Numerische Mathematik manuscript No.
(will be inserted by the editor)

On the H^1-stability of the L_2-projection onto finite element spaces

Randolph E. Bank · **Harry Yserentant**

May 28, 2013

Abstract We study the stability in the H^1-seminorm of the L_2-projection onto finite element spaces in the case of nonuniform but shape regular meshes in two and three dimensions and prove, in particular, stability for conforming triangular elements up to order twelve and conforming tetrahedral elements up to order seven.

Keywords L_2-projection · H^1-stability · Finite Elements

Mathematics Subject Classification (2000) 65N30, 65N50

1 Introduction

The L_2-projection onto a finite element space is a valuable tool in many areas of finite element analysis. It plays, for example, often an important role in the convergence analysis of multigrid methods. In [2], a simple proof of H^1-stability of the L_2-projection was given in the case of quasiuniform meshes with this application in mind. With the widespread use of adaptive and more general classes of nonuniform meshes, there is interest in generalizing this result to the nonuniform mesh case. At first glance, it seems that this should not be a difficult problem. The mass matrix, while no longer comparable to the identity matrix independent of the (now local)

Randolph E. Bank was supported by the Alexander von Humboldt Foundation through a Humboldt Research Award.

Harry Yserentant was supported by the DFG-Priority Program 1324 and the DFG-Research Center MATHEON.

Randolph E. Bank
Department of Mathematics, University of California San Diego, La Jolla, California 92093-0112
E-mail: rbank@ucsd.edu

Harry Yserentant
Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany
E-mail: yserentant@math.tu-berlin.de

mesh size, does remain comparable to its own diagonal. One expects that the exponential decay of matrix elements away from the diagonal in the inverse of the mass matrix, shown by Douglas, Dupont and Wahlbin in [14], should also remain valid even in the nonuniform mesh case. However, the central difficulty is that this exponential decay might potentially be offset by exponential growth due to grading of the finite element mesh. The works of Crouzeix and Thomée [10], Bramble, Pasciak and Steinbach [6], and Carstensen [7,8] all address this issue by imposing certain local or global growth constraints on the mesh. In this work, we must also impose some constraints, but ours are weaker, and allow the inclusion of high order elements and meshes generated by many commonly used adaptive meshing strategies.

Although our technique easily transfers to more general situations and can be applied to a large variety of different finite element spaces, we restrict ourselves in this article for the ease of presentation to the classical case of piecewise polynomial conforming elements. Starting point is a conforming triangulation $\mathscr{T}$ of a polygonal domain Ω in two or three space dimensions, built up from triangles in two space dimensions and tetrahedrons in the three-dimensional case. Associated with $\mathscr{T}$ is a conforming finite element space $\mathscr{S}$ of the usual kind, consisting of continuous, piecewise polynomial functions of at first fixed degree, determined by their nodal values. Our object of study is the L_2-orthogonal projection

$$Q : L_2(\Omega) \to \mathscr{S} \tag{1.1}$$

from $L_2(\Omega)$ onto the finite element space $\mathscr{S}$. We want to estimate the H^1-seminorm of the projection Qu, the L_2-norm of its first order derivatives, of a function u in the Sobolev space H^1 by the H^1-seminorm of u itself.

To each finite element $T \in \mathscr{T}$ we assign a nonnegative integer $k = k(T)$, the level of the element, such that $2^{-k(T)}$ is roughly proportional to the diameter $h(T)$ of T, in the sense that there are constants $\alpha > 0$ and $\beta > 0$ with

$$\alpha\, 2^{-k(T)} \le h(T) \le \beta\, 2^{-k(T)}. \tag{1.2}$$

The actual size of these constants is of no significance; only their ratio β/α will enter into our estimates. The triangulation $\mathscr{T}$ can be highly nonuniform and can contain finite elements from a very wide range of levels. We require, however, that the level of two neighboring elements differs at most by one. By neighbor, we refer to all elements that share a vertex with a given element.

Originally we had grids in mind that are generated by common adaptive meshing schemes based on bisection, for example the red-green refinement scheme in two [4] or three space dimensions [5]. In this case, the level of an element T counts the number of refinement steps that are needed to generate T from its ancestor in the initial triangulation. This is consistent with our current definition of level. The ratio of the constants α and β reflects in such cases the degree of non-uniformity of the initial triangulation but remains independent of the degree of refinement. Consider, for example, red-green refinement in two space dimensions. Red refinement of a triangle means subdivision into four congruent subtriangles of half the diameter. Since a green refinement into two subtriangles takes place only in the final step and is reversed in case of an additional refinement, the condition (1.2) holds with $\beta/\alpha = 2$

if one restricts the attention to a single triangle of the initial triangulation. In three space dimensions, the situation is basically the same. Red refinement of a tetrahedron means subdivision into eight tetrahedrons of equal volume that fall in at most three congruence classes, independent of the number of refinement steps, and green refinement takes again only place in the final step.

In red-green refinement, it is typical to impose a difference of at most one level between elements sharing a common edge in two or face in three space dimension. This is not sufficient to guarantee the same kind of constraint for the vertices, that is, that the levels of neighboring elements in the sense introduced above differs only by one. This condition can, however, be explicitly imposed on red-green meshes at little additional cost, even if such meshes often satisfy such a condition simply due to the refinement pattern generated by the local error indicators used in the refinement process. Experience shows that the local element sizes change even more moderately in case of mesh generators for nonuniform meshes that employ vertex-based mesh smoothing to locally optimize some measure of shape regularity [1], at least up to very few elements in small exceptional regions around strong singularities.

We will show that subject to the given conditions on the mesh the estimate

$$|Qu|_1 \leq c\,|u|_1 \tag{1.3}$$

for the H^1-seminorm of the projection Qu of a function u in H^1 holds for elements up to order twelve in two and up to order seven in three space dimensions. The constant c depends only on the order of the elements, on the ratio β/α of the constants in (1.2), and on the shape regularity of the triangles or tetrahedrons, but neither on the quasiuniformity of the mesh nor on any other such global property. For a sufficiently smooth grading of the mesh our argumentation works without any restriction to the polynomial order of the finite elements or any further condition to the triangulation up to shape regularity. It is even possible to extend our proof to hp-like adaptive finite element methods as long as the maximum order of the elements remains limited.

