \ell(v) + f_{u_j}(v), \tag{2.30}$$

it is easily checked that the desired solution u_j satisfies the variational equality

$$a_{u_j}(u_j, v) = \ell_{u_j}(v), \qquad v \in \mathcal{S}_j^\circ, \tag{2.31}$$

where the reduced subspace $\mathcal{S}_j^\circ \subset \mathcal{S}_j$ is defined by

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

If M^ν is regular and we asymptotically have $M^\nu = M^*$, then the reduced set

$$M^\circ = \{\mu \in M^* \mid \mu(p) = 0,\ p \in \mathcal{N}_j^\bullet(u_j)\} \subset M^*,$$

is inducing an extended relaxation method for the iterative solution of (2.31). The corresponding corrections $v_l^* \in V_l$ in the direction of $\mu_l \in M^\circ$ are computed from the linear local subproblems

$$v_l^* \in V_l: \qquad a_{u_j}(v_l^*, v) = \ell_{u_j}(v) - a(w_{l-1}, v), \quad v \in V_l. \tag{2.32}$$

Assuming that the original discrete problem (1.6) is non–degenerate, it is easily seen that an extended relaxation induced by a regular sequence $(M^\nu)_{\nu\geq 0}$ is asymptotically reducing to the linear scheme (2.32). In order to obtain a related result for extended underrelaxations, we have to impose further restrictions on the local approximations.

A sequence of monotone approximations $(\partial\Psi_l^\nu)_{\nu\geq 0}$ is called *quasioptimal*, if the convergence of the intermediate iterates w_l^ν and of their critical values $\mathcal{N}_j^\bullet(w_l^\nu)$ implies that there is a $\nu_0 \geq 0$ and an open interval $I \subset \mathbb{R}$, which contains 0 and is not depending on ν, l, such that

$$\partial\Psi_l^\nu(z) = \partial\Phi_l^\nu(z), \quad z \in I, \qquad \nu \geq \nu_0, \tag{2.33}$$

holds for all μ_l^ν with $\mu(p) = 0$, $p \in \mathcal{N}_j^\bullet(u_j)$.

Now we are ready to state the main result of this section.

Theorem 2.2 *Assume that the discrete problem (1.6) is non–degenerate. Then the extended underrelaxation induced by regular search directions $(M^\nu)_{\nu\geq 0}$ and quasioptimal local approximations $(\partial\Psi_l^\nu)_{\nu\geq 0}$ is reducing to the extended relaxation (2.32) for $\nu \geq \nu_0$ and some $\nu_0 \geq 0$.*

Proof. It follows from Lemma 2.3 that $\mathcal{N}_j^\bullet(w_l^\nu) = \mathcal{N}_j^\bullet(u_j)$ holds for $\nu \geq \nu_1$ and some suitable $\nu_1 \geq 0$. The exact local corrections $v_l^{*,\nu} = z_l^{*,\nu}\mu_l^\nu$ tend to zero. Hence, we can find a $\nu_0 \geq \nu_1$ so that $z_l^{*,\nu} \in I$, $\nu \geq \nu_0$. Then it follows from (2.33) that $z_l^\nu = z_l^{*,\nu}$, $\nu \geq \nu_0$. This completes the proof. ∎

Theorem 2.2 states that for non–degenerate problems all extended underrelaxations, which are induced by a fixed sequence $(M^\nu)_{\nu\geq 0}$ and various quasioptimal approximations, asymptotically coincide. This includes the original extended relaxation itself. In the case of good initial iterates ("good" with respect to the stability of the actual critical set $\mathcal{N}_j^\bullet(u_j)$), this optimal asymptotic behavior dominates the whole iteration process. We refer to the numerical experiments reported below.

Chapter 3

Monotone Multigrid Methods

Assume that $\mathcal{T}_j$ is resulting from j refinements of an intentionally coarse triangulation $\mathcal{T}_0$. In this way, we obtain a sequence of triangulations $\mathcal{T}_0, \ldots, \mathcal{T}_j$ and corresponding nested finite element spaces $\mathcal{S}_0 \subset \ldots \subset \mathcal{S}_j$. Though the algorithms and convergence results to be presented can be easily generalized to the non–uniform case, we assume for notational convenience that the triangulations are uniformly refined. More precisely, each triangle $t \in \mathcal{T}_k$ is subdivided in four congruent subtriangles in order to produce the next triangulation $\mathcal{T}_{k+1}$.

