COMPUTATIONAL METHODS IN APPLIED MATHEMATICS, Vol.6(2006), No.2, pp.221–232

THE INFLUENCE OF HIGHER ORDER FEM DISCRETISATIONS ON MULTIGRID CONVERGENCE

M. KÖSTER[1] and S. TUREK[2]

Abstract — Quadratic and even higher order finite elements are interesting candidates for the numerical solution of partial differential equations (PDEs) due to their improved approximation properties in comparison to linear approaches. The systems of equations that arise from the discretisation of the underlying (elliptic) PDEs are often solved by iterative solvers like preconditioned Krylow-space methods, while multigrid solvers are still rarely used – which might be caused by the high effort that is associated with the realisation of the necessary data structures as well as smoothing and intergrid transfer operators.

In this note, we discuss the numerical analysis of quadratic conforming finite elements in a multigrid solver. Using the "correct" grid transfer operators in conjunction with a quadratic finite element approximation allows to formulate an *improved approximation property* which enhances the (asymptotic) behaviour of multigrid: If m denotes the number of smoothing steps, the convergence rates behave asymptotically like $\mathcal{O}(1/m^2)$ in contrast to $\mathcal{O}(1/m)$ for linear FEM.

2000 Mathematics Subject Classification: 65F10, 65N30, 65N55.

Keywords: higher order finite element method, multigrid, W-cycle.

1. Introduction

Multigrid methods for solving linear systems that arise from FEM discretisations for PDEs have been described and analysed by several authors since many years. There are "classical" multigrid proofs for W-cycle [5] and V-cycle [3] multigrid, for linear conforming as well as nonconforming [4] finite elements, for elliptic PDEs as well as for extensions like nonsymmetry, saddle-point formulations, nonlinear systems, etc. Of course, instead of linear finite elements, also higher order elements can be used for discretisation. But although the convergence of the multigrid method is often clear in this situation (the W-cycle proof of Hackbusch/Braess holds for all conforming finite elements of order $\geqslant 1$), it is not yet fully understood whether there are further quantitative effects for the convergence behavior.

In this note, we concentrate on quadratic FEM if applied to elliptic 2nd order PDEs. We modify the classical multigrid proof of Hackbusch/Braess to obtain a sharper result for this situation. Moreover, the analysis indicates that these results may be also valid for even

[1] *Institute of Applied Mathematics, University of Dortmund*, Vogelpothsweg 87, D-44227 Dortmund, Germany. E-mail: michael.koester@mathematik.uni-dortmund.de

[2] *Institute of Applied Mathematics, University of Dortmund*, Vogelpothsweg 87, D-44227 Dortmund, Germany. E-mail: stefan.turek@mathematik.uni-dortmund.de

higher order finite elements, leading to our conjecture that multigrid convergence rates might further improve for higher order elements which gives a new viewpoint to hp-FEM.

The remainder of this paper is organised as follows: In section 2 we introduce our notations. We formulate the smoothing property for the multigrid algorithm, as it was already formulated by other authors [1,2,9], and we repeat the key ingredients of the classical multigrid W-cycle proof. Section 3 specialises this proof for the situation of quadratic finite elements. Furthermore, we point out a possible generalisation for higher order finite elements. Finally, in section 4 we perform a numerical analysis of multigrid convergence rates which shows that the result of the proof is sharp.

2. Key ingredients of the multigrid proof

We start with investigating the classical W-cycle proof and show its key ingredients. However, we make some stronger assumptions here than in the original proof which will allow us to extend it to quadratic finite elements.

For our investigations, we consider a typical selfadjoint elliptic boundary value problem in a bounded domain $\Omega \subset \mathbb{R}^2$ with boundary $\partial\Omega$, for instance,

$$-\Delta u = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega. \tag{2.1}$$

Assumptions and notations. 1. Throughout this paper, c, $c' > 0$ denote generic constants that can vary from equation to equation. $(\cdot,\cdot)$ represents the standard L_2 product.

2. $\mathcal{T} = \{\mathcal{T}_h\}$, with $\mathcal{T}_h \subseteq \Omega$ with a mesh size parameter $h > 0$ denotes a family of uniform decompositions of Ω in the sense of [2], i.e., there is $\kappa > 0$ so that every $T \in \mathcal{T}_h$ contains a ball with radius $\rho_h \geqslant h/\kappa$ for every h.

3. Let $V := H_0^1(\Omega)$. Furthermore $\{V_h\}$ with $V_h \subset V$ denotes a nested family of affine conforming finite elements in the sense of [2]. In particular, $V_{2h} \subset V_h \subset V$; V_h has dimension $n = n_h$, two elements $T_1 \neq T_2,\ T_1, T_2 \in \mathcal{T}_h$ intersect at most in common corners or edges.

4. Let $a(\cdot,\cdot)$ be a positive definite, symmetric bilinear form, for instance,

$$a(v,w) = (\nabla v, \nabla w) \quad \forall\ v, w \in V,$$

that is H_0^1-coercive and continuous, i.e., there exist $c, \alpha > 0$ such that

$$|a(v,w)| \leqslant c\|v\|_1\,\|w\|_1, \quad a(v,v) \geqslant \alpha\|v\|_1^2 \quad \forall\, v, w \in H_0^1(\Omega).$$

By definition, the norm induced by this bilinear form $\|v\|_a := \sqrt{a(v,v)}$ is equivalent to the H^1-norm, i.e., with appropriate constants $c, c' > 0$ for all $v \in V$, the following estimate holds:

$$c\|v\|_1 \leqslant \|v\|_a \leqslant c'\|v\|_1 \quad \text{(or shorter: } \|v\|_1 \sim \|v\|_a).$$

5. The problem is to find a weak solution $u \in V$ of the boundary value problem

$$a(u,\varphi) = (f,\varphi) \quad \forall\ \varphi \in V \tag{2.2}$$

for a given $f \in L_2(\Omega)$. This problem is replaced by a discrete analogon. Find $u_h \in V_h$ such that

$$a(u_h,\varphi_h) = (f,\varphi_h) \quad \forall\ \varphi_h \in V_h. \tag{2.3}$$

6. In this section we assume the problem to be $H^2(\Omega)$-regular and linear finite elements to be used for the discretisation. In section 3, we assume the problem to be $H^3(\Omega)$-regular and quadratic finite elements are used.

