A note on optimal multigrid convergence for higher-order FEM

Michael Köster* Stefan Turek†

August 30, 2005

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 numerical realisation of smoothing and intergrid transfer operators.

In this note, we discuss the numerical analysis of quadratic conforming finite elements in a multigrid solver. Numerical tests indicate that – if the 'correct' grid transfer is used – quadratic elements provide much better (asymptotic) convergence rates than linear finite element spaces: If m denotes the number of smoothing steps, the convergence rates behave asymptotically like $\mathcal{O}(\frac{1}{m^2})$ in contrast to $\mathcal{O}(\frac{1}{m})$ for linear FEM. The corresponding proof is explained in this note.

1 Introduction

Multigrid methods for solving linear systems that arise from FEM discretisations for PDEs are 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 ≥ 1), it is not

*Institute of Applied Mathematics, University of Dortmund, Vogelpothsweg 87, D-44227 Dortmund, Germany, michael.koester@mathematik.uni-dortmund.de

†Institute of Applied Mathematics, University of Dortmund, Vogelpothsweg 87, D-44227 Dortmund, Germany, stefan.turek@mathematik.uni-dortmund.de

yet fully understood if there are further quantitative effects for the convergence behaviour – at least there is no article known to the authors.

In this note, we concentrate on effects for 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 higher order finite elements, leading to our supposition that multigrid convergence rates might even 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, 8], 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 to point out 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:

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

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

b) $\mathcal{T} = \{\mathcal{T}_h\}$, 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, p. 58], i.e. there is $\kappa > 0$ so that every $T \in \mathcal{T}_h$ contains a ball with radius $\rho_h \geq \frac{h}{\kappa}$ for every h.

c) 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, p. 68]. In particular

- *$V_{2h} \subset V_h \subset V$; V_h has dimension $n = n_h$.*
- *Two elements $T_1 \neq T_2 \in \mathcal{T}_h$ intersect at most in common corners or edges.*

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

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

which is H_0^1-coercive and continuous, i.e. for all $v,w \in H_0^1(\Omega)$ there are $c,\alpha > 0$ such that

$$|a(v,w)| \leq c\|v\|_1 \cdot \|w\|_1, \qquad a(v,v) \geq \alpha\|v\|_1^2.$$

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

$$c\|v\|_1 \;\leq\; \|v\|_a \;\leq\; c'\|v\|_1 \quad (\textit{or shorter: } \|v\|_1 \sim \|v\|_a)$$

e) *The problem is to find a weak solution* $u \in V$ *of the boundary value problem*

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

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

$$a(u_h,\varphi_h) \;=\; (f_h,\varphi_h) \qquad \forall\, \varphi_h \in V_h. \tag{2.3}$$

f) *We assume the problem to be* $H^3(\Omega)$*-regular if* $f \in H^1(\Omega)$*.*

g) *The FEM approximation is done by quadratic finite elements.*

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:

2.2 Algorithm (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 behaviour of the algorithm:

- (μ,ν) = number of pre-/postsmoothing steps
- p =cycle ($p = 2$ describes the W-cycle)

Let $\mathcal{R}_h : V_h \to V_h$ denote a special mapping, called *smoothing operator*, and $h_{\max}$ the resolution of the coarsest grid in the family $\{V_h\}$.

function MultigridCycle(h, u_h^0, f_h**) :** u_h^*

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

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

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

b) *Presmoothing:* Compute: $u_h^i := \mathcal{R}_h u_h^{i-1} \qquad i = 1, ..., \mu$

c) *Coarse grid correction:* Set up the *coarse grid problem* to find $u_{2h} \in V_{2h}$,

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

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

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

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

$$u_{2h}^i := \textbf{MultigridCycle}(2h, u_{2h}^{i-1}, f_{2h}) \qquad i = 1, ..., p,$$

and correct the current approximative solution:

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

d) *Postsmoothing:* Compute: $u_h^{\mu+1+i} := \mathcal{R}_h u_h^{\mu+i} \qquad i = 1, ..., \nu$
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 to the case of a two-grid algorithm, as the multigrid proof follows by a perturbation argument (cf. [1, 2, 4, 5]). The above assumptions allow to formulate a couple of (algebraic) statements which all 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].