2 An iterative method

The main ingredient of our proof for the H^1-stability of the L_2-projection, as well as other proofs in the literature, are its strong localization properties, or roughly speaking, the fact that the inverse of the mass matrix decays very rapidly away from the diagonal. See Douglas, Dupont and Wahlbin [14] and the related work [11,12] of Demko and coauthors for band matrices. The results of this section will serve to quantify this effect in a fashion needed for our analysis.

Let $\mathscr{S}_0$ be the subspace of the finite element space $\mathscr{S}$ under consideration that consists of the functions in $\mathscr{S}$ that are L_2-orthogonal to the functions in $\mathscr{S}$ that vanish on the boundaries of the single elements. The functions in $\mathscr{S}_0$ are determined by their values at the nodes on the boundaries of the finite elements. For finite element spaces of polynomial order d or less, $d = 2$ or $d = 3$ the space dimension, no interior nodes, inside the elements, exist and $\mathscr{S}_0$ is the original space $\mathscr{S}$. In all other cases, we denote by $\mathscr{S}_1$ the space of the functions in $\mathscr{S}$ that vanish on the boundaries of the

finite elements. The L_2-orthogonal projection Q onto $\mathscr{S}$ splits then into the sum

$$Q = Q_0 + Q_1 \tag{2.1}$$

of the L_2-orthogonal projection Q_0 onto the subspace $\mathscr{S}_0$ and the L_2-orthogonal projection Q_1 onto $\mathscr{S}_1 = \mathscr{S}_0^{\perp}$. As the contributions from the single elements to Q_1 do not interact, the localization properties of the projection Q are completely determined by those of Q_0, the operator that will be studied in the rest of this section.

We label the vertices of the finite elements by the integers $i = 1, 2, \ldots n$. The vertex i is surrounded by the patch U_i, the union of the finite elements that share this vertex. Let $\mathscr{V}_i$ be the space that consists of the functions in the space $\mathscr{S}_0$ that vanish outside U_i. Let P_i be the L_2-orthogonal projection onto $\mathscr{V}_i$ and let

$$C = P_1 + P_2 + \ldots + P_n. \tag{2.2}$$

We will construct with help of this operator approximations of the projection $Q_0 u$ of a given square integrable function u onto $\mathscr{S}_0$. For that purpose, we first define finite element functions $u^{(\nu)} \in \mathscr{S}_0$ recursively by

$$u^{(\nu+1)} = u^{(\nu)} + C\,(u - u^{(\nu)}), \tag{2.3}$$

with $u^{(0)} \in \mathscr{S}_0$ an arbitrary starting value, and combine them to weighted averages

$$w^{(\ell)} = \sum_{\nu=0}^{\ell} \alpha_{\ell\nu} u^{(\nu)}, \quad \sum_{\nu=0}^{\ell} \alpha_{\ell\nu} = 1. \tag{2.4}$$

This procedure can be seen as a polynomially accelerated additive subspace correction method in the sense of [16]; see also the survey article [18].

The analysis of this iterative method starts from two abstract assumptions [16, 18] that still need to be verified. First, we assume that every finite element function

$$v = v_1 + v_2 + \ldots + v_n \tag{2.5}$$

in $\mathscr{S}_0$ can be decomposed into functions v_i in the subspaces $\mathscr{V}_i$ such that

$$\sum_i \|v_i\|_0^2 \;\le\; K_1 \|v\|_0^2. \tag{2.6}$$

Secondly, we assume that for every such linear combination (2.5) the estimate

$$\|v\|_0^2 \;\le\; K_2 \sum_i \|v_i\|_0^2 \tag{2.7}$$

holds. Our main tool is the following lemma that is an adaption of the general result for subspace decomposition methods [16, 18] to the present situation.

Lemma 2.1 *The restriction of the operator C to the subspace $\mathscr{S}_0$ is self-adjoint with respect to the L_2-inner product. Its eigenvalues λ range in the interval*

$$1/K_1 \;\le\; \lambda \;\le\; K_2. \tag{2.8}$$

Proof The self-adjointness of the operator C immediately results from its definition, that is, the self-adjointness of the projections P_i. Let $v = v_1 + v_2 + \ldots + v_n$ be a decomposition of the finite element function $v \in \mathscr{S}_0$ as in the assumption (2.6). Then it follows from the Cauchy-Schwarz inequality

$$(v,v) = \sum_i (v_i, v) = \sum_i (v_i, P_i v) \leq \Big(\sum_i \|v_i\|_0^2\Big)^{1/2} \Big(\sum_i \|P_i v\|_0^2\Big)^{1/2}$$

and therefore, by assumption (2.6),

$$(v,v) \leq K_1 \sum_i \|P_i v\|_0^2 = K_1 \sum_i (v, P_i v) = K_1 (v, Cv).$$

The operator C is therefore positive definite and its eigenvalues are bounded from below by the constant $1/K_1$. By assumption (2.7) conversely

$$\|Cv\|_0^2 = \|\sum_i P_i v\|_0^2 \leq K_2 \sum_i \|P_i v\|_0^2$$

and, using once more that P_i is an L_2-orthogonal projection,

$$\|Cv\|_0^2 \leq K_2 \sum_i (v, P_i v) = K_2 (v, Cv).$$

The constant K_2 is therefore an upper bound for the eigenvalues of C. □

The error between the projection $Q_0 u$ and its approximation (2.4) can be bounded in terms of the ratio $\lambda_n/\lambda_1 \leq K_1 K_2$ of the maximum and the minimum eigenvalue of the operator C restricted to the subspace $\mathscr{S}_0$. This results from:

Lemma 2.2 *Let $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ be the eigenvalues of the operator C restricted to the space $\mathscr{S}_0$ and introduce the polynomial*

$$P(\lambda) = \sum_{\nu=0}^{\ell} \alpha_{\ell\nu} (1-\lambda)^\nu \tag{2.9}$$

of degree ℓ. Then the error estimate

$$\|Q_0 u - w^{(\ell)}\|_0 \leq \max_{k=1,\ldots,n} |P(\lambda_k)| \, \|Q_0 u - w^{(0)}\|_0. \tag{2.10}$$

holds for the approximation (2.4) of the projection $Q_0 u$ of $u \in L_2$ onto $\mathscr{S}_0$.