We now formulate a typical multigrid algorithm which follows a two-grid approach in the usual way by recursively replacing the solution process of the coarse grid problem:

Algorithm 2.1 (One multigrid iteration step on V_h). *Purpose:* For an iterate u_h^0 and a right hand side f_h, compute a new iterate u_h^* approximating the solution $u_h \in V_h$. The following parameters configure the behavior of the algorithm:
(μ, ν) = number of presmoothing and postsmoothing steps, respectively;
p =cycle ($p = 2$ describes the W-cycle).
Let $\mathcal{S}_h : V_h \to V_h$ denote a special mapping, called *smoothing operator*, and $h_{\max}$ the resolution of the coarsest grid in the family $\{V_h\}$. Set $f_h := f$ on the finest mesh level h.

`FUNCTION MultigridCycle(`h, u_h^0, f_h`):` u_h^*
`BEGIN`

1. *Coarse grid solution:* If $h = h_{\max}$, then solve the coarse grid problem

$$a(u_{h_{\max}}, \varphi_{h_{\max}}) = (f_{h_{\max}}, \varphi_{h_{\max}}) \quad \forall\, \varphi_{h_{\max}} \in V_{h_{\max}}.$$

Return $u_h^* := u_{h_{\max}}$. Otherwise

2. *Presmoothing:* Compute $u_h^i := \mathcal{S}_h u_h^{i-1}$, $i = 1, \ldots, \mu$.
3. *Coarse grid correction:* Set up the *coarse grid problem* to find $u_{2h} \in V_{2h}$,

$$a(u_{2h}, \varphi_{2h}) = (f_{2h}, \varphi_{2h}) \quad \forall\, \varphi_{2h} \in V_{2h},$$

for the function $f_{2h} \in V$ defined by

$$(f_{2h}, \varphi_{2h}) = (f_h, \varphi_{2h}) - a(u_h^{\mu}, \varphi_{2h}) \quad \forall\, \varphi_{2h} \in V_{2h}.$$

Set $u_{2h}^0 := 0$. Solve the coarse grid problem approximately and recursively with p multigrid steps on the lower level,

$$u_{2h}^i := \texttt{MultigridCycle}\,(2h, u_{2h}^{i-1}, f_{2h}), \quad i = 1, \ldots, p,$$

and correct the current approximate solution

$$u_h^{\mu+1} := u_h^{\mu} + u_{2h}^{p}.$$

4. *Postsmoothing:* Compute $u_h^{\mu+1+i} := \mathcal{S}_h u_h^{\mu+i}$, $i = 1, \ldots, \nu$.

Return $u_h^* := u_h^{\mu+\nu+1}$.
`END MultigridCycle`

In the following, we only focus on the case of $\mu > 0$ and $\nu = 0$, thus ignoring any postsmoothing. Furthermore, we restrict ourselves to the case of a two-grid algorithm, as the multigrid proof for the W-cycle follows by a perturbation argument (see [1, 2, 4, 5]). The assumptions above allow to formulate a couple of (algebraic) statements that are independent of the finite element spaces, i.e., they hold for all conforming FEM spaces. We will briefly summarise the results in the following. For additional and more detailed information, see [1, 2, 5].

Definition 2.1 (Scale of norms). 1. We define a linear continuous operator $\mathcal{A}_h : V_h \to V_h$ by:

$$(\mathcal{A}_h v_h, w_h) = a(v_h, w_h) \quad \forall\, v_h, w_h \in V_h. \tag{2.4}$$

2. For $s \in \mathbb{R}$ and $v_h = \sum c_i \psi_h^i \in V_h$ with $\{\psi_h^i\} \subset V_h$ a set of orthonormal eigenfunctions to eigenvalues $\{\lambda_i\}$ of $\mathcal{A}_h$, using a spectral decomposition [1], we define the following scale of norms:

$$|||v_h|||_s := \sqrt{(\mathcal{A}_h^s v_h, v_h)} = \left(\sum_i \lambda_i^s |c_i|^2 \right)^{1/2}. \tag{2.5}$$

Remark 2.1. 1. For $v_h \in V_h$, there is the norm equivalence

$$|||v_h|||_1 = \sqrt{(\mathcal{A}_h v_h, v_h)} = \sqrt{a(v_h, v_h)} = \|v_h\|_a \sim \|v_h\|_1, \tag{2.6}$$

and by definition, there holds

$$|||v_h|||_0 = \|v_h\|_0. \tag{2.7}$$

2. For $v_h \in V_h$ and $s \in \mathbb{R}$, the symmetry of the matrix leads to

$$|||\mathcal{A}_h^{s/2} v_h|||_0 = \sqrt{(\mathcal{A}_h^{s/2} v_h, \mathcal{A}_h^{s/2} v_h)} = \sqrt{(\mathcal{A}_h^s v_h, v_h)} = |||v_h|||_s. \tag{2.8}$$

3. For $r, t \in \mathbb{R}$, $s = (r+t)/2$ and $v_h, w_h \in V_h$ we have the *logarithmic convexity*

$$|(\mathcal{A}_h^s v_h, w_h)| \;=\; |(\mathcal{A}_h^{r/2} v_h, \mathcal{A}_h^{t/2} w_h)| \;\leqslant\; |||v_h|||_r \, |||w_h|||_t \tag{2.9}$$

4. The eigenvalues $\{\lambda\}$ of $\mathcal{A}_h$ and eigenfunctions $v_h \in V_h$ which are determined by the generalised eigenvalue problem

$$a(v_h, \varphi_h) = \lambda (v_h, \varphi_h) \quad \forall\, \varphi_h \in V_h$$

can be bounded as follows (see [1]): $c \leqslant \lambda(\mathcal{A}_h) \leqslant c' h^{-2}$.