Collecting the nodal basis functions from all refinement levels, we define the multilevel nodal basis Λ,

$$\Lambda = \{\lambda_{p_1}^{(j)}, \lambda_{p_2}^{(j)} \ldots, \lambda_{p_{n_j}}^{(j)}, \ldots, \lambda_{p_1}^{(0)}, \ldots, \lambda_{p_{n_0}}^{(0)}\}, \tag{3.1}$$

with $m = n_j + \ldots + n_0$ elements. As indicated in (3.1), Λ is ordered from fine to coarse. An extended underrelaxation induced by a regular sequence $(M^\nu)_{\nu \geq 0}$ and quasioptimal local approximations $(\Psi_l^\nu)_{\nu \geq 0}$ is called *monotone multigrid method*, if the reduced multilevel nodal basis $\Lambda^\circ = \{\lambda \in \Lambda | \lambda(p) = 0, p \in \mathcal{N}_j^\bullet(u_j)\} \subset \Lambda$ is contained in the corresponding reduced set M°.

We first consider the constant search directions $M^\nu = \Lambda$, $\nu \geq 0$, with coarse grid functions given by $\Lambda_c = \Lambda \setminus \Lambda_j$. In this way, we will generalize the *standard monotone multigrid method* proposed in the first part of this paper [18]. It is clear that Λ is regular.

Due to the ordering of the search directions Λ, each iteration step starts with a fine grid smoothing of the given iterate u_j^ν, involving the search directions $\lambda_l \in \Lambda_j$. Recall that the corresponding local fine grid corrections can be easily computed from (2.12).

Then, we basically want to improve the resulting intermediate iterate $w_{n_j}^\nu$ by successive minimization of the energy $\mathcal{J} + \phi_j$ in the coarse grid directions $\lambda_l \in \Lambda_c$. To take advantage of the simple representation of linear operators and linear functionals on the coarse spaces $\mathcal{S}_k \subset \mathcal{S}_j$, $0 \leq k < j$, which is crucial for the optimal complexity of classical multigrid methods, we want to *restrict the scalar corrections z_l to such intervals, on which the subdifferentials $\partial \Phi_l^\nu(z) = \partial \phi_j(w_{l-1}^\nu + z\lambda_l)(\lambda_l)$ are linear.* In this case, we can evaluate the coarse grid corrections $v_l = z_l \lambda_l$ *without visiting the fine grid.*

Following this basic idea, we define the closed, convex subset $\mathcal{K}_j^\nu \subset \mathcal{S}_j$,

$$\mathcal{K}_j^\nu = \{v \in \mathcal{S}_j | \, \underline{\varphi}_j^\nu(p) \leq v(p) \leq \overline{\varphi}_j^\nu(p), \; p \in \mathcal{N}_j\},$$

where the obstacles $\underline{\varphi}_j^\nu, \overline{\varphi}_j^\nu \in \mathcal{S}_j$ are given by

$$\underline{\varphi}_j^\nu(p) = \begin{cases} -\infty, & w_{n_j}^\nu(p) < \theta_0 \\ \theta_0, & w_{n_j}^\nu(p) \geq \theta_0 \end{cases}, \quad \overline{\varphi}_j^\nu(p) = \begin{cases} \theta_0, & w_{n_j}^\nu(p) \leq \theta_0 \\ \infty, & w_{n_j}^\nu(p) > \theta_0 \end{cases} \tag{3.2}$$

for all $p \in \mathcal{N}_j$. As usual, the index ν will be frequently skipped in the sequel. By construction of the obstacles $\underline{\varphi}_j$ and $\overline{\varphi}_j$, the functional ϕ_j on $\mathcal{K}_j$ can be rewritten in the form

$$\phi_j(v) = \tfrac{1}{2} b_{w_{n_j}}(v, v) - f_{w_{n_j}}(v), \quad v \in \mathcal{K}_j. \tag{3.3}$$

The bilinear form $b_{w_{n_j}}(\cdot,\cdot)$ and the functional $f_{w_{n_j}}$ on $\mathcal{S}_j$ are defined by (2.28) and (2.29), respectively, replacing u_j by w_{n_j}. Observe that the underlying approximate splitting

$$\mathcal{N}_j = \mathcal{N}_j^-(w_{n_j}) \cup \mathcal{N}_j^\bullet(w_{n_j}) \cup \mathcal{N}_j^+(w_{n_j}) \tag{3.4}$$

is fixed by the fine grid smoothing.