2.3 Definitions (Scale of norms)**.** *a) We define a linear continuous operator* $\mathcal{A}_h : V_h \to V_h$ *by*

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

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

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

2.4 Remarks. a) 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

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

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

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

c) For $r, t \in \mathbb{R}$, $s = \frac{r+t}{2}$ and $v_h, w_h \in V_h$ we have the *logarithmic convexity*:

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

d) The eigenvalues $\{\lambda\}$ of $\mathcal{A}_h$ and eigenfunctions $v_h \in V_h$ which are computed via the generalised eigenvalue problem

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

can be characterised following [1]:

$$c \;\leq\; \lambda(\mathcal{A}_h) \;\leq\; c' h^{-2}$$

2.5 Lemma (Smoothing property of damped Richardson iteration)**.** *Let $\lambda_{max} = \lambda_{max}(\mathcal{A}_h)$ denote the maximum 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 $\mathcal{R}_h : V_h \to V_h$ *with $u_h^{i+1} := \mathcal{R}_h u_h^i$ being the solution of*

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

Then, with $e_i = e_h^i := u_h^i - u_h$, for arbitrary $s, t \in \mathbb{R}$, $s \geq t$, this scheme fulfills the smoothing property, *i.e. after n Richardson smoothing steps:*

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

Proof: We follow [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 one shows that

$$e_n = \sum_i (1 - \lambda_i/\lambda_{\max})^n c_i z_i.$$

Then, from $0 < \lambda_i/\lambda_{\max} \leq 1$, it follows:

$$\begin{aligned} |||e_n|||_s^2 &= (\mathcal{A}_h^s e_n, e_n) \;=\; \sum_i \lambda_i^s [(1 - \lambda_i/\lambda_{\max})^n c_i]^2 \\ &= \lambda_{\max}^{s-t} \sum_i (\lambda_i/\lambda_{\max})^{s-t} [(1 - \lambda_i/\lambda_{\max})^{2n}]\; \lambda_i^t c_i^2 \\ &\leq \lambda_{\max}^{s-t} \max_{0 \leq \xi \leq 1} [\xi^{s-t}(1-\xi)^{2n}] \;\sum_i \lambda_i^t c_i^2 \end{aligned}$$

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

In particular, if we set $s = 2$, $t = 0$, resp., $s = 3$, $t = -1$, we obtain:

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

$$|||e_\mu|||_2 \;\leq\; \frac{c}{\mu} h^{-2} |||e_0|||_0 \tag{2.11}$$

$$|||e_\mu|||_3 \;\leq\; \frac{c}{\mu^2} h^{-4} |||e_0|||_{-1} \tag{2.12}$$

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

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

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

with $\mathcal{I}_{2h}^h : V_{2h} \to V_h$, $\mathcal{I}_{2h}^h v_{2h} = v_{2h} \in V_h\ \forall\, v_{2h} \in V_{2h}$ *being the natural inclusion.*

2.8 Remarks. a) The coarse grid correction in step d) of the given two-grid algorithm is defined by $u_h^{\mu+1} = u_h^\mu + u_{2h}$ with u_{2h} being the solution of

$$\begin{aligned} a(u_{2h}, \varphi_{2h}) &= (f_{2h}, \varphi_{2h}) = (f_h, \varphi_{2h}) - a(u_h^\mu, \varphi_{2h}) \\ &= a(u_h - u_h^\mu, \varphi_{2h}) = a(-e_h^\mu, \mathcal{I}_{2h}^h \varphi_{2h}) \qquad \forall\, \varphi_{2h} \in V_{2h}, \end{aligned}$$

thus $u_{2h} = \mathcal{P}_h^{2h}(-e_h^\mu)$.