Now, we repeat the classical multigrid proof for linear finite elements, starting with the *smoothing property.*

Lemma 2.1 (Smoothing property of damped Richardson iteration). *Let* $\lambda_{\max} = \lambda_{\max}(\mathcal{A}_h)$ *denote the maximal eigenvalue of the operator* $\mathcal{A}_h$ *and* $u_h \in V_h$ *the exact solution of problem* (2.3). *Furthermore, with arbitrary* $u_h^0 \in V_h$, *for* $i \in \mathbb{N}_0$ *we define the* damped Richardson smoother $\mathcal{S}_h : V_h \to V_h$ *with* $u_h^{i+1} := \mathcal{S}_h u_h^i$ *being the solution of*

$$(u_h^{i+1}, \varphi_h) = (u_h^i, \varphi_h) - \lambda_{\max}^{-1}((f, \varphi_h) - a(u_h^i, \varphi_h)) \quad \forall\, \varphi_h \in V_h.$$

Then, with $e_i = e_h^i := u_h^i - u_h$, *for arbitrary* $s, t \in \mathbb{R}$, $s \geqslant t$, $\mathcal{S}_h$ *satisfies the* smoothing property, *i.e., after n Richardson smoothing steps*

$$|||e_n|||_s \leqslant c n^{(t-s)/2} h^{(t-s)} |||e_0|||_t. \tag{2.10}$$

Proof. Following [1, 2], let $\{z_i\} \subset V_h$ denote an orthonormal basis of eigenvectors of V_h, $\{\lambda_i\}$ the corresponding eigenvalues and $e_0 = \sum_i c_i z_i$ the representation of e_0 w.r.t. this basis. By induction, $e_n = \sum_i (1 - \lambda_i/\lambda_{\max})^n c_i z_i$. Then, from $0 < \lambda_i/\lambda_{\max} \leqslant 1$, it follows:

$$|||e_n|||_s^2 \overset{(2.5)}{=} (\mathcal{A}_h^s e_n, e_n) = \sum_i \lambda_i^s [(1 - \lambda_i/\lambda_{\max})^n c_i]^2 =$$

$$\lambda_{\max}^{s-t} \sum_i (\lambda_i/\lambda_{\max})^{s-t}[(1-\lambda_i/\lambda_{\max})^{2n}]\lambda_i^t c_i^2 \leqslant \lambda_{\max}^{s-t} \max_{0\leqslant\xi\leqslant 1}[\xi^{s-t}(1-\xi)^{2n}] \sum_i \lambda_i^t c_i^2.$$

Now using $\lambda_{\max}^{s-t} \leqslant ch^{-2(s-t)}$, $\max_{0\leqslant\xi\leqslant1}[\xi^{s-t}(1-\xi)^{2n}] \leqslant 1/n^{s-t}$ and $\sum_i \lambda_i^t c_i^2 = |||e_0|||_t^2$, taking the square root, we obtain the desired result. □

As a special case, we set $s = 2$, $t = 0$, and $s = 3$, $t = -1$, respectively.

Corollary 2.1. *After $\mu \in \mathbb{N}$ smoothing steps, we obtain*

$$|||e_\mu|||_2 \leqslant c\mu^{-1}h^{-2}|||e_0|||_0, \tag{2.11}$$

$$|||e_\mu|||_3 \leqslant c\mu^{-2}h^{-4}|||e_0|||_{-1}. \tag{2.12}$$

Next, we need for the analysis of the coarse grid correction.

Definition 2.2 (Coarse grid operator). The *coarse grid operator* is defined as the mapping $\mathcal{P}_h^{2h} : V_h \to V_{2h}$ with

$$a(\mathcal{P}_h^{2h}v_h, v_{2h}) = a(v_h, \mathcal{I}_{2h}^h v_{2h}) \quad \forall\, v_h \in V_h,\ v_{2h} \in V_{2h}, \tag{2.13}$$

where $\mathcal{I}_{2h}^h : V_{2h} \to V_h$, $\mathcal{I}_{2h}^h v_{2h} = v_{2h} \in V_h\ \forall\, v_{2h} \in V_{2h}$ denotes the natural injection.

Remark 2.2. *The coarse grid correction in step* 3) *of the given two-grid algorithm is defined as $u_h^{\mu+1} = u_h^\mu + u_{2h}$ with $u_{2h} = -\mathcal{P}_h^{2h}(e_h^\mu)$. Subtracting u_h leads to*

$$e_h^{\mu+1} = e_h^\mu - \mathcal{P}_h^{2h}e_h^\mu \tag{2.14}$$

and with the definition of $\mathcal{I}_{2h}^h$, this leads to the following essential orthogonality relation*:*

$$a(e_h^{\mu+1}, \varphi_{2h}) = a(e_h^\mu - \mathcal{P}_h^{2h}e_h^\mu, \varphi_{2h}) = 0 \quad \forall\, \varphi_{2h} \in V_{2h}. \tag{2.15}$$