We will impose the condition $w_l \in \mathcal{K}_j$ on the remaining intermediate iterates w_l, $l = n_j + 1, \cdots, m$. Equivalently, *the coarse grid corrections must not cause a change of phase*. In particular, the values $w_{n_j}(p) = \theta_0$ at the critical points $p \in \mathcal{N}_j^\bullet(w_{n_j})$ remain invariant. We emphasize, that *the coupling of the phases by the coarse grid correction is not excluded.*

The restricted successive minimization of the energy functional $\mathcal{J} + \phi_j$ on $\mathcal{K}_j$ in the directions $\lambda_l \in \Lambda_c$ leads to the same type of local obstacle problems as we have already considered in the first part of this paper [18]. Hence, we can directly apply all the arguments and algorithms presented therein.

In particular, the exact solution of the resulting local obstacle problems is still not available at reasonable cost. For an approximation we use quasioptimal local obstacles $\underline{\psi}_l$, $\overline{\psi}_l \in V_l = \operatorname{span}\{\lambda_l\}$ generated by monotone recursive restriction of the defect obstacles $\underline{\varphi}_j - w_{l-1}$, $\overline{\varphi}_j - w_{l-1} \in \mathcal{S}_j$. Introducing the bilinear form $a_{w_{n_j}}(\cdot,\cdot)$ and the functional $\ell_{w_{n_j}}$ on $\mathcal{S}_j$ according to (2.30) and the local constraints $\mathcal{D}_l \subset V_l$,

$$\mathcal{D}_l = \{v \in V_l \mid \underline{\psi}_l(p) \leq v(p) \leq \overline{\psi}_l(p),\ p \in \mathcal{N}_j\},$$

the (approximate) coarse grid corrections v_l are finally computed from

$$v_l \in \mathcal{D}_l: \quad a_{w_{n_j}}(v_l, v - v_l) \geq \ell_{w_{n_j}}(v - v_l) - a_{w_{n_j}}(w_{l-1}, v - v_l), \quad v \in \mathcal{D}_l, \tag{3.5}$$

for all $l = n_j + 1, \cdots, m$. Note that the resulting *standard monotone multigrid method* can be implemented as a classical V–cycle. We refer to [18] for details.

To apply the convergence theory developed in the preceding section, we reformulate (3.5) as a scalar inclusion of the form (2.14). For this reason, we define the scalar, convex functions Ψ_l,

$$\Psi_l(z) = \phi_j(w_{l-1} + z\lambda_l) + \chi_l(z), \quad z \in \mathbb{R}, \qquad \lambda_l \in \Lambda_c, \tag{3.6}$$

with χ_l denoting the characteristic function of $I_l = \{z \in \mathbb{R} \mid z\lambda_l \in \mathcal{D}_l\} \subset \mathbb{R}$. Then, it is easily checked that (3.5) can be reformulated as

$$z_l \in I_l : \qquad 0 \in a_{ll} z_l - r_l + \partial\Psi_l(z_l) \tag{3.7}$$

and $v_l = z_l\lambda_l$. Recall the notation $a_{ll} = a(\lambda_l, \lambda_l)$ and $r_l = \ell(\lambda_l) - a(w_{l-1}, \lambda_l)$.

Lemma 3.1 *The subdifferentials of the scalar functions $(\Psi_l^\nu)_{\nu\geq 0}$ defined in (3.6) are quasioptimal approximations $(\partial\Psi_l)_{\nu\geq 0}$.*

Proof. Consider some arbitrary, fixed $\nu \geq 0$ and a fixed l, $\lambda_l \in \Lambda_c$. Being monotone restrictions of the defect obstacles $\underline{\varphi}_j - w_{l-1}$ and $\overline{\varphi}_j - w_{l-1}$, the local defect obstacles $\underline{\psi}_l$ and $\overline{\psi}_l$ satisfy

$$\underline{\varphi}_j - w_{l-1} \leq \underline{\psi}_l \leq 0 \leq \overline{\psi}_l \leq \overline{\varphi}_j - w_{l-1}. \tag{3.8}$$

Hence, $0 \in I_l$. Now the monotonicity (2.5) follows from

$$\Psi_l(z) = \Phi_l(z) + \chi_l(z), \qquad z \in \mathbb{R}, \tag{3.9}$$

and simple arguments from convex analysis.