b) Using the definition $u_h^{\mu+1} = u_h^\mu + u_{2h}$, subtracting u_h leads to $e_h^{\mu+1} = e_h^\mu - \mathcal{P}_h^{2h} e_h^\mu$, and with a) and the definition of $\mathcal{I}_{2h}^h$, the following *orthogonality relation* is essential:

$$a(e_h^{\mu+1}, \varphi_{2h}) = a(e_h^\mu - \mathcal{P}_h^{2h} e_h^\mu, \varphi_{2h}) = 0 \qquad \forall\, \varphi_{2h} \in V_{2h} \tag{2.14}$$

The standard W-cycle proof with linear finite elements estimates the $|||\cdot|||_2$-norm by the $|||\cdot|||_0$-norm. Using μ smoothing steps, this allows to gain a factor $\frac{1}{\mu}$, although this also "looses" a factor of h^{-2}:

$$|||e_\mu|||_2 \le \frac{c}{\mu} h^{-2} |||e_0|||_0$$

This is compensated by the following *approximation property.* To prove it, we first need a standard duality argument for linear finite elements:

2.9 Lemma (Duality argument for linear finite elements)**.** *When approximating a function* $u \in V$ *by the discrete finite element solution* $u_h \in V_h$ *being piecewise linear, the following estimate holds:*

$$\|u - u_h\|_0 \le ch \|u - u_h\|_1 \tag{2.15}$$

Proof: For $g \in L_2(\Omega)$, let $z \in V$ be the solution of the auxiliary problem

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

and $z_h \in V_h$ the approximating function of z in V_h. With $\varphi := u - u_h \in V$, $a(z, u - u_h) = (g, u - u_h)$ follows. Furthermore, $a(u - u_h, \varphi_h) = 0\ \forall \varphi_h \in V_h$ and Bramble-Hilbert Lemma for linear finite elements lead to

$$(u - u_h, g) = a(u - u_h, z - z_h) \le c\|u - u_h\|_1\ ch\|z\|_2 \le c\|u - u_h\|_1\ ch\|g\|_0$$

and therefore with $g := u - u_h$,

$$\|u - u_h\|_0^2 = (u - u_h, g) \le c\|u - u_h\|_1\ ch\|g\|_0 \le ch\|u - u_h\|_1 \|u - u_h\|_0.$$

□

2.10 Corollary (Duality argument für $e_h^{\mu+1}$)**.** *By formally setting* $V := V_h$*,* $u := e_h^n \in V_h$*,* $u_h := \mathcal{P}_h^{2h} e_h^n \in V_{2h}$ *we obtain:*

$$\|e_h^{\mu+1}\|_0 \le ch \; \|e_h^{\mu+1}\|_1 \tag{2.16}$$

2.11 Proposition (Approximation property for linear finite elements)**.**

$$|||e_h^{\mu+1}|||_0 \;\le\; ch^2 |||e_h^{\mu}|||_2 \tag{2.17}$$

Proof: From the norm equivalences and the duality argument (2.16) follows

$$|||e_h^{\mu+1}|||_0 \;\le\; ch \; |||e_h^{\mu+1}|||_1 . \tag{2.18}$$

Because of $\mathcal{P}_h^{2h} e_h^{\mu} \in V_{2h}$, the latter term can be estimated as

$$\begin{aligned} |||e_h^{\mu+1}|||_1^2 &\overset{\text{Def.}}{=} a(e_h^{\mu+1}, e_h^{\mu} - \mathcal{P}_h^{2h} e_h^{\mu}) \overset{(2.14)}{=} a(e_h^{\mu+1}, e_h^{\mu}) \\ &\overset{(2.4)}{=} (\mathcal{A}_h e_h^{\mu+1}, e_h^{\mu}) \overset{(2.9)}{\le} |||e_h^{\mu+1}|||_0 \; |||e_h^{\mu}|||_2 \\ &\overset{(2.18)}{\le} ch |||e_h^{\mu+1}|||_1 \; |||e_h^{\mu}|||_2 . \end{aligned}$$

Canceling out redundant terms and using (2.18) completes the proof. □

Now combining (2.11) and (2.17) for linear finite elements, we obtain the convergence of the two-grid algorithm:

$$|||e_h^{\mu+1}|||_0 \;\le\; ch^2 |||e_h^{\mu}|||_2 \;\le\; \frac{c}{\mu} |||e_0|||_0$$