When using linear finite elements for the discretisation, we use estimate (2.11) to characterise the effect of the smoothing. Formula (2.11) states that as a result of the smoothing procedure with μ smoothing steps we gain a factor $1/\mu$, while "loosing" a factor of h^{-2}. This has to be compensated by an appropriate *approximation property*, which is based on a standard duality argument for linear finite elements. The proof then follows the following "roadmap" — a more detailed description of the different steps can be found, e.g., in [2]:

At first, one shows a *duality argument for linear finite elements*,

$$\|u - u_h\|_0 \leqslant ch\|u - u_h\|_1, \tag{2.16}$$

which is used in a special case to show the inequality

$$\|e_h^{\mu+1}\|_0 \leqslant ch\|e_h^{\mu+1}\|_1. \tag{2.17}$$

Applying the norm equivalences to this inequality and using orthogonality (2.15) and the logarithmic convexity (2.9) then leads to the *approximation property for linear finite elements*:

$$|||e_h^{\mu+1}|||_0 \leqslant ch^2|||e_h^\mu|||_2. \tag{2.18}$$

Finally, combining (2.11) and (2.18) for linear finite elements allows to obtain the well-known convergence result of the two-grid algorithm:

$$|||e_h^{\mu+1}|||_0 \leqslant ch^2|||e_h^\mu|||_2 \leqslant c\mu^{-1}|||e_0|||_0. \tag{2.19}$$

The multigrid proof for the W-cycle follows by a perturbation argument (see [1, 2, 4, 5]).

3. Approximation property for quadratic FEM

In this section, we investigate the case that quadratic finite elements are used. Before we start with a detailed investigation, we first discuss our objectives and give an outline of the arguments.

Starting from (2.12), we will estimate the $|||\cdot|||_3$-norm by the $|||\cdot|||_{-1}$-norm for the smoothing property. Using μ presmoothing steps, this will give us a factor $1/\mu^2$, while we will also "loose" a factor h^{-4}. This will be compensated by an appropriate approximation property as in the linear case. Therefore, the aim will be to generalise the approximation property (2.18) to the case of quadratic FEM, where $|||e_h^{\mu+1}|||_{-1} \leqslant ch^4|||e_h^{\mu}|||_3$. For this, we will need a kind of norm equivalence between the $|||\cdot|||_{-1}$ and the $\|\cdot\|_{-1}$-norm. Using $\|v_h\|_1 \leqslant c\,|||v_h|||_1$, we will be able to show one part of this equivalence

$$|||v_h|||_{-1} \leqslant c\,\|v_h\|_{-1} \quad \forall\, v_h \in V_h.$$

Then, using a duality argument for quadratic finite elements will give

$$|||e_h^{\mu+1}|||_{-1} \leqslant c\|e_h^{\mu+1}\|_{-1} \leqslant ch^2\|e_h^{\mu+1}\|_1 \leqslant ch^2|||e_h^{\mu+1}|||_1.$$

Combining the square of this term with $|||e_h^{\mu+1}|||_1^2 \leqslant |||e_h^{\mu+1}|||_{-1}|||e_h^{\mu}|||_3$ and cancelling equal terms on both sides will lead to the desired result, i.e. the approximation property

$$|||e_h^{\mu+1}|||_{-1} \leqslant ch^4|||e_h^{\mu}|||_3.$$

Having shown this small "roadmap" for the convergence proof in the case of quadratic finite elements, we give a detailed study, starting with the following Lemma.

Lemma 3.1.

$$|||e_h^{\mu+1}|||_1^2 \;\leqslant\; |||e_h^{\mu+1}|||_{-1}|||e_h^{\mu}|||_3 \tag{3.1}$$

Proof. Using the logarithmic convexity and the orthogonality of the Ritz-projection $a(e_h^{\mu+1}, \varphi_{2h}) = 0\;\forall\,\varphi_{2h} \in V_{2h}$ leads to:

$$|||e_h^{\mu+1}|||_1^2 = a(e_h^{\mu+1}, e_h^{\mu} - \mathcal{P}_h^{2h}e_h^{\mu}) = a(e_h^{\mu+1}, e_h^{\mu}) = (\mathcal{A}_h e_h^{\mu+1}, e_h^{\mu}) \overset{(2.9)}{\leqslant} |||e_h^{\mu+1}|||_{-1}|||e_h^{\mu}|||_3. \qquad \square$$

Next, we formulate the required inequality for the $|||\cdot|||_{-1}$-norm.

Proposition 3.1 (Norm estimate). *There is a constant $c > 0$ such that for all $v_h \in V_h$ the following inequality holds:*

$$|||v_h|||_{-1} \leqslant c\|v_h\|_{-1}. \tag{3.2}$$

Proof. The $\|\cdot\|_{-1}$ norm is defined by $\|v\|_{-1} = \sup\limits_{\varphi\in H^1(\Omega)} (v,\varphi)/\|\varphi\|_1$. Without loss of generality, we assume $v_h \neq 0$. Then, we can write:

$$|||v_h|||_{-1}^2 \overset{(2.5)}{=} (v_h, \mathcal{A}_h^{-1}v_h) = \frac{(v_h, \mathcal{A}_h^{-1}v_h)}{\|\mathcal{A}_h^{-1}v_h\|_1}\;\|\mathcal{A}_h^{-1}v_h\|_1 \overset{\mathcal{A}_h^{-1}v_h\in V_h}{\leqslant} \max_{w_h\in V_h}\frac{(v_h,w_h)}{\|w_h\|_1}\|\mathcal{A}_h^{-1}v_h\|_1 \overset{V_h\subset H^1(\Omega)}{\leqslant}$$

$$\sup_{w\in H^1(\Omega)}\frac{(v_h,w)}{\|w\|_1}\|\mathcal{A}_h^{-1}v_h\|_1 \overset{\text{def}}{=} \|v_h\|_{-1}\|\mathcal{A}_h^{-1}v_h\|_1 \overset{(2.6)}{\leqslant} c\|v_h\|_{-1}|||\mathcal{A}_h^{-1}v_h|||_1 \overset{(2.8)}{=} c\|v_h\|_{-1}|||v_h|||_{-1}.$$

$\square$

Then, we discuss the duality argument for quadratic finite elements.