Proof Since $\sum_{\nu=0}^{\ell} \alpha_{\ell\nu} = 1$, $w^{(0)} = u^{(0)}$, and $Cu = CQ_0 u$,

$$Q_0 u - w^{(\ell)} = \sum_{\nu=0}^{\ell} \alpha_{\ell\nu} \, (I - C)^\nu (Q_0 u - w^{(0)}).$$

Expanding the initial error into L_2-orthogonal eigenfunctions of the self-adjoint operator C from the finite element space into itself, the proposition follows. □

The polynomial P of degree ℓ that satisfies the normalization condition $P(0) = 1$ and attains the minimum maximal value on the interval $\lambda_1 \leq \lambda \leq \lambda_n$ is, up to a linear transformation of the variable and the multiplication by a constant factor, the Chebyshev polynomial of degree ℓ, a fact that is widely used in the analysis of semi-iterative and Krylov-space methods; see [13], for example. Choosing the averaging coefficients $\alpha_{\ell\nu}$ accordingly, one obtains from (2.10) and the estimate (2.8) for the eigenvalues of C the final estimate for the convergence rate of our semiiterative method:

Lemma 2.3 *If the coefficients $\alpha_{\ell\nu}$ are optimally chosen,*

$$\|Q_0 u - w^{(\ell)}\|_0 \leq \frac{2\,q^\ell}{1+q^{2\ell}}\,\|Q_0 u - w^{(0)}\|_0, \tag{2.11}$$

where the convergence rate

$$q = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \tag{2.12}$$

is determined by the condition number $\kappa = \lambda_n/\lambda_1 \leq K_1 K_2$ of the operator C.

It remains to bound the constants K_1 and K_2. The optimum decomposition (2.5) of a given finite element function in $\mathscr{S}_0$ can be found solving a global saddle point problem, but this observation is not truly helpful. Our estimate for K_1 is based on a local, element-wise construction and utilizes the nodal basis representation of the finite element functions. Let φ_i be the linear nodal basis function assigned to the vertex i. The φ_i form a partition of unity. We decompose the finite element functions $v \in \mathscr{S}_0$ therefore into the finite element functions $v_i \in \mathscr{V}_i$ given by their values

$$v_i(x_k) = \varphi_i(x_k) v(x_k) \tag{2.13}$$

at the boundary nodes x_k of the finite elements. The estimate (2.6) on the overall region then follows from the local estimates

$$\sum_i \|v_i\|_{0,T}^2 \leq K_1 \|v\|_{0,T}^2 \tag{2.14}$$

on the single elements T in the triangulation $\mathscr{T}$ by summation over all T. As the local mass matrices for the single elements differ only by a scalar area respectively volume factor and L_2-orthogonality is not affected by the size or shape of the elements, thus one can restrict oneself in the calculation of K_1 to the examination of a single reference element. This leads to a small eigenvalue problem

$$Bx = \zeta M x, \tag{2.15}$$

where M is the element mass matrix and B the sum of the three or four matrices that arise from the multiplication of the values of the finite element functions at the boundary nodes by the values of the corresponding functions φ_i. The maximum eigenvalue ζ represents then an upper bound for the best possible constant K_1.

Upper bounds for the constants K_2 can be derived in a similar way. As above, one can restrict oneself again to the consideration of a reference element T. If

$$\|v_1 + v_2 + \ldots + v_{d+1}\|_{0,T}^2 \leq K_2 \sum_{i=1}^{d+1} \|v_i\|_{0,T}^2 \tag{2.16}$$

p	K_1	K_2	q
1	2.000000000000000	2.000000000000000	0.333333333333333
2	1.632455532033676	2.720759220056126	0.356393958692601
3	1.393486807238790	2.644675210593510	0.315002511332227
4	1.295003216312185	2.636962512818568	0.297737526545759
5	1.222972165878670	2.594459484027661	0.280906146388308
6	1.302765305805047	2.593404439622517	0.295302231149967
7	1.299140221691548	2.565323271153285	0.292178820394059
8	1.512553465736873	2.567071358864272	0.326710513903202
9	1.576386303728820	2.547084406182064	0.334175300880362
10	2.028522383753149	2.549982029590451	0.389192884071447
11	2.269932525316121	2.535127973622464	0.411568856476138
12	3.300684993647074	2.538378961496179	0.486461159099767
13	4.089468949259074	2.526997823375298	0.525466578179801

Table 1 The constants K_1 and K_2 and the convergence rate q for triangular elements of order $p = 1$ to 13

holds for every set of finite element functions v_i that vanish on the edge respectively the face opposite to the vertex i and that are orthogonal to all polynomials of given order that vanish on the boundary of T, (2.7) holds with the same constant. Therefore

$$K_2 \leq d + 1 \tag{2.17}$$

independent of the order of the elements, as follows from the triangle and the Cauchy-Schwarz inequality. The calculation of of the minimum possible constant K_2 in (2.16) leads again to a matrix eigenvalue problem.

The corresponding eigenvalues have been calculated with help of the computer algebra program MAPLE, where exact arithmetic over the rational numbers has been used to set up the element matrices and to factor the resulting characteristic polynomials as far as possible. The zeroes of the single factors have then been calculated numerically with very high precision. All these zeroes turned out to be simple. To guarantee the accuracy of the zeroes up to the given number of digits, it has been shown afterward using again rational arithmetic that the factors change their sign in correspondingly small neighborhoods of the computed floating point numbers.

Table 1 shows the results of these computations, rounded to sixteen decimal digits, for triangular elements of order $p = 1$ to $p = 13$ together with the resulting upper bound for the convergence rate (2.12) of the polynomially accelerated iteration. For elements up to order twelve, the convergence rate remains below the bound

$$q < \frac{1}{2} \tag{2.18}$$

that will turn out to be critical and will be needed to counterbalance the mesh grading.

The situation for tetrahedral elements in three space dimensions is somewhat less favorable. The results of the corresponding computations are shown in Table 2. The convergence rate (2.12) remains below the critical bound (2.18) for elements of order one to seven. The condition (2.18) is violated for elements of order eight.

p	K_1	K_2	q
1	2.000000000000000	2.500000000000000	0.381966011250105
2	1.707106781186548	3.414213562373095	0.414213562373095
3	2.259093807508443	3.838087488839953	0.492976225239085
4	2.165373305836483	3.796924809950539	0.482851496718484
5	1.805566330903338	3.793859892383240	0.447095775974396
6	2.070521001184189	3.767800697161865	0.472724376154808
7	1.973130199031725	3.766914533817588	0.463271950206115
8	2.483457333030411	3.750763186028049	0.506419704289367

Table 2 The constants K_1 and K_2 and the convergence rate q for tetrahedral elements of order $p = 1$ to 8

3 A decay property of the $\mathbf{L_2}$-projection

The convergence rate q determines the localization properties of the L_2-projection. The faster the iteration converges, that is, the smaller q is, the faster local influences decay with distance from the source. The aim of this section is to quantify this effect. We group the finite elements by their levels and denote by Ω_i the union of the elements of level i. The next lemma describes how fast the projections of functions that vanish outside these sets decay on the exterior of these sets.