The multigrid proof 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. We start from formula (2.12), i.e. we estimate the $|||\cdot|||_3$-norm by the $|||\cdot|||_{-1}$-norm for the smoothing property. Using μ presmoothing steps, this gives us a factor $\frac{1}{\mu^2}$, while we also "loose" a factor h^{-4}. This has to be compensated by an appropriate approximation property. Therefore, the aim is to generalise the approximation property (2.17) to the case of quadratic FEM. This will lead us to the relation

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

For this, we need a kind of norm equivalence between the $|||\cdot|||_{-1}$ and the $\|\cdot\|_{-1}$-norm. Using $\|v_h\|_1 \le c\,|||v_h|||_1$, we will be able to show one part of this equivalence:

$$|||v_h|||_{-1} \le c\,\|v_h\|_{-1} \quad \forall\; v_h \in V_h$$

Then, using a duality argument for quadratic finite elements gives

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

Combining the square of this term with

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

and canceling out equal terms on both sides will give the desired result, i.e. the approximation property

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

Now, we start our investigations with the following Lemma:

3.1 Lemma.

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

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:

$$\begin{aligned} |||e_h^{\mu+1}|||_1^2 &= a(e_h^{\mu+1}, e_h^{\mu} - \mathcal{P}_h^{2h} e_h^{\mu}) \;=\; a(e_h^{\mu+1}, e_h^{\mu}) \;=\; (\mathcal{A}_h e_h^{\mu+1}, e_h^{\mu}) \\ &\overset{(2.9)}{\leq} |||e_h^{\mu+1}|||_{-1}|||e_h^{\mu}|||_3 \end{aligned}$$

□

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

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

$$|||v_h|||_{-1} \;\leq\; c\,\|v_h\|_{-1} \tag{3.2}$$

Proof: By definition, there holds

$$\|v\|_{-1} = \sup_{\varphi \in V} \frac{(v,\varphi)}{\|\varphi\|_1}.$$

Without loss of generality we assume $v_h \neq 0$, and by using the symmetry of $\mathcal{A}_h$:

$$\begin{aligned} |||v_h|||_{-1}^2 &\overset{(2.8)}{=} |||\mathcal{A}_h^{-\frac{1}{2}} v_h|||_0^2 \overset{(2.7)}{=} \|\mathcal{A}_h^{-\frac{1}{2}} v_h\|_0^2 \\ &= (\mathcal{A}_h^{-\frac{1}{2}} v_h, \mathcal{A}_h^{-\frac{1}{2}} v_h) \;=\; (v_h, \mathcal{A}_h^{-1} v_h) \\ &= \frac{(v_h, \mathcal{A}_h^{-1} v_h)}{\|\mathcal{A}_h^{-1} v_h\|_1}\, \|\mathcal{A}_h^{-1} v_h\|_1 \\ &\overset{\mathcal{A}_h^{-1} v_h \in V_h}{\leq} \sup_{w_h \in V_h} \frac{(v_h, w_h)}{\|w_h\|_1}\, \|\mathcal{A}_h^{-1} v_h\|_1 \\ &\overset{V_h \subset V}{\leq} \sup_{w \in V} \frac{(v_h, w)}{\|w\|_1}\, \|\mathcal{A}_h^{-1} v_h\|_1 \overset{\text{Def.}}{=} \|v_h\|_{-1}\, \|\mathcal{A}_h^{-1} v_h\|_1 \\ &\overset{(2.6)}{\leq} c\,\|v_h\|_{-1}\, |||\mathcal{A}_h^{-1} v_h|||_1 \overset{(2.8)}{=} c\,\|v_h\|_{-1}\, |||v_h|||_{-1} \end{aligned}$$

□

Then, we still need a duality argument for quadratic finite elements:

3.3 Proposition (Duality argument for quadratic finite elements)**.** *When approximating a function $u \in V$ by the discrete finite element solution $u_h \in V_h$ being piecewise quadratic, the following estimate holds:*

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

***Proof*:** The proof is similar to the linear case. For $g \in H^1(\Omega)$ let $z \in V$ be the solution of the auxiliary problem