Proposition 3.2 (Duality argument for quadratic finite elements). *Let $u_h \in V_h$ be the piecewise quadratic approximation to $u \in V$ in the sense of* (2.3)*. Then, the following estimate holds for a constant $c > 0$:*

$$\|u - u_h\|_{-1} \leqslant ch^2 \|u - u_h\|_1. \tag{3.3}$$

Proof. For arbitrary $g \in H^1(\Omega)$, let $z \in V$ be the solution of the auxiliary problem

$$a(z, \varphi) = (g, \varphi) \quad \forall\, \varphi \in V.$$

As the primal problem is assumed to be $H^3(\Omega)$-regular, this dual problem is also $H^3(\Omega)$-regular, so we have $z \in H^3(\Omega)$. Let $z_h \in V_h$ be the piecewise quadratic solution of the corresponding discrete problem

$$a(z_h, \varphi_h) = (g, \varphi_h) \quad \forall\, \varphi_h \in V_h.$$

By using $\varphi := u - u_h \in V$, $a(u - u_h, \varphi_h) = 0\ \forall\, \varphi_h \in V_h$ and the Bramble — Hilbert Lemma for quadratic finite elements, we derive

$$(g, u - u_h) = a(z - z_h, u - u_h) \leqslant ch^2 \|z\|_3 \|u - u_h\|_1 \leqslant ch^2 \|g\|_1 \|u - u_h\|_1$$

for all $g \in H^1(\Omega)$. Therefore, we obtain

$$\|u - u_h\|_{-1} = \sup_{g \in H^1(\Omega)} \frac{(u - u_h, g)}{\|g\|_1} \leqslant ch^2 \|u - u_h\|_1.$$

□

Corollary 3.1. *By formally setting $V := V_h$, $u := e_h^\mu \in V_h$, $u_h := \mathcal{P}_h^{2h} e_h^\mu \in V_{2h}$ we obtain with* (2.14)

$$\|e_h^{\mu+1}\|_{-1} \leqslant ch^2\ \|e_h^{\mu+1}\|_1. \tag{3.4}$$

After these preparations, we can formulate the desired approximation property.

Proposition 3.3 (Approximation property). *Assuming $H^3(\Omega)$-regularity for problem* (2.2) *and $\mu \in \mathbb{N}$ large enough, the error $e_h^{\mu+1}$ after one two-grid cycle behaves like*

$$|||e_h^{\mu+1}|||_{-1} \leqslant ch^4 |||e_h^\mu|||_3 \tag{3.5}$$

for a constant $c > 0$.

Proof. By the preceding lemmas and the duality argument for quadratic finite elements it follows

$$|||e_h^{\mu+1}|||_{-1} \overset{(3.2)}{\leqslant} c\|e_h^{\mu+1}\|_{-1} \overset{(3.3)}{\leqslant} ch^2 \|e_h^{\mu+1}\|_1 \overset{(2.6)}{\leqslant} ch^2 |||e_h^{\mu+1}|||_1$$

and therefore

$$|||e_h^{\mu+1}|||_{-1}^2 \leqslant ch^4 |||e_h^{\mu+1}|||_1^2 \overset{(3.1)}{\leqslant} ch^4 |||e_h^{\mu+1}|||_{-1} |||e_h^\mu|||_3.$$

Cancelling redundant terms gives the desired formula. □

Finally, we combine the previous results and achieve the proof for the optimal two-grid convergence for quadratic finite elements.

Theorem 3.1 (Two-grid convergence with quadratic finite elements). *Assuming $H^3(\Omega)$-regularity for problem* (2.2) *and $\mu \in \mathbb{N}$ large enough, the error $e_h^{\mu+1}$ after one two-grid cycle behaves like*

$$|||e_h^{\mu+1}|||_{-1} \leqslant c\mu^{-2}|||e_h^0|||_{-1} \tag{3.6}$$

for a constant $c > 0$.

Proof.

$$|||e_h^{\mu+1}|||_{-1} \overset{(3.5)}{\leqslant} ch^4|||e_h^{\mu}|||_3 \overset{(2.12)}{\leqslant} ch^4 c\mu^{-2}h^{-4}|||e_h^0|||_{-1}.$$

□

The convergence of the W-cycle multigrid algorithm with the same inequality for the error follows again by a perturbation argument similar to the linear case.

Remark 3.1. One essential detail in the above proof is the fact that for the grid transfer the natural injection $\mathcal{I}_{2h}^h v_{2h} = v_{2h} \in V_h \; \forall\, v_{2h} \in V_{2h}$ and its adjoint are used. The corresponding discrete transfer operators in a finite element implementation are the piecewise quadratic interpolation and restriction. Using different grid transfer operators like, e.g., the *piecewise linear interpolation on a once more refined grid* typically results in $\mathcal{I}_{2h}^h v_{2h} \neq v_{2h}$, thus destroying the essential Galerkin orthogonality (2.15). Therefore it cannot be guaranteed in general that equation (3.6) is valid in this case.