Lemma 3.1 *Let v_k be a square integrable function vanishing outside Ω_k. Then*

$$\|Qv_k\|_{0,\Omega_i} \leq \gamma(|i-k|)\|v_k\|_0, \tag{3.1}$$

where the prefactors decay like $q^{|i-k|}$ with the distance from i to k and are given by

$$\gamma(\ell) = \min(1, 2q^{\ell-1}). \tag{3.2}$$

Proof Let the $w^{(\ell)} \in \mathscr{S}_0$ be the iteratively generated approximations of $Q_0 v_k$ as in (2.4), with starting value $w^{(0)} = 0$. By (2.11) then

$$\|Q_0 v_k - w^{(\ell)}\|_0 \leq 2q^\ell \|Q_0 v_k\|_0 \leq 2q^\ell \|v_k\|_0.$$

As the operator (2.2) is composed of the local projections onto the spaces of functions in $\mathscr{S}_0$ that vanish outside the patches consisting of the elements surrounding a given vertex, information propagates in the transition from step ℓ to step $\ell+1$ from a given set of elements not further than to the union of the patches that cover at least one of these elements. As the level of neighboring elements, where neighbors were in our definition elements that share a vertex, differs by assumption at most by one, the approximation $w^{(\ell)}$ of $Q_0 v_k$ thus vanishes on the union Ω_i of the elements of level i as long as the number ℓ of iterations remains less than $|i-k|$. For these ℓ,

$$\|Q_0 v_k\|_{0,\Omega_i} = \|Q_0 v_k - w^{(\ell)}\|_{0,\Omega_i} \leq 2q^\ell \|v_k\|_0.$$

As $Qv_k = Q_0 v_k$ outside Ω_k, the proposition follows choosing $\ell = |i-k|-1$. □

4 The $\mathbf{H^1}$-stability

If the convergence rate q given by (2.12) is less than $1/2$, then the decay behavior expressed by (3.1) counterbalances the possible decrease of the element diameter from a given element to its neighbors. This is the key to the proof of the following theorem, that forms the basis of our proof for the H^1-stability of the L_2-projection.

Theorem 4.1 *If the convergence rate q given by (2.12) is less than $1/2$, there is a constant c such that for all square integrable functions v*

$$\|h^{-1}Qv\|_0 \le c\,\|h^{-1}v\|_0, \tag{4.1}$$

where the function h denotes the local meshwidth on the single elements. This constant exclusively depends on the convergence rate q and the ratio β/α of the constants in the assumption (1.2) that links the levels of the elements to their diameters.

Proof We first introduce the weighted L_2-norm given by the expression

$$|||v|||^2 = \sum_i 4^i\,\|v\|^2_{0,\Omega_i}, \tag{4.2}$$

where Ω_i is as above the union of the elements of level i. By assumption (1.2),

$$\alpha\,\|h^{-1}v\|_0 \le |||v||| \le \beta\,\|h^{-1}v\|_0$$

holds for all square integrable functions v. Thus it suffices to prove an estimate

$$|||Qv||| \le c_0\,|||v|||. \tag{4.3}$$

Let v_k attain the same values as v on Ω_k and the value 0 elsewhere. Then

$$|||Qv|||^2 = |||\sum_k Qv_k|||^2 = \sum_i 4^i \sum_{k,l} (Qv_k, Qv_\ell)_{\Omega_i},$$

where all sums extend over the finite set of nonempty levels i, k, and ℓ in the given triangulation. The estimate from Lemma 3.1 leads therefore to the estimate

$$|||Qv|||^2 \le \sum_{k,l}\sum_i 4^i\,\gamma(|i-k|)\,\gamma(|i-\ell|)\,\|v_k\|_0\|v_\ell\|_0$$

or, going back to the definition of v_k and v_ℓ and of the weighted L_2-norm (4.2),

$$|||Qv|||^2 \le \sum_{k,l} a_{k\ell}\,|||v_k|||\,|||v_\ell|||,$$

where the coefficients $a_{k\ell}$ are given by

$$a_{k\ell} = \sum_i 2^{i-k}\gamma(|i-k|)\,2^{i-\ell}\gamma(|i-\ell|).$$

If Λ denotes the maximum eigenvalue of the symmetric matrix with these entries,

$$|||Qv|||^2 \le \Lambda \sum_k |||v_k|||^2 = \Lambda \sum_k 4^k\,\|v\|^2_{0,\Omega_k} = \Lambda\,|||v|||^2$$

follows. The eigenvalue Λ is bounded by the maximum row sum

$$\sum_{\ell} |a_{k\ell}| = \sum_{i} 2^{i-k}\gamma(|i-k|) \sum_{\ell} 2^{i-\ell}\gamma(|i-\ell|)$$

of this matrix, that is, in terms of the maximum level j of the finite elements in the given triangulation, by the square of the expression

$$\sum_{\ell=0}^{j} 2^{-\ell}\gamma(\ell) + \sum_{\ell=1}^{j} 2^{\ell}\gamma(\ell).$$

If $q < 1/2$, one can let tend j to infinity and obtains the desired upper bound

$$\sum_{\ell=0}^{\infty} 2^{-\ell}\gamma(\ell) + \sum_{\ell=1}^{\infty} 2^{\ell}\gamma(\ell) = \frac{4}{1-2q} - \frac{1}{2} + \frac{q}{2-q}$$

for the constant c_0 in the estimate (4.3). In the limit case $q = 1/2$,

$$\sum_{\ell=0}^{j} 2^{-\ell}\gamma(\ell) + \sum_{\ell=1}^{j} 2^{\ell}\gamma(\ell) = 4j - \frac{1}{6} - \frac{4}{3}4^{-j}.$$