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

and $z_h \in V_h$ the approximating function of z 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

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

for all $g \in H^1(\Omega)$. Therefore, with $g := u - u_h$ we have

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

□

3.4 Corollary. *By 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:*

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

After these preparations we can formulate:

3.5 Proposition (Approximation property)**.** *Using the above notations, the following approximation property holds:*

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

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

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

and therefore

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

Canceling out redundant terms gives the desired formula. □

3.6 Theorem (Two-grid convergence with quadratic finite elements)**.** *With the above notations, after one two-grid cycle consisting of $\mu \in \mathbb{N}$ damped Richardson smoothing steps and one coarse grid correction, there holds*

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

Proof:

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

□

3.7 Remarks. a) Obtaining multigrid convergence can be derived from this two-grid theorem again with the help of a perturbation argument.
b) Up to now, it is not clear to us whether the norm estimate

$$|||v_h|||_t \leq c\,\|v_h\|_t$$

holds for arbitrary $v_h \in V_h$ and $t < 0, |t| \leq k+1$ (with k being the polynomial degree of the FEM space). If this holds for $t < -1$, then the proof can be formulated even more general. In this case, for $s > 2$, the approximation property could be reformulated as

$$|||e_h^{\mu+1}|||_{1-s} \leq ch^{2s}|||e_h^{\mu}|||_{1+s}$$

for a problem being $H^{s+1}(\Omega)$-regular and using finite elements of local degree s. Then, a similar duality argument would lead again to

$$\begin{aligned}|||e_h^{\mu+1}|||_{1-s}^2 &\leq c\|e_h^{\mu+1}\|_{1-s}^2 \leq ch^{2s}\|e_h^{\mu+1}\|_1^2 \leq ch^{2s}|||e_h^{\mu+1}|||_1^2\\ &\leq ch^{2s}|||e_h^{\mu+1}|||_{1-s}|||e_h^{\mu}|||_{1+s}\end{aligned}$$

so that the proposed formula follows again by canceling out redundant terms. Using the general formulation of the smoothing property (2.10) results in the two-grid convergence:

$$|||e_h^{\mu+1}|||_{1-s} \leq \frac{c}{\mu^s}|||e_h^0|||_{1-s}$$

c) If $\|v_h\|_t \leq c\,|||v_h|||_t$ holds for $t > 1$, then with similar techniques as before the required inequality $|||v_h|||_{-t} \leq c\,\|v_h\|_{-t}$ from b) can be proven. The ingredients are the same as in Proposition 3.2, there are only different powers of $\mathcal{A}_h$ involved. However, this relation between $\|\cdot\|_t$ and $|||\cdot|||_t$ for $t \geq 2$ is not yet clear.

4 Numerical examples

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

4.1 Definition (Smoothing efficiency index)**.** *For $m \in \mathbb{N}$ we denote by Φ_m the multigrid algorithm with m smoothing steps. Then, for $i, j \in \mathbb{N},\ i < j$, we set:*

$$t := t(i,j) := \log_2\left(\frac{j}{i}\right)$$

Let $\rho(\Phi_i)$ denote the asymptotic convergence rate[1] *of the algorithm using i smoothing steps. Then, we define:*

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

By definition, the smoothing efficiency index describes approximatively the mean improvement of the convergence rate when doubling the number of smoothing steps. In particular, if $j = 2^k i$ for $k \in \mathbb{N}$, $t = k$ 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 following numerical tests, we use $j = 2i, j = 4i$ and also $j = 8i$, which intuitively gives a more detailed picture about what happens when doubling the number of smoothing steps, as these settings build a mean value for two or three doublings.

We set up the following test configuration:

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

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

with $u_{|\partial\Omega} = 0$, $u(x,y) := \sin(xy)\ \sin((1-x)(1-y))$ and appropriate right hand side f. Following [6, p. 203], this problem is $H^{4-\alpha}$-regular, $0 < \alpha < 1$.

b) We use a two-grid algorithm on level 6/7, i.e. we regularly refine the unit square 5 and 6 times, respectively, by connecting opposite midpoints.

c) The iteration stops if the relative residuum (measured in the l_2-norm) drops below 10^{-28}, or if a maximum of 30 two-grid iterations is reached.

d) Instead of the Richardson smoother in weak formulation as used for the proof, we apply the damped Jacobi-iteration as smoother (with damping parameter $\omega = 0.7$).