We finish this section with some final remarks about the case of higher order. For this purpose, we assume the problem (2.2) to be even $H^{s+1}(\Omega)$-regular with $s > 2$ and the discretisation to be done with piecewise polynomials of order s. Then, formulas (2.19) and (3.6) suggest the inequality $|||e_h^{\mu+1}|||_{1-s} \leqslant c\mu^{-s}|||e_h^0|||_{1-s}$ in this situation. To derive this, we take $t := s - 1$ and assume the inequality

$$\|v_h\|_t \leqslant c\,|||v_h|||_t \quad \forall\, v_h \in V_h \tag{3.7}$$

which holds for $t = 0, 1$ even in the stronger sense (2.6) and (2.7). The same technique as in Proposition 3.1 allows to derive the inequality

$$|||v_h|||_{-t} \leqslant c\,\|v_h\|_{-t} \quad \forall\, v_h \in V_h.$$

With the help of a similar duality argument, this leads to

$$|||e_h^{\mu+1}|||_{1-s}^2 \leqslant c\|e_h^{\mu+1}\|_{1-s}^2 \leqslant ch^{2s}\|e_h^{\mu+1}\|_1^2 \leqslant ch^{2s}|||e_h^{\mu+1}|||_1^2 \leqslant ch^{2s}|||e_h^{\mu+1}|||_{1-s}|||e_h^{\mu}|||_{1+s}$$

which gives the inequality $|||e_h^{\mu+1}|||_{1-s} \leqslant ch^{2s}|||e_h^{\mu}|||_{1+s}$ by cancelling redundant terms. Then, using the general formulation of the smoothing property (2.10) would indeed result in the above optimal two-grid convergence for finite element approximations of order s. However, whether or not equation (3.7) is valid is an open question at present.

4. Numerical examples

In this section we perform some numerical tests in order to illustrate the above results. For this purpose, we define the *smoothing efficiency index*.

Definition 4.1 (Smoothing efficiency index). For $m \in \mathbb{N}$ we denote by Φ_m the multigrid algorithm with m smoothing steps. Let $\rho(\Phi_i)$ denote the asymptotic convergence rate which is measured from the last three multigrid iterations in the solution process. After a sufficiently large number of iterations, this is approximately a constant number. Then, we define

$$G(i,j) := (\rho(\Phi_i)/\rho(\Phi_j))^{1/t}, \tag{4.1}$$

where $t := \log_2(j/i)$, $i, j \in \mathbb{N}$, $i < j$.

Thus, the smoothing efficiency index describes approximately the mean improvement of the convergence rate when doubling the number of smoothing steps. In particular, if $j = 2^k i$ for $k \in \mathbb{N}$, then $t = k$ gives the number of doublings. We expect $G(i,j) \approx 2$ for an approximation with linear finite elements and $G(i,j) \approx 4$ if quadratic finite elements are used. For the numerical experiments below, we use $k = 1, 2$.

Example 4.1. We perform the following experiment.

1. On the unit square $\Omega = [0,1]^2 \subset \mathbb{R}^2$ we solve numerically

$$a(u,v) = (\nabla u, \nabla v) = (f, v) \quad \forall\, v \in V.$$

With $u|_{\partial\Omega} := 0$, following [6, p. 203], this problem is $H^{4-\alpha}$-regular, $0 < \alpha < 1$. As a right hand side, we prescribe f corresponding to the analytical solution $u(x,y) := \sin(xy)\sin((1-x)(1-y))$.

2. We use the triangular elements P_1 and P_2. As a coarse grid, we use the unit square $[0,1]^2$ cut into two triangles. The algorithm used is a two-grid algorithm on level 6/7, i.e., we regularly refine the coarse grid 5 and 6 times, respectively.

3. The iteration is stopped if the relative residual (measured in the l_2-norm) drops below 10^{-20} or if a maximum of 30 two-grid iterations is reached.

4. As a smoothing operator, we use the discrete Richardson smoothing $x_{n+1} := x_n - \omega(b - Ax_n)$ with parameter $\omega = 0.02$. (The parameter was chosen on the one hand to fulfil $\omega < 1/\lambda_{\max}$ and on the other hand to show a clear asymptotic convergence behaviour.) Furthermore, we test with the (slightly more advanced) Jacobi smoother with smoothing parameter $\omega = 0.2$, which in fact gives a similar behaviour as the Richardson smoother. All tests are performed with pure postsmoothing and no presmoothing. (Note that the above theoretical analysis can similarly be formulated for postsmoothing instead of presmoothing.)

The first test in Table 4.1 shows the resulting convergence rates ρ and the smoothing efficiency indices $G(\cdot,\cdot)$ corresponding to the number of smoothing steps m. In this test, we used

Table 4.1. **Numerical test, P_1, linear interpolation (i.e., natural injection)**

m	Richardson(0.02)			Jacobi(0.2)		
	ρ	$G(m/2,m)$	$G(m/4,m)$	ρ	$G(m/2,m)$	$G(m/4,m)$
8	7.13E-01			4.25E-01		
12	6.05E-01			2.81E-01		
16	5.14E-01	1.39		2.23E-01	1.90	
24	3.71E-01	1.63		1.53E-01	1.84	
32	2.71E-01	1.89	1.62	1.16E-01	1.92	1.91
48	1.93E-01	1.92	1.77	7.91E-02	1.93	1.89
64	1.46E-01	1.86	1.88	5.99E-02	1.94	1.93
96	9.89E-02	1.95	1.94	4.04E-02	1.96	1.94
128	7.50E-02	1.94	1.90	2.94E-02	2.04	1.99
192	5.01E-02	1.98	1.96	2.01E-02	2.01	1.98
256	3.81E-02	1.97	1.96	1.48E-02	1.99	2.02

the linear finite element space P_1 for the discretisation. The results show that we obtain the expected smoothing efficiency index of 2 as predicted by the multigrid proof for linear finite elements.