In this case, it is no longer possible to limit the weighted L_2-norm of the projection operator independent of j. It increases, however, at most like the logarithm of the quotient of the maximum and the minimum diameter of the finite elements. $\square$

The stability (4.1) of Q with respect to the given weighted L_2-norm implies the H^1-stability of Q. In fact, to finish our proof we need only a quasi-interpolation operator $\Pi : L_2 \to \mathscr{S}$ with the following properties. First, we assume that the estimate

$$\|h^{-1}(u - \Pi u)\|_0 \lesssim |u|_1 \tag{4.4}$$

holds for all functions $u \in H^1$, which is a kind of error estimate, and secondly, that

$$|\Pi u|_1 \lesssim |u|_1 \tag{4.5}$$

for these u. Since its introduction by Clément [9], the use of such quasi-interpolation operators is standard in finite element theory. One possible example is the operator

$$\Pi u = \sum_{i} \frac{(u, \varphi_i)}{(1, \varphi_i)} \varphi_i \tag{4.6}$$

that reproduces locally constant functions; the φ_i are the piecewise linear hat functions assigned to the vertices of the finite elements. It is analyzed in more detail in [17]. The proof of (4.4) and (4.5) is based on a local version of the Poincaré inequality and the local quasiuniformity of the triangulation.

Theorem 4.2 *If the convergence rate q given by (2.12) is less than $1/2$, and the L_2-orthogonal projection Q onto the finite element space $\mathscr{S}$ is thus stable with respect to the weighted L_2-norm from (4.1), there is a constant c such that for all $u \in H^1$*

$$|Qu|_1 \le c\,|u|_1. \tag{4.7}$$

Proof Since Πu is reproduced by the projection Q,

$$|Qu|_1 \leq |Q(u-\Pi u)|_1 + |\Pi u|_1.$$

The second term on the right hand side can, by means of (4.5), be estimated as desired. To estimate the first term, we first apply the inverse inequality

$$|Q(u-\Pi u)|_1 \lesssim \|h^{-1}Q(u-\Pi u)\|_0$$

and utilize Theorem 4.1, the stability of Q in the weighted L_2-norm:

$$\|h^{-1}Q(u-\Pi u)\|_0 \lesssim \|h^{-1}(u-\Pi u)\|_0.$$

Applying the error estimate (4.4), one obtains the estimate

$$|Q(u-\Pi u)|_1 \lesssim |u|_1$$

for the first term, which proves the theorem. □

The assumptions of Theorem 4.2 are in particular fulfilled for conforming triangular elements up to order twelve in two space dimensions and conforming tetrahedral elements up to order seven in three space dimensions. We conclude:

Theorem 4.3 *The L_2-projection onto the given finite element spaces is H^1-stable in the case of elements up to order twelve in two and up to order seven in three space dimensions, regardless the quasiuniformity of the triangulation.*

5 An analysis of a simple hp-method

Finite elements of a fixed order greater than three or four are rarely used in practice, in contrast to hp-methods that became meanwhile very popular. The recent version [1] of the finite element code PLTMG is based on such an approach. More details on such kinds of finite element methods can be found in [3]. Our theory applies to the hp-case, at least if the maximum order of the elements is limited.

We restrict ourselves here to the two-dimensional case and start from the same type of conforming triangulations $\mathscr{T}$ as described in the introduction, to which a conforming finite element space $\mathscr{S}$ is associated. The functions in $\mathscr{S}$ are globally continuous but their polynomial degree can now differ from triangle to triangle. The triangulations are composed of regular elements and a special kind of transition elements. Regular elements are elements as considered before. The restrictions of the finite element functions to regular elements of order p are the polynomials of degree p. These are determined by their values at the nodes of order p in the given triangle. The transition elements serve to connect elements of a given order p to a single element of order $p+1$. The finite element functions are determined on such triangles by their values at the nodes of order p, except for those on the transition edge that are there replaced by the nodes of order $p+1$. The restriction of the finite element space to such a triangle is spanned by the polynomials of degree p and a further polynomial of degree $p+1$ that enables the continuous transition across the transition edge and

p	K_1	K_2
1	2.973780511330725	2.264830035664230
2	1.857129338976809	2.723363613342249
3	1.615801659121557	2.661483393318108
4	1.321501990999704	2.649121326689785
5	1.290337433437715	2.619472429444019
6	1.303493086453341	2.608376639571315
7	1.499020700986966	2.597820625553555
8	1.690529874327097	2.584547510600745
9	2.088008464403344	2.583492578167884
10	2.604264691222322	2.574531675438955
11	3.530390137827354	2.572006470993573

Table 3 The constants K_1 and K_2 for the 2d-transition elements connecting triangles of order p and $p+1$

vanishes on the other two edges. The transition elements play in the given context the same role as the "green" triangles in red-green h-refinement [4].

The analysis of the preceding sections completely covers this construction. The corresponding constants K_1 and K_2 from (2.14) and (2.16) for the transition elements, calculated and checked in the same way as in Sect. 2, can be found in Table 3. As there, the interior nodes of the transition elements have been eliminated to speed up the convergence of the iteration underlying our proof. Since in the hp-case elements of different kind are combined, the global constants

$$K_1 = \max_{T\in\mathscr{T}} K_1(T), \quad K_2 = \max_{T\in\mathscr{T}} K_2(T) \tag{5.1}$$

are the maxima of the corresponding constants for the single elements respectively reference elements. Remarkably, the global constants K_1 and K_2 are dominated by those of low-order transition elements, K_1 by that of the transition elements from order 1 to 2 and K_2 by that of the transition elements from order 2 to 3. Nevertheless we can conclude that the convergence rate of the iteration from Sect. 2 is at least

$$q = 0.4799547\ldots \tag{5.2}$$

if we limit the maximum order of the elements to eleven and remains therefore below the critical bound $q = 1/2$. The L_2-projection thus remains also for this case H^1-stable without further conditions to the mesh or any other assumption.