Table 4.1 depicts the resulting number of iteration steps ITE, convergence rates ρ and smoothing efficiency indices $G(\cdot,\cdot)$, corresponding to the number of presmoothing steps m (no postsmoothing is performed). For this table, we used the (quadrilateral) bilinear space Q_1 for the discretisation. As one can see, we obtain the expected smoothing efficiency index of 2 as predicted by the multigrid proof for linear finite elements.

In the next test, we use the biquadratic space Q_2 for the discretisation. As visible in Table 4.2, we obtain the expected factor of 4 which we derived from our theoretical investigation. We clearly point out that we used the *fully biquadratic interpolation* as grid transfer operator.

Alternatively, we calculated the same problem with Q_2, but this time using only a *bilinear interpolation* as grid transfer operator. As can be seen in Table 4.3, the property that "doubling the number of smoothing steps quarters the

[1] This means the convergence rate measured in the last three steps before the algorithm terminates. Heuristically, this measures the asymptotic behaviour better than the standard definition of the convergence rate which is usually very small in the first couple of steps before it 'stabilises' towards the asymptotic rate.

m	**ITE**	ρ	**G(m/2,m)**	**G(m/4,m)**	**G(m/8,m)**
4	25	7,712E-02			
6	22	5,282E-02			
8	20	4,022E-02	1,917		
12	18	2,719E-02	1,943		
16	17	2,044E-02	1,968	1,942	
24	16	1,340E-02	2,029	1,985	
32	15	1,030E-02	1,985	1,976	1,957
48	14	6,536E-03	2,051	2,040	2,007
64	13	5,026E-03	2,049	2,017	2,000
96	12	2,757E-03	2,371	2,205	2,144
128	12	2,200E-03	2,284	2,163	2,102
192	11	1,463E-03	1,884	2,114	2,092

Table 4.1: Numerical test, Q_1, bilinear interpolation

convergence rate" is clearly lost in this case! Only a factor of 2 can be seen here. So, as a rule of thumb we can state what many practitioners have already observed in numerical simulations: *Using biquadratic finite elements with bilinear grid transfer in multigrid only behaves like bilinear finite elements concerning the convergence rate.*

m	**ITE**	ρ	**G(m/2,m)**	**G(m/4,m)**	**G(m/8,m)**
4	30	1,426E-01			
6	26	8,112E-02			
8	23	6,391E-02	2,232		
12	20	4,032E-02	2,012		
16	18	2,554E-02	2,502	2,363	
24	14	1,011E-02	3,988	2,833	
32	13	4,971E-03	5,138	3,586	3,062
48	11	1,990E-03	5,079	4,501	3,442
64	10	1,167E-03	4,259	4,678	3,797
96	9	5,515E-04	3,609	4,281	4,181
128	9	3,370E-04	3,463	3,841	4,232
192	8	1,576E-04	3,499	3,553	4,003

Table 4.2: Numerical test, Q_2, fully biquadratic interpolation

Next, we perform numerical tests for a more complex configuration which is coming from a typical CFD simulation, namely 'flow around a cylinder (in 2D)'. Figure 4.1 shows the typical coarse mesh which has been successively refined by connection opposite midpoints (see Table 4.4 for the geometrical details).

m	ITE	ρ	G(m/2,m)	G(m/4,m)	G(m/8,m)
4	30	2,931E-01			
6	30	2,258E-01			
8	30	1,808E-01	1,621		
12	30	1,287E-01	1,754		
16	28	1,032E-01	1,752	1,685	
24	25	7,368E-02	1,747	1,751	
32	23	5,702E-02	1,810	1,781	1,726
48	20	3,916E-02	1,881	1,813	1,793
64	19	2,983E-02	1,911	1,860	1,823
96	17	2,021E-02	1,938	1,910	1,854
128	16	1,503E-02	1,985	1,948	1,901
192	15	1,032E-02	1,958	1,948	1,926