The second test in Table 4.2 now shows the smoothing efficiency indices when using the quadratic element P_2. As expected from the theoretical investigations, the smoothing efficiency indices approach a factor of 4 with increasing number of smoothing steps. We emphasise that we use the *fully quadratic interpolation* (i.e., natural injection) as a grid transfer operator.

Table 4.2. **Numerical test, P_2, fully quadratic interpolation (i.e., natural injection)**

m	Richardson(0.02)			Jacobi(0.2)		
	ρ	$G(m/2,m)$	$G(m/4,m)$	ρ	$G(m/2,m)$	$G(m/4,m)$
8	6.44E-01			4.29E-01		
12	5.17E-01			3.27E-01		
16	4.15E-01	1.55		3.05E-01	1.41	
24	3.44E-01	1.50		2.01E-01	1.63	
32	3.03E-01	1.37	1.46	1.34E-01	2.28	1.79
48	2.02E-01	1.71	1.60	6.40E-02	3.14	2.26
64	1.37E-01	2.20	1.74	3.50E-02	3.83	2.95
96	6.70E-02	3.01	2.27	1.58E-02	4.06	3.57
128	3.68E-02	3.73	2.87	9.31E-03	3.76	3.79
192	1.60E-02	4.19	3.55	4.22E-03	3.74	3.89
256	9.09E-03	4.05	3.89	2.57E-03	3.62	3.69

The effect of a lower order interpolation on the convergence rates is analysed in Table 4.3 (here only with the Richardson smoother). We tested what happens when one or both of the grid transfer operators are changed from fully quadratic interpolation to *piecewise bilinear interpolation on a once more refined mesh*. It is clearly seen that the property of "doubling the number of smoothing steps quarters the convergence rate" is lost in this case! When both operators are replaced that way, a clear factor of 2 can be seen here. When only one of the operators is replaced, the smoothing efficiency indices reduce to factors of $2-3$. So, as a rule of thumb we can state what many practitioners have already observed in numerical simulations:

Using quadratic finite elements with piecewise linear grid transfer in multigrid, the convergence rates asymptotically behave like $\mathcal{O}(1/m)$, m denoting the number of smoothing steps, thus loosing the optimal convergence rates for quadratic finite elements.

Table 4.3. **Numerical test, P_2; left: linear prolongation/restriction; center: quadratic prolongation, linear restriction; right: linear prolongation, quadratic restriction, Richardson smoother**

Prl./Rst.:	lin./lin.		quad./lin.		lin./quad.	
m	ρ	$G(m/2,m)$	ρ	$G(m/2,m)$	ρ	$G(m/2,m)$
8	6.39E-01		6.43E-01		6.39E-01	
12	5.14E-01		5.17E-01		5.14E-01	
16	4.28E-01	1.49	4.15E-01	1.55	4.15E-01	1.54
24	3.70E-01	1.39	3.61E-01	1.43	3.68E-01	1.40
32	3.04E-01	1.41	3.01E-01	1.38	3.01E-01	1.38
48	2.16E-01	1.71	2.04E-01	1.77	2.03E-01	1.81
64	1.61E-01	1.89	1.43E-01	2.11	1.41E-01	2.13
96	1.06E-01	2.04	6.55E-02	3.12	7.95E-02	2.56
128	7.26E-02	2.22	4.43E-02	3.22	5.27E-02	2.68
192	4.66E-02	2.28	2.51E-02	2.61	3.22E-02	2.47
256	3.52E-02	2.06	1.83E-02	2.42	2.30E-02	2.30

The numerical tests in Table 4.5 and Table 4.6 are quite similar to the first two tests. Again, we use the unit square $[0,1]^2$ as a coarse mesh, but this time we discretise with the quadrilateral bilinear Q_1 and biquadratic Q_2 element. Although we have not given a theoretical analysis for quadrilateral elements, we expect a similar behaviour, and the computed convergence rates confirm that expectation. In using Q_2, the smoothing efficiency indices are generally better than 2 (using the Jacobi smoother shows a clear tendency to a factor of about 4), while in using Q_1 only a factor of 2 is reached.

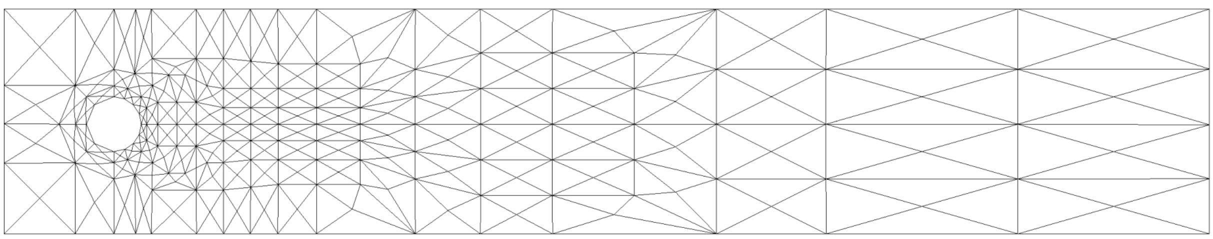
Table 4.4. **Numerical test, Q_1 with bilinear interpolation (natural injection)**

m	Richardson(0.02)		Jacobi(0.2)	
	ρ	$G(m/2,m)$	ρ	$G(m/2,m)$
8	7.14E-01		2.70E-01	
12	6.07E-01		1.41E-01	
16	5.16E-01	1.39	8.32E-02	3.24
24	3.72E-01	1.63	4.94E-02	2.85
32	2.68E-01	1.92	3.67E-02	2.27
48	1.40E-01	2.67	2.43E-02	2.03
64	8.17E-02	3.28	1.82E-02	2.02
96	4.73E-02	2.95	1.20E-02	2.03
128	3.47E-02	2.35	8.88E-03	2.05
192	2.29E-02	2.06	5.83E-03	2.05
256	1.71E-02	2.03	4.33E-03	2.05