Assume that the intermediate iterates w_l^ν and their critical points $\mathcal{N}_j^\bullet(w_l^\nu)$ converge to u_j and $\mathcal{N}_j^\bullet(u_j)$, respectively. Choose $\nu_0 \geq 0$ such that $\mathcal{N}_j^\bullet(w_l^\nu) = \mathcal{N}_j^\bullet(u_j)$ for $\nu \geq \nu_0$, $l = 1, \ldots, m$. Then the obstacles $\underline{\varphi}_j^\nu$, $\overline{\varphi}_j^\nu$ and the corresponding constraints $\mathcal{K}_j^\nu$ remain invariant, say $\mathcal{K}_j^\nu = \mathcal{K}_j^*$ for $\nu \geq \nu_0$. It is easily checked that u_j is the solution of the double obstacle problem

$$u_j \in \mathcal{K}_j^* : \qquad a_{u_j}(u_j, v - u_j) \geq \ell_{u_j}(v - u_j), \qquad v \in \mathcal{K}_j^*.$$

Note that the corresponding active set of u_j coincides with the critical set $\mathcal{N}^\bullet(u_j)$. By the definition of the quasioptimality of $\underline{\psi}_l^\nu$, and $\overline{\psi}_l^\nu$ (c.f. [18, 19]), there is a positive number $\psi^* \in \mathbb{R}$ and a threshold $\nu_1 \geq \nu_0$, such that

$$\underline{\psi}_l^\nu(p) \leq -\psi^* < 0 < \psi^* \leq \overline{\psi}_l^\nu(p), \quad p \in \mathcal{N}_j \cap \text{int supp } \lambda_l, \quad \nu \geq \nu_1, \tag{3.10}$$

holds if λ_l is vanishing on $\mathcal{N}_j^\bullet(u_j)$. Setting $I = (-\psi^*, \psi^*)$, it is obvious that $0 \in I \subset I_l$ so that

$$\partial\Psi_l^\nu(z) = \partial\Phi_l^\nu(z), \quad z \in I, \quad \nu \geq \nu_1, \tag{3.11}$$

is valid for all l with int supp $\lambda_l \cap \mathcal{N}_j^\bullet(u_j) = \emptyset$. This completes the proof. ■

Exploiting recent estimates of the convergence rates for the linear reduced problem (c.f. [18, 19]), the following theorem is an immediate consequence of Lemma 3.1 and Theorem 2.2.

Theorem 3.1 *The standard monotone multigrid method induced by the local coarse grid problems (3.5) is globally convergent.*
If additionally the discrete problem (1.6) is non–degenerate, then the phases also converge and the a posteriori error estimate

$$\|u_j - u_j^{\nu+1}\| \leq (1 - c(j+1)^{-3})\|u_j - u_j^\nu\| \tag{3.12}$$

holds for $\nu \geq \nu_0$ with suitable $\nu_0 \geq 0$. Here $\|\cdot\|^2 = a(\cdot,\cdot)$ denotes the energy norm and the positive constant $c < 1$ depends only on the ellipticity of $a(\cdot,\cdot)$ and on the initial triangulation $\mathcal{T}_0$.

Note that the error estimate (3.12) requires no additional regularity assumptions. On the other hand, this result is restricted to two space dimensions. We refer to [18, 19] for a detailed discussion.

Obviously, there are no contributions from coarse grid functions $\lambda_l \in \Lambda_c \setminus \Lambda^\circ$, once the correct phases are fixed. However, the reduced splitting induced by Λ° may be rather poor, leading to unsatisfying asymptotic convergence rates (c.f. [18, 19]). Following [18], we will extend the set Λ° by suitable *truncations* of the coarse grid functions $\lambda_l \in \Lambda_c \setminus \Lambda^\circ$.

In each iteration step, we adapt Λ_c to the critical set $\mathcal{N}_j^\bullet(w_{n_j}^\nu)$ of the smoothened iterate $w_{n_j}^\nu$. More precisely, the actual coarse grid search directions $\tilde{\Lambda}_c^\nu$ are given by

$$\tilde{\Lambda}_c^\nu = \{\tilde{\lambda} \mid \tilde{\lambda} = T_{j,k}^\nu \lambda_p^{(k)},\ \lambda_p^{(k)} \in \Lambda_c,\ p \in \mathcal{N}_j \setminus \mathcal{N}_j^\bullet(w_{n_j})\}. \tag{3.13}$$