6 The influence of the mesh grading

The crucial estimate that links the structure of the mesh and the type of the finite element functions is (3.1). It represents a worst case estimate that allows the size of the elements to change rather quickly. If one assumes a smoother grading of the finite element mesh, one can relax our conditions considerably. To quantify this effect, we

attribute a graph to the triangulation. The single elements are the nodes of this graph. Two nodes are connected by an edge if the corresponding two elements are neighbors. The distance of two elements is then the distance of the corresponding nodes in the graph. The proof of (3.1) uses that the distance of two elements of levels i and k is at least $|i-k|$. If one assumes that the distance of two such elements is at least $m|i-k|$ with m a real parameter greater than one for levels that differ sufficiently much, that is, that the meshsize changes at least in some average less abruptly, one obtains instead of (3.1) a potentially much stronger estimate

$$\|Qv_k\|_{0,\Omega_i} \leq \widetilde{\gamma}(|i-k|)\|v_k\|_0, \tag{6.1}$$

where the prefactor is, for ℓ greater than a fixed ℓ_0, now given by

$$\widetilde{\gamma}(\ell) = 2\,q^{m\ell-1}. \tag{6.2}$$

It depends now on m and decays like $q^{m|i-k|}$ with the distance from i to k. The proof of Theorem 4.1 works then as before, with the only difference that the condition that the convergence rate q is less than $1/2$ can be replaced by the less restrictive condition

$$q^m < \frac{1}{2}. \tag{6.3}$$

One can even go a step further modifying the concept of level. Assume that the diameter $h(T)$ of two elements in $\mathscr{T}$ that have at least one point in common differs at most by the factor $\mu \geq 1$; the existence of such a factor follows from the shape regularity of the elements. Then to each finite element T in our triangulation $\mathscr{T}$ we assign a nonnegative integer $k = k(T)$, the new kind of level of the element, such that

$$h_0\,\mu^{-k(T)} \leq h(T) < h_0\,\mu^{-k(T)+1}, \tag{6.4}$$

where h_0 is the maximum element diameter. The level of two neighboring elements differs then by construction at most by one. One can therefore show along the given lines that the L_2-projection remains H^1-stable as long as the condition

$$q^m < \frac{1}{\mu} \tag{6.5}$$

is satisfied. The condition on the convergence rate q of the iterative method from Sect. 2 thus becomes less and less restrictive the more μ approaches the value one or the larger constant m introduced as above is, that is, the smoother the grading of the mesh becomes. Shape regularity of the elements and a sufficiently smooth grading of the mesh provided, our argumentation thus works without any restriction to the type or order of the finite elements. There is therefore less doubt that the L_2-projection remains H^1-stable in almost all cases of practical interest.

7 A counterexample

The following simple example strongly suggests that the L_2-projection onto finite element spaces cannot be H^1-stable without any condition to the grid and that assumptions on the grading of the mesh cannot be avoided. We start from the sequence

$$x_0 = 0, \quad x_{k+1} = x_k + h_k, \quad k = 1,2,3,\dots, \tag{7.1}$$

of gridpoints in the interval $[0,1)$, where the local gridwidths

$$h_k = (1-\rho)\,\rho^k, \quad 0 < \rho < 1, \tag{7.2}$$

tend exponentially to zero and are chosen such that the x_k tend to 1 as k goes to infinity. We consider the infinite-dimensional space $\mathscr{S}$ that consists of the on $[0,1)$ continuous and square integrable functions u whose restrictions to the intervals $[x_k, x_{k+1}]$ are linear. As the spaces of the restrictions of the functions in $\mathscr{S}$ to all closed subintervals of the half-open interval $[0,1)$ are finite-dimensional, $\mathscr{S}$ is a closed subspace of $L_2(0,1)$ and the L_2-orthogonal projection Q from $L_2(0,1)$ to $\mathscr{S}$ exists. The nodal values u_k of the projection Qu of a functions $u \in L_2$ to $\mathscr{S}$ satisfy the equations

$$\frac{h_{k-1}}{6}\,u_{k-1} + \frac{h_{k-1}+h_k}{3}\,u_k + \frac{h_k}{6}\,u_{k+1} = (u,\varphi_k) \tag{7.3}$$

assigned to the gridpoints x_k for $k = 1,2,\dots$ and the additional equation

$$\frac{h_0}{3}\,u_0 + \frac{h_0}{6}\,u_1 = (u,\varphi_0), \tag{7.4}$$

assigned to x_0; the φ_k are the hat functions assigned to the x_k, the piecewise linear functions that take the value 1 at x_k and the value 0 at all other gridpoints.

In the following, we are only interested in the projections Qu of functions u that vanish for $x \geq x_1$. To represent these in closed form, we introduce the constants

$$\lambda_1 = -\,\frac{1+\rho-\sqrt{1+\rho+\rho^2}}{\rho}, \quad \lambda_2 = -\,\frac{1+\rho+\sqrt{1+\rho+\rho^2}}{\rho}. \tag{7.5}$$

Lemma 7.1 *The nodal values u_k of the projection Qu of a square integrable function u that takes the value zero for $x \geq x_1$ are, for $k \geq 1$, given by*

$$u_k = \frac{6\,(u,\varphi_0) - 12\,(u,\varphi_1)}{(1-\rho)(2+\lambda_1)}\,\lambda_1^k. \tag{7.6}$$

Proof The right hand sides of the difference equation (7.3) vanish in this case for all indices $k \geq 2$. The u_k satisfy therefore, for $k \geq 2$, the difference equation

$$u_{k-1} + 2\,(1+\rho)\,u_k + \rho\,u_{k+1} = 0.$$

Since λ_1 and λ_2 are the roots λ of the characteristic polynomial

$$1 + 2\,(1+\rho)\,\lambda + \rho\,\lambda^2$$

of this equation, the general solution of this equation is

$$u_k \;=\; a\,\lambda_1^{k-1} + b\,\lambda_2^{k-1}, \quad k = 1, 2, \ldots .$$

Since $\rho\lambda_1^2 < 1$ and $\rho\lambda_2^2 > 1$ for all ρ in the interval $(0,1]$ and because of

$$\int_{x_k}^{x_{k+1}} (\lambda^k \varphi_k + \lambda^{k+1}\varphi_{k+1})^2 \,\mathrm{d}x \;=\; \frac{(1-\rho)(1+\lambda+\lambda^2)}{3}\,(\rho\lambda^2)^k,$$

the u_k are nodal values of a square integrable function in $\mathscr{S}$ if and only if $b = 0$. This observation fixes the coefficient b. To determine a, we use that

$$(3h_0 + 4h_1)u_1 + 2h_1 u_2 \;=\; 12\,(u,\varphi_1) - 6\,(u,\varphi_0),$$

as follows from (7.3) and (7.4). Since λ_1 is a root of the characteristic polynomial,