Table 4.5. **Numerical test, Q_2 with biquadratic interpolation (natural injection)**

m	Richardson(0.02)		Jacobi(0.2)	
	ρ	$G(m/2,m)$	ρ	$G(m/2,m)$
8	6.64E-01		3.95E-01	
12	5.42E-01		2.48E-01	
16	4.42E-01	1.50	1.56E-01	2.53
24	2.94E-01	1.84	6.18E-02	4.02
32	1.95E-01	2.26	5.60E-02	2.79
48	8.70E-02	3.37	3.36E-02	1.84
64	4.32E-02	4.52	2.01E-02	2.78
96	3.34E-02	2.60	7.14E-03	4.71
128	2.23E-02	1.94	3.42E-03	5.89
192	1.02E-02	3.28	1.52E-03	4.69
256	5.75E-03	3.88	8.96E-04	3.82

We complete this section with a convergence rate test in a more complex situation. On a domain with two boundary components, we choose an unstructured triangular mesh as a coarse grid which is coming from a typical CFD simulation, namely "flow around a cylinder (in 2D)" (see [8]). The coarse grid can be seen in Figure below. For the discretisation, we use P_1 and P_2. The stopping criteria and boundary conditions are the same as above, while we use $f = 1$ as the right hand side of the Laplace equation. The two-grid algorithm takes place at level 4/5 for P_1 and at level 3/4 for P_2, so the number of unknowns (~ 67000) is the same. As a smoother, we choose the ILU(1) smoother here (i.e., the ILU-smoother with one level of fill-in according to [7]). The smoothing parameter was set to $\omega = 0.7$ and the Cuthill — McKee algorithm was used to sort the matrices/vectors.



CFD benchmark (coarse) grid for numerical tests with triangular elements

As can be seen from Table 4.6, the multigrid algorithm behaves similarly to the above model problems. The smoothing efficiency indices in the test with P_1 rapidly approach a factor of 2 while those in the test with P_2 approach a factor of 4 with convergence rates that are of similar quality or even better than the convergence rates in the case of P_1.

Table 4.6. **Numerical test, unstructured mesh, P_1 with linear interpolation (natural injection) and P_2 with quadratic interpolation (natural injection), ILU(1) smoother, $\omega = 0.7$**

	P_1		P_2	
m	ρ	$G(m/2,m)$	ρ	$G(m/2,m)$
2	3.37E-01		3.17E-01	
3	2.35E-01		2.09E-01	
4	1.77E-01	1.90	1.42E-01	2.23
6	1.16E-01	2.03	7.35E-02	2.85
8	8.53E-02	2.08	4.51E-02	3.15
12	5.30E-02	2.19	1.94E-02	3.78
16	3.49E-02	2.44	9.41E-03	4.79
24	2.01E-02	2.64	4.31E-03	4.50
32	1.45E-02	2.41	2.27E-03	4.14
48	9.15E-03	2.20	1.14E-03	3.78

5. Conclusions

We have shown that using higher order finite elements for the discretisation of a PDE is not only advantageous for the accuracy — which is already well-known — but also for the solution process of the discretised linear systems using standard (geometrical) multigrid algorithms. In fact, using quadratic finite elements, the convergence rates behave like $\mathcal{O}(1/m^2)$, for m being the number of smoothing steps, in contrast to the factor $\mathcal{O}(1/m)$ which is well-known for linear FEM. This property can be clearly seen in practical examples, but only if *"fully quadratic grid transfer"*, i.e., natural injection, is used; a lower order interpolation seems to destroy this property, which unfortunately happens in many existing codes.

The theoretical considerations indicate that this result might be extended to even higher order finite elements. If this is the case, geometrical multigrid solvers for the FEM discretisation of Poisson-like problems — with the corresponding grid transfer — could asymptotically lead to much faster convergence rates of order $\mathcal{O}(1/m^s)$, with s denoting the polynomial order.

References

1. R. E. Bank and T. F. Dupond, *An optimal order process for solving finite element equations*, Math. Comput., **36** (1981), No. 153, pp. 35–51.
2. D. Braess, *Finite Elemente*, 3rd edn., Springer, Berlin, 2003.
3. D. Braess and W. Hackbusch, *A new convergence proof for the multigrid method including the V–cycle*, SIAM Journal on Numerical Analysis, **20** (1983), No. 5, pp. 967–975.
4. S. C. Brenner, *Convergence of nonconforming V–cycle and F–cycle multigrid algorithms for second order elliptic boundary value problems*, Math. Comput., **73** (2003), No. 247, pp. 1041–1066, http://www.ams.org/mcom/2004-73-247/S0025-5718-03-01578-3/home.html.
5. W. Hackbusch, *Multi–Grid Methods and Applications*, Springer, Berlin, 1985.
6. W. Hackbusch, *Theorie und Numerik elliptischer Differentialgleichungen, 2. Auflage*, Oxford, Clarendon Press, 1996.
7. D. Hysom and A. Pothen, *Level–based Incomplete LU Factorization: Graph Model and Algorithms*, Preprint UCRL-JC-150789, U.S. Department of Energy, 2002, http://www.cs.odu.edu/~pothen/papers.html; to appear in: SIAM Journal On Matrix Analysis and Applications.
8. S. Turek, *Efficient Solvers for Incompressible Flow Problems: An Algorithmic and Computational Approach*, Springer, Berlin, 1999.
9. H. Yserentant, *Old and new convergence proofs for multigrid methods*, Acta Numerica, (1992), pp. 1–44.