The truncation operators $T_{j,k}^\nu$,

$$T_{j,k}^\nu = I_{\mathcal{S}_j^\nu} \ldots I_{\mathcal{S}_{k+1}^\nu},\quad k = 0, \ldots, j-1, \tag{3.14}$$

are resulting from recursive $\mathcal{S}_k^\nu$–interpolation denoted by $I_{\mathcal{S}_k^\nu} : \mathcal{S}_j \to \mathcal{S}_k^\nu$. The reduced spaces $\mathcal{S}_k^\nu \subset \mathcal{S}_k$,

$$\mathcal{S}_k^\nu = \{v \in \mathcal{S}_k \mid v(p) = 0,\ p \in \mathcal{N}_k^\nu\},\quad k = 0, \ldots, j, \tag{3.15}$$

consist of the functions $v \in \mathcal{S}_k$ vanishing on the restricted critical sets $\mathcal{N}_k^\nu = \mathcal{N}_k \cap \mathcal{N}_j^\bullet(w_{n_j}^\nu)$, $k = 0, \ldots, j$. The ordering of $\tilde{\Lambda}_c^\nu = \{\tilde{\lambda}_{n_j+1}, \ldots, \tilde{\lambda}_{m^\nu}\}$ is inherited from Λ_c. It is easily checked that $\tilde{\Lambda}^\nu = \Lambda_j \cup \tilde{\Lambda}_c^\nu$, $\nu \geq 0$, is regular.

In particular, we have $\tilde{\Lambda}^\nu = \tilde{\Lambda}^\circ$, $\nu \geq \nu_0$, with some fixed $\tilde{\Lambda}^\circ$, if the phases remain invariant for $\nu \geq \nu_0$. Note that $\Lambda^\circ \subset \tilde{\Lambda}^\circ$ holds by construction.

As before, we use quasioptimal restrictions $\underline{\tilde{\psi}}_l$ and $\overline{\tilde{\psi}}_l$ of the defect obstacles $\underline{\varphi}_j - w_{l-1}$ and $\overline{\varphi}_j - w_{l-1}$ to define the local constraints $\tilde{\mathcal{D}}_l \subset \tilde{V}_l = \mathrm{span}\{\tilde{\lambda}_l\}$,

$$\tilde{\mathcal{D}}_l = \{v \in \tilde{V}_l \mid \underline{\tilde{\psi}}_l(p) \leq v(p) \leq \overline{\tilde{\psi}}_l(p),\ p \in \mathcal{N}_j\},\ \ \tilde{\lambda}_l \in \tilde{\Lambda}_c.$$

For all $\tilde{\lambda}_l \in \tilde{\Lambda}_c$, the coarse grid corrections $\tilde{v}_l$ are computed from

$$\tilde{v}_l \in \tilde{\mathcal{D}}_l :\ \ a_{w_{n_j}}(\tilde{v}_l, v - \tilde{v}_l) \geq \ell_{w_{n_j}}(v - \tilde{v}_l) - a_{w_{n_j}}(w_{l-1}, v - \tilde{v}_l),\ \ v \in \tilde{\mathcal{D}}_l, \quad (3.16)$$

In this way, we have derived a *truncated monotone multigrid method.*

The next theorem follows almost literally in the same way as Theorem 3.1.

Theorem 3.2 *The truncated monotone multigrid method induced by the local coarse grid problems (3.16) is globally convergent.*
If additionally the discrete problem (1.6) is non–degenerate, then the phases also converge and the a posteriori error estimate

$$\|u_j - u_j^{\nu+1}\| \leq (1 - c(j+1)^{-3})\|u_j - u_j^\nu\| \quad (3.17)$$

holds for $\nu \geq \nu_0$ with suitable $\nu_0 \geq 0$. The positive constant $c < 1$ depends only on the ellipticity of $a(\cdot,\cdot)$ and on the initial triangulation $\mathcal{T}_0$.

Both the standard and the truncated version can be implemented as a V–cycle with non–linear Gauss–Seidel smoothing (2.12) on the fine grid and projected Gauss–Seidel smoothing on the coarse levels. This carries over to the adaptive case. Other variants including W–cycles or symmetric Gauss–Seidel smoothing can be obtained in a similar way.

Chapter 4

Numerical Experiments

The non–linear evolution equation

$$\frac{\partial}{\partial t}\mathcal{H}(U) - \Delta U = F, \qquad \text{in } \Omega \times (0,T), \tag{4.1}$$

with suitable initial and boundary conditions describes the heat conduction in Ω undergoing a change of phase. $\mathcal{H}$ is a generalized enthalpy or heat content, U is a generalized temperature and F is a body heating term. The enthalpy $\mathcal{H}$ is a scalar maximal monotone multifunction,

$$\mathcal{H}(z) = \begin{cases} c_1(z-\theta_0)/\kappa_1 & \text{if} \quad z < \theta_0 \\ [0,L] & \text{if} \quad z = \theta_0, \\ c_2(z-\theta_0)/\kappa_2 + L & \text{if} \quad z > \theta_0 \end{cases} \qquad z \in \mathbb{R}, \tag{4.2}$$

which is set–valued at the phase change temperature θ_0. The positive constants c_i, κ_i, $i = 1,2$, describe the thermal properties in the two different phases and $L > 0$ stands for the latent heat.