$$(3h_0 + 4h_1 + 2h_1\lambda_1)\,\lambda_1 \;=\; -\,(1-\rho)(2+\lambda_1).$$

Inserting the representation $u_1 = a$, $u_2 = a\,\lambda_1$, the proposition follows. The value

$$u_0 \;=\; \frac{6\,(u,\varphi_0) + 6\,\lambda_1(u,\varphi_1)}{(1-\rho)(2+\lambda_1)}$$

of u_0 can then be derived from (7.4). $\square$

Not much surprisingly, $u_k = 0$ for $k \geq 1$ if $u = \varphi_0$. As

$$\int_{x_k}^{x_{k+1}} (\lambda^k \varphi_k' + \lambda^{k+1}\varphi_{k+1}')^2 \,\mathrm{d}x \;=\; \frac{(1-\lambda)^2}{1-\rho}\,\left(\frac{\lambda^2}{\rho}\right)^k, \tag{7.7}$$

the H^1-seminorms over the subintervals of a function represented by the values (7.6) sum up, for $(u,\varphi_0) \neq 2\,(u,\varphi_1)$, to a finite value if and only if

$$\frac{\lambda_1^2}{\rho} \;<\; 1, \tag{7.8}$$

regardless the smoothness of u. Since λ_1 behaves asymptotically like

$$\lambda_1 \;=\; -\frac{1}{2} + \frac{3}{8}\rho + \mathscr{O}(\rho^2), \quad \rho \to 0+, \tag{7.9}$$

this condition is violated if ρ approaches zero. That means, the L_2-projection can only be H^1-stable if the length of the subintervals does not decrease too fast in comparison to the speed of decay of the projection from gridpoint to gridpoint, the property on which the proof of Theorem 4.1 and with that our entire theory is based.

Appendix. The one-dimensional case

The one-dimensional case is particularly simple. The subspace $\mathscr{S}_0$ can in this case be built up from a single type of basis functions that are assigned to the boundary points of the subintervals. Let $[-1,1]$ be the reference interval and let $\varphi : [-1,1] \to \mathbb{R}$ be the polynomial of given degree n that takes the values $\varphi(-1) = 1$ and $\varphi(1) = 0$ and is L_2-orthogonal to all polynomials of degree n that vanish at -1 and 1. The basis functions on the reference interval are then the functions

$$\varphi_1(x) = \varphi(x), \quad \varphi_2(x) = \varphi(-x).$$

The reduced mass matrix of the overall problem becomes in this case a tridiagonal matrix. The condition number of its diagonally scaled counterpart can be estimated by the condition number κ of the (2×2)-element matrix

$$\begin{pmatrix} (\varphi_1,\varphi_1) & (\varphi_1,\varphi_2) \\ (\varphi_2,\varphi_1) & (\varphi_2,\varphi_2) \end{pmatrix}.$$

As the application of the operator C from Sect. 2 corresponds here to diagonal preconditioning of the reduced mass matrix, it determines the convergence rate

$$q \;=\; \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$$

of the iterative method studied in Sect. 2 and with that the upper bound for the factor μ from equation (6.5) by which the length of neighbored subintervals can differ.

To calculate the above element matrix, we use the Jacobi polynomials $P_k = P_k^{(\alpha,\beta)}$ of degree k for the indices $\alpha,\beta = 2$. They are given by the Rodrigues formula

$$P_k(x) \;=\; \frac{(-1)^k}{2^k\,k!\,(1-x^2)^2}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k\Big\{(1-x^2)^{k+2}\Big\}$$

and satisfy the orthogonality relations

$$\int_{-1}^{1}(1-x)^2(1+x)^2\,P_k(x)P_\ell(x)\,\mathrm{d}x \;=\; \frac{32}{2k+5}\,\frac{(k+1)(k+2)}{(k+3)(k+4)}\,\delta_{k\ell}.$$

For even k, they are symmetric and for odd k antisymmetric:

$$P_k(-x) \;=\; (-1)^k\,P_k(x).$$

More comprehensive information can be found in [15].

Lemma *The above element matrix is a scalar multiple of the matrix*

$$\begin{pmatrix} n+1 & (-1)^{n+1} \\ (-1)^{n+1} & n+1 \end{pmatrix}.$$

Proof We first expand the projection of the linear function that takes the value 1 at the left and the value 0 at the right endpoint onto the space of polynomials of degree at most n that vanish at both endpoints in terms of the L_2-orthonormal polynomials

$$\psi_k(x) \;=\; \left(\frac{2k+5}{32}\,\frac{(k+3)(k+4)}{(k+1)(k+2)}\right)^{1/2}(1-x^2)\,P_k(x).$$

Since the boundary terms vanish, multiple integration by parts yields

$$\int_{-1}^{1}\frac{1-x}{2}\,(1-x^2)\,P_k(x)\,\mathrm{d}x \;=\; \frac{1}{2^{k+1}k!}\int_{-1}^{1}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k\left\{\frac{1}{1+x}\right\}(1-x^2)^{k+2}\,\mathrm{d}x$$

and thus after some rather obvious intermediate steps the explicit representation

$$\int_{-1}^{1}\frac{1-x}{2}\,(1-x^2)\,P_k(x)\,\mathrm{d}x \;=\; (-1)^k\,\frac{8}{(k+3)(k+4)}$$

of the integral. The expansion coefficients are therefore

$$c_k \;=\; (-1)^k\left(\frac{2k+5}{32}\,\frac{(k+3)(k+4)}{(k+1)(k+2)}\right)^{1/2}\frac{8}{(k+3)(k+4)}$$

and the initially introduced shape function φ of polynomial order $n \geq 2$ is given by

$$\varphi(x) \;=\; \frac{1-x}{2} \;-\; \sum_{k=0}^{n-2} c_k\,\psi_k(x)$$

or directly in terms of the Jacobi polynomials by

$$\varphi(x) \;=\; \frac{1-x}{2} \;-\; \sum_{k=0}^{n-2}\frac{(-1)^k}{4}\,\frac{2k+5}{(k+1)(k+2)}\,(1-x^2)\,P_k(x).$$

Using the L_2-orthonormality of the ψ_k, the diagonal entries of the matrix become

$$\int_{-1}^{1}\varphi(x)^2\,\mathrm{d}x \;=\; \frac{2}{3} \;-\; \sum_{k=0}^{n-2} c_k^2 \;=\; \frac{2}{n\,(n+2)}.$$