Table 4.3: Numerical test, Q_2, bilinear interpolation

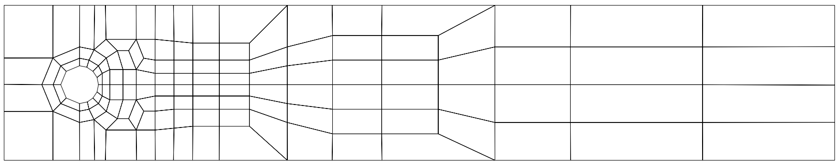

Figure 4.1: CFD benchmark grid for numerical tests.

Q_1			Q_2		
LEV	NEL	NEQ	LEV	NEL	NEQ
3	2080	2184	2	520	2184
4	8320	8528	3	2080	8528
5	33280	33696	4	8320	33696
6	133120	133952	5	33280	133952
7	532480	534144	6	133120	534144

Table 4.4: Statistical data for the 'CFD benchmark grid': Number of elements (NEL) and number of degrees of freedom (NEQ) on different mesh levels (LEV).

Dirichlet boundary conditions are prescribed on all boundary components, the outer rectangle as well as the inner circle. Then, solving Poisson problems with the bilinear Q_1 and biquadratic Q_2 (with fully biquadratic grid transfer) finite elements, respectively, leads to the multigrid convergence results in Table 4.5 in which case we measured the averaged rates for gaining 12 digits inside of the multigrid solver, starting from zero. The results show that multigrid for Q_2 behaves (at least) similar to fast multigrid solvers for bilinear finite elements, even on unstructured meshes which are prototypical for more realistic configurations (see also [7]).

	Q_1				Q_2			
LEV \ SMS	1	2	4	8	1	2	4	8
3	0.186	0.059	0.020	0.009	0.205	0.066	0.016	0.007
4	0.190	0.071	0.027	0.012	0.221	0.082	0.020	0.010
5	0.200	0.088	0.033	0.017	0.270	0.108	0.027	0.010
6	0.207	0.109	0.040	0.019	0.274	0.110	0.031	0.008
7	0.208	0.119	0.045	0.020	0.289	0.117	0.033	0.009

Table 4.5: Convergence rates on different mesh levels (LEV) with different numbers of smoothing steps (SMS).

5 Conclusions

We have shown that using higher order finite elements for the discretisation of PDEs 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, 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 *full biquadratic grid transfer* is used; a lower order interpolation destroys this property which unfortunately is happening in many codes.

The theoretical results indicate that this result might be extended to higher order finite elements. If this is the case, geometrical multigrid solvers for Poisson-like problems – with the corresponding grid transfer – could asymptotically lead to much faster convergence rates of order $\mathcal{O}(1/m^p)$, with p denoting the polynomial order, which might give a new viewpont on the efficiency of *hp*-FEM techniques in future.

References

[1] R. E. Bank and T. F. Dupond. An optimal order process for solving finite element equations. *Math. Comput.*, 36(153):35–51, 1981.

[2] D. Braess. *Finite Elemente*. Springer, Berlin, 3 edition, 2003. 3., korrigierte und ergänzte Auflage, ISBN 3-540-00122-0.

[3] D. Braess and W. Hackbusch. A new convergence proof for the multigrid method including the V–cycle. *SIAM Journal on Numerical Analysis*, 20(5):967–975, 1983.

[4] S. C. Brenner. Convergence of nonconforming V–cycle and F–cycle multigrid algorithms for second order elliptic boundary value problems. *Math. Comput.*, 73(247):1041–1066, August 2003. `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. Berlin-Heidelberg, ISBN 3540127615.

[6] W. Hackbusch. *Theorie und Numerik elliptischer Differentialgleichungen, 2. Auflage.* Oxford, Clarendon Press, 1996.

[7] M. Köster. Robuste Mehrgitter-Krylowraum-Techniken für FEM-Verfahren. Diplomarbeit, Universität Dortmund, 2004.

[8] H. Yserentant. Old and new convergence proofs for multigrid methods. *Acta Numerica*, pages 1–44, 1992.