Discretizing (4.1) in time by the backward Euler scheme with respect to a uniform step size $\tau > 0$, the spatial problems at the different time levels $t_k = k\tau$ can be identified with problems of the form (1.1). The solution $u = U_\tau(\cdot, t_k)$ is the approximation at the actual time step, the bilinear form $a(v,w) = \tau(\nabla v, \nabla w)$ is generated by the Laplacian and the functional ℓ is given by $\ell(v) = (\tau F_k + H_{k-1}, v)$ with $F_k = F(\cdot, t_k)$ and a suitable function $H_{k-1} \in \mathcal{H}(U_\tau(\cdot, t_{k-1}))$. The brackets $(\cdot,\cdot)$ denote the canonical scalar product in $L^2(\Omega)$. Finally, we choose $a_i = c_i/\kappa_i$, $i = 1,2$, and $s_1 = 0$, $s_2 = L$ so that the piecewise quadratic function Φ defined in (1.3) satisfies $\partial\Phi = \mathcal{H}$. This semi–discretization has been used by Jerome [17] to establish existence and uniqueness of the continuous solution U and also provides a general framework for a variety of numerical methods. We refer to Hoppe [14] and the literature cited therein.

To illustrate the numerical properties of our monotone multigrid methods, we will concentrate on a simple model problem, which has been already considered by Hoppe and Kornhuber [15] and Hoppe [14]. The space–time domain $\Omega \times (0,T)$ is specified by $\Omega = (0,1)^2$ and $T = 0.5$, while the physical data are $c_1 = 2$, $\kappa_1 = 1$, $c_2 = 6$, $\kappa_2 = 2$ and $\theta_0 = 0$, $L = 1$. Using the (physical) temperature θ,

$$\theta(x_1, x_2, t) = (x_1 - 0.5)^2 + (x_2 - 0.5)^2 - \ \exp(-4t)/4, \ \ (x_1, x_2) \in \Omega, \ t > 0,$$

the source term F is given by

$$F(x_1, x_2, t) = \begin{cases} c_1 \exp(-4t) - 4\kappa_1 & \text{if } \theta < 0 \\ c_2 \exp(-4t) - 4\kappa_2 & \text{if } \theta > 0 \end{cases}, \quad (x_1, x_2) \in \Omega,\ t > 0.$$

Then the generalized temperature U,

$$U = \kappa_1 \theta \ \text{ if } \theta \leq 0, \qquad U = \kappa_2 \theta \ \text{ if } \theta \geq 0,$$

is the solution of (4.1). Initial and boundary conditions taken from the exact solution U.

As in [14, 15], we choose the time step $\tau = 0.0125$. To obtain an initial triangulation $\mathcal{T}_0$, a partition of Ω in two triangles is regularly refined. Starting with $\mathcal{T}_0$, we apply successive uniform refinement to obtain a sequence of triangulations $\mathcal{T}_0, \ldots, \mathcal{T}_7$. The resulting discrete problems (1.6) are solved iteratively by the standard monotone multigrid method STDKH (c.f. Theorem 3.1) and the truncated version TRCKH (c.f. Theorem 3.2). The implementation was carried out in the framework of the finite element code KASKADE (c.f. Erdmann, Lang and Roitzsch [1]) and we used a SPARC IPX Workstation for the computations.