As $\psi_k(-x) = (-1)^k\psi_k(x)$, its off-diagonal entries are correspondingly

$$\int_{-1}^{1}\varphi(x)\varphi(-x)\,\mathrm{d}x \;=\; \frac{1}{3} \;-\; \sum_{k=0}^{n-2}(-1)^k\,c_k^2 \;=\; \frac{2\,(-1)^{n+1}}{n\,(n+1)(n+2)}.$$

As the final results remain true for $n = 1$, this proves the proposition. □

The reduced total mass matrix is composed of the element mass matrices and is the sum of these. If the polynomial order is everywhere the same the following holds:

Lemma *If the polynomial order is n on all subintervals, the minimum and the maximum eigenvalue of the diagonally scaled reduced total mass matrix are*

$$1-\frac{1}{n+1}, \quad 1+\frac{1}{n+1}.$$

Proof We start from the generalized eigenvalue problem

$$\begin{pmatrix} (\varphi_1,\varphi_1) & (\varphi_1,\varphi_2) \\ (\varphi_2,\varphi_1) & (\varphi_2,\varphi_2) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} (\varphi_1,\varphi_1) & 0 \\ 0 & (\varphi_2,\varphi_2) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

for the just considered element matrix and its diagonal. The minimum and the maximum eigenvalue λ are by our first lemma the two values given above. Therefore

$$\left(1-\frac{1}{n+1}\right) v^T D v \le v^T M_0 v \le \left(1+\frac{1}{n+1}\right) v^T D v$$

for all coefficient vectors v of functions in $\mathscr{S}_0$, where M_0 is the reduced total mass matrix and D its diagonal. The given values thus represent a lower respectively upper bound for the eigenvalues of the diagonally scaled reduced total mass matrix.

The eigenvectors for the minimum and the maximum respectively the maximum and the minimum eigenvalue of the element matrix are the multiples of the vectors

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

depending on the polynomial degree n. Inserting above the coefficient vectors v of the functions in $\mathscr{S}_0$ with value 1 and alternately the values ± 1 at the intermediate points, one sees that both bounds are attained. This proves the proposition. $\square$

After diagonal scaling the condition number of the reduced total mass matrix is thus

$$\kappa = \frac{n+2}{n}$$

if the polynomial order is n on all subintervals. In this case therefore

$$q = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} < \frac{1}{2n}.$$

The L_2-projection Q onto the full space $\mathscr{S}$ thus remains H^1-stable as long as the length of neighbored subintervals differs at most by the factor of $2n$. The larger the polynomial degree is, the less stringent this condition becomes. By the same arguments as in Sect. 7 one can, however, show that the length of neighbored subintervals must not differ arbitrarily.

Even more interesting than the case of fixed polynomial degree is the hp-case in which the polynomial degree can change from subinterval to subinterval. Independent of that, $1/2$ remains a lower and $3/2$ an upper bound for the eigenvalues of the diagonally scaled reduced element matrices. Thus the estimate

$$\kappa \le 3$$

holds for the condition number of the diagonally scaled reduced total mass matrix. This implies that the L_2-projection onto such hp-spaces remains stable with respect to the weighted L_2-norm from Theorem 4.1 independent of the involved polynomial degrees as long as the length of neighbored subintervals differs at most by a factor

$$\mu < \frac{\sqrt{3}+1}{\sqrt{3}-1}.$$

The limiting factor in the hp-case is therefore the inverse inequality.

References

1. Bank, R.E.: PLTMG: A software package for solving elliptic partial differential equations, users' guide 11.0. Tech. rep., Department of Mathematics, University of California at San Diego (2012)
2. Bank, R.E., Dupont, T.: An optimal order process for solving finite element equations. Math. Comput. **36**, 35–51 (1981)
3. Bank, R.E., Nguyen, H.: *hp* adaptive finite elements based on derivative recovery and superconvergence. Computing and Visualization in Science **14**, 287–299 (2012)
4. Bank, R.E., Sherman, A.H., Weiser, A.: Refinement algorithms and data structures for regular local mesh refinement. In: Scientific Computing (Applications of Mathematics and Computing to the Physical Sciences) (R. S. Stepleman, ed.), pp. 3–17. North Holland (1983)
5. Bey, J.: Tetrahedral grid refinement. Computing **55**, 355–378 (1995)
6. Bramble, J., Pasciak, J., Steinbach, O.: On the stability of the L_2-projection in $H^1(\Omega)$. Math. Comp. **71**, 147–156 (2001)
7. Carstensen, C.: Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomée criterion for H^1-stability of the L_2-projection onto finite element spaces. Math. Comput. **71**, 157–163 (2001)
8. Carstensen, C.: An adaptive mesh-refining algorithm allowing for an H^1 stable L_2 projection onto Courant finite element spaces. Constr. Approx. **20**, 549–564 (2004)
9. Clément, P.: Approximation by finite element functions using local regularization. RAIRO Sér. Rouge Anal. Numér. **R-2**, 77–84 (1975)
10. Crouzeix, M., Thomée, V.: The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces. Math. Comput. **48**, 521–532 (1987)
11. Demko, S.: Inverses of band matrices and local convergence of spline projection. SIAM J. Numer. Anal. **14**, 616–619 (1977)
12. Demko, S., Moss, W.F., Smith, P.W.: Decay rates for inverses of band matrices. Math. Comput. **43**, 491–499 (1984)
13. Deuflhard, P., Hohmann, A.: Numerical Analysis in Modern Scientific Computing: An Introduction, *Texts in Applied Mathematics*, vol. 43. Springer, Berlin Heidelberg (2003)
14. Douglas, J., Dupont, T., Wahlbin, L.: The stability in L^q of the L^2-projection into finite element spaces. Numer. Math. **23**, 193–197 (1975)
15. Olver, F., Lozier, D., Boisvert, R., Clark, C. (eds.): NIST Handbook of Mathematical Functions. Cambridge University Press, Cambridge (2010). `http://dlmf.nist.gov/`
16. Xu, J.: Iterative methods by space decomposition and subspace correction. SIAM Review **34**, 581–613 (1992)
17. Yserentant, H.: Two preconditioners based on the multi-level splitting of finite element spaces. Numer. Math. **58**, 163–184 (1990)
18. Yserentant, H.: Old and new convergence proofs for multigrid methods. Acta Numerica **2**, 285–326 (1993)