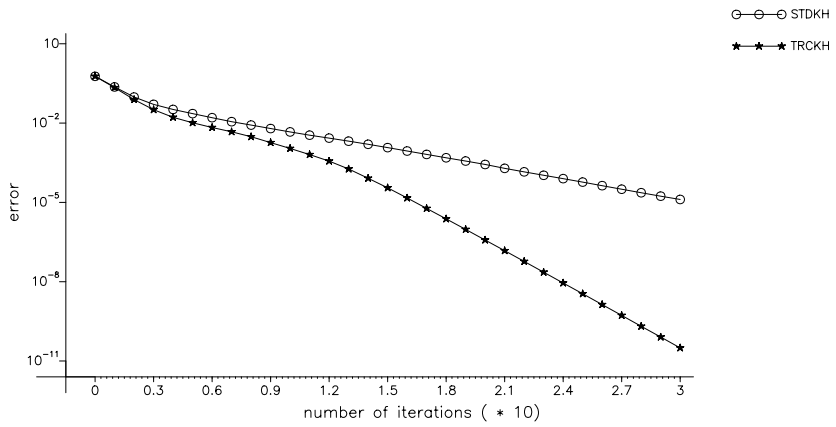


Figure 4.1: Iteration History

Let us consider the convergence behavior for the spatial problem resulting from the initial time step. In our first experiment the refinement level is fixed to $j = 6$ and we apply both multigrid methods to the initial iterate $u^0 = 0$. The resulting iterative errors with respect to the energy norm are depicted in Figure 4.1. Obviously, the iteration history can be separated in three different parts. First, we observe a rapid decrease due to the fast elimination

of the high frequent terms. In the following transient phase the algorithm determines the correct free boundary until finally the asymptotic behavior of the reduced linear iteration is reached. Obviously, TRCKH heavily benefits from the adaptive truncation of the standard search directions, providing a tremendous improvement of the asymptotic convergence rates.

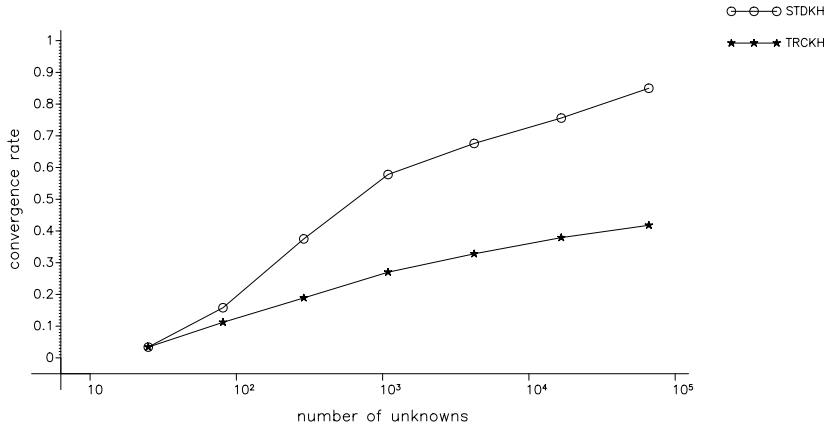


Figure 4.2: Asymptotic Convergence Rates

We now concentrate on the variation of the convergence behavior with increasing refinement level j. For the fixed initial iterate $u^0 = 0$ the transient convergence rates seem to be uniformly bounded but the number of transient steps grows considerably with increasing j. However, using reasonable initial iterates as resulting from nested iteration, we found that the transient steps were vanishing completely or (for large j) were reduced to a very small number. Starting with the interpolated solution from the previous level, we consider the asymptotic convergence rates ρ_j given by

$$\rho_j = \sqrt[\nu_0]{\varepsilon_j^{\nu_0} / \varepsilon_j^0}, \quad j = 0, \ldots, 7, \tag{4.3}$$

where ε_j^ν denotes the iterative error after ν iteration steps. To be compatible with [14, 15], the error is measured in the l^2–norm and we choose ν_0 such that $\varepsilon_j^{\nu_0} < 10.^{-8}$. The resulting asymptotic convergence rates of STDKH and TRCKH over the levels $j = 1, \ldots, 7$ are shown in Figure 4.2. Obviously, the convergence rates only slightly deteriorate with increasing j.

To compare TRCKH with previous multigrid methods, we consider the algorithms MGSTEF2 (c.f. [15]) and the dampened version DMGSTEF (c.f.[14]). As a basic construction principle of both methods, the coarse grid correction is restricted to the *interior* of the (approximate) phases, which have been

fixed by fine grid smoothing. In addition, DMGSTEF uses advanced relaxation strategies in the spirit of Hackbusch and Reusken [11, 12], leading to global convergence results and significantly improved asymptotic efficiency rates. The asymptotic efficiency rates q_j are obtained by multiplying the number ν_0 of iterations appearing in (4.3) by a certain work unit. A work unit corresponds to one symmetric Gauss–Seidel step on the finest level j. Table 1 below displays the resulting asymptotic efficiency rates q_5 for TRCKH, MGSTEF2 and DMGSTEF at the time levels $t = 10k\tau$, $k = 1, \ldots, 5$. The values for MGSTEF2 and DMGSTEF are taken from [14]. Similar results are obtained for the remaining time steps.

	t=0.10	t=0.20	t=0.30	t=0.40	t=0.50
TRCKH	0.20	0.23	0.21	0.19	0.19
DMGSTEF	0.34	0.31	0.33	0.30	0.29
MGSTEF2	0.50	0.45	0.50	0.44	0.43

Table 4.1: Asymptotic Efficiency Rates

Though we did not (yet) apply a suitable ordering of the unknowns or additional relaxation techniques, TRCKH performs best for all time levels. Unlike the other two methods, TRCKH allows the coupling of the phases by the (truncated) search directions. This leads to a larger coarse grid space, which is the reason for the improved convergence.

Acknowledgements. The author wants to thank R. Roitzsch for computational assistance.

Bibliography

[1] B. Erdmann, J. Lang, R. Roitzsch: *KASKADE Manual, Version 2.0* Technical Report TR 93–5, Konrad–Zuse–Zentrum Berlin (1993)

[2] F.H. Clarke: *Optimization and Nonsmooth Analysis.* John Wiley and Sons, New York (1983)

[3] P. Deuflhard, P. Leinen, H. Yserentant: *Concepts of an Adaptive Hierarchical Finite Element Code.* IMPACT Comput. Sci. Engrg. **1**, p. 3–35 (1989)

[4] G. Duvaut, J.L. Lions: *Les inéquations en mécanique et en physique.* Dunaud, Paris (1972)

[5] I. Ekeland, R. Temam: *Convex Analysis and Variational Problems.* North Holland, Amsterdam (1976)

[6] C.M. Elliot: *On the Finite Element Approximation of an Elliptic Variational Inequality Arising from an Implicit Time Discretization of the Stefan Problem.* IMA J. Numer. Anal. **1**, p. 115–125 (1981)

[7] C.M. Elliot, J.R. Ockendon: *Weak and Variational Methods for Moving Boundary Problems.* Research Notes in Mathematics **53**, Pitman, London (1982)

[8] R. Glowinski: *Numerical Methods for Nonlinear Variational Problems* Springer, New York (1984)

[9] R. Glowinski, J.L. Lions, R. Trémolières: *Numerical Analysis of Variational Inequalities.* North–Holland, Amsterdam (1981)

[10] W. Hackbusch: *Multi–Grid Methods and Applications.* Springer, Berlin (1985)

[11] W. Hackbusch, A. Reusken: Analysis of a Dampened Nonlinear Multilevel Method. Numer. Math. **55**, p. 225–246 (1989)

[12] W. Hackbusch, A. Reusken: On Global Multigrid Convergence for Nonlinear Problems. In: Robust Multigrid Methods, W. Hackbusch (ed.), Vieweg, Braunschweig, p. 105–113 (1988)

[13] R.H.W. Hoppe: *Numerical Solution of Multicomponent Alloy Solidification by Multigrid Techniques.* IMPACT Comput. Sci. Engrg. **2**, p. 125-156 (1990)

[14] R.H.W. Hoppe: *A Globally Convergent Multi–Grid Algorithm for Moving Boundary Problems of Two–Phase Stefan Type.* IMA J. Numer. Anal. **13**, p. 235–253 (1993)

[15] R.H.W. Hoppe, R. Kornhuber: *Multigrid Methods for the Two–Phase Stefan Problem.* In: Multigrid Methods: Theory, Applications and Supercomputing, St. McCormick (ed.), Marcel Dekker, New York, p. 267–297 (1988)

[16] R.H.W. Hoppe, R. Kornhuber: *Multi–Grid Solution of Two Coupled Stefan Equations Arising in Induction Heating of Large Steel Slabs.* Int. J. Numer. Methods Eng., Vol. **30**, p. 779–801 (1990)

[17] J.W. Jerome: *Approximation of Nonlinear Evolution Equations.* Academic Press, New York (1983)

[18] R. Kornhuber: *Monotone Multigrid Methods for Elliptic Variational Inequalities I.* To appear in Numer. Math. (1994)

[19] R. Kornhuber, H. Yserentant: *Multilevel Methods for Elliptic Problems on Domains not Resolved by the Coarse Grid.* Preprint SC 93–21, Konrad–Zuse–Zentrum Berlin (1993)

[20] J. Xu: *Iterative Methods by Space Decomposition and Subspace Correction.* SIAM Reviews **34**, p. 581–613 (1992)

[21] H. Yserentant: *Old and New Convergence Proofs of Multigrid Methods.* Acta Numerica, p. 285–326 (1993)