Numer. Math. 72: 367–389 (1996)

Numerische Mathematik

Mixed *hp* finite element methods for problems in elasticity and Stokes flow

Rolf Stenberg[1], **Manil Suri**[2,⋆]

[1] Faculty of Mechanical Engineering, Helsinki University of Technology, 02150 Esbo, Finland
[2] Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21228, USA

Received March 4, 1994 / Revised version received February 12, 1995

Summary. We consider the mixed formulation for the elasticity problem and the limiting Stokes problem in $\mathbb{R}^d$, $d = 2, 3$. We derive a set of sufficient conditions under which families of mixed finite element spaces are simultaneously stable with respect to the mesh size h and, subject to a maximum loss of $O(k^{\frac{d-1}{2}})$, with respect to the polynomial degree k. We obtain asymptotic rates of convergence that are optimal up to $O(k^{\epsilon})$ in the displacement/velocity and up to $O(k^{\frac{d-1}{2}+\epsilon})$ in the "pressure", with $\epsilon > 0$ arbitrary (both rates being optimal with respect to h). Several choices of elements are discussed with reference to properties desirable in the context of the hp-version.

Mathematics Subject Classification (1991): 65N30

1. Introduction

We consider the equations of linear elasticity, given by[1]

$$2\mu(\varepsilon(\boldsymbol{u}), \varepsilon(\boldsymbol{v})) + \lambda(\operatorname{div} \boldsymbol{u}, \operatorname{div} \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{V} = [H_0^1(\Omega)]^d, \tag{1.1}$$

where $\boldsymbol{u}$ is the unknown displacement and μ, λ are the Lamé parameters. Here, $\Omega \subset \mathbb{R}^d, d = 2, 3$, is a polygonal or polyhedral domain. For the two-dimensional case, this is the plain strain problem. For notational simplicity, we consider only homogeneous Dirichlet boundary conditions, which is the "worst case" with respect to the stability and convergence of the methods to be analyzed.

⋆ The work of this author was supported in part by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant F49620-92-J-0100. It was performed while visiting the Helsinki University of Technology

[1] We will use the standard notation $H^k(S), H_0^k(S)$ for Sobolev spaces on S. $\|\cdot\|_{k,S}$ will denote the norm of $H^k(S)$ and $(\cdot,\cdot)_S$ the $L^2(S)$ inner product. The subscript S will be dropped when $S = \Omega$

The finite element analysis of problem (1.1) is complicated by the fact that when the second Lamé parameter is near the incompressible limit ($\lambda \to \infty$), one observes so-called "locking" phenomena. This occurs due to an inability of the finite element subspace for $\boldsymbol{u}$ to satisfy the limiting constraint of incompressibility,

$$\operatorname{div} \boldsymbol{u} = 0, \tag{1.2}$$

and still retain the required approximability properties. As a result, there can be a marked decrease in accuracy in computed values of interest such as the displacement $\boldsymbol{u}$ and the "pressure"

$$p = -\lambda \operatorname{div} \boldsymbol{u}. \tag{1.3}$$

This loss will be particularly evident when the standard formulation (1.1) is discretized using polynomials of low degree k, say $k = 1, 2$ (see e.g. [4]).

There are two main strategies for overcoming the effects of locking. The first is to reduce the severity of the constraint (1.2) by having it satisfied only approximately, using a mixed method. For this, (1.1) is written in the Herrmann variational form by taking the pressure p (given by (1.3)) as an independent unknown:

$$\begin{aligned} 2\mu(\varepsilon(\boldsymbol{u}), \varepsilon(\boldsymbol{v})) - (p, \operatorname{div} \boldsymbol{v}) &= (\boldsymbol{f}, \boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \\ \lambda^{-1}(p, q) + (\operatorname{div} \boldsymbol{u}, q) &= 0 \quad \forall q \in W. \end{aligned} \tag{1.4}$$

Here, $W = L^2_0(\Omega)$ is the subset of $L^2(\Omega)$ consisting of functions with zero mean value. (The pressure will have zero mean value due to Dirichlet boundary conditions.) Using an appropriate combination of *stable* finite element spaces for $\boldsymbol{u}$ and p can then lead to approximations for $\boldsymbol{u}$ and p that are uniformly optimal with respect to λ, even for low degree polynomials.

In the limit $\lambda \to \infty$, we obtain from (1.4) the equations

$$\begin{aligned} 2\mu(\varepsilon(\boldsymbol{u}), \varepsilon(\boldsymbol{v})) - (p, \operatorname{div} \boldsymbol{v}) &= (\boldsymbol{f}, \boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \\ (\operatorname{div} \boldsymbol{u}, q) &= 0 \quad \forall q \in W. \end{aligned} \tag{1.5}$$

For applications in fluid flow, this is the Stokes problem and then μ is the viscosity, $\boldsymbol{u}$ the velocity and p the pressure of the fluid. A locking-free method for (1.4) then automatically yields a corresponding method for (1.5) with the same convergence properties (and vice versa), so that our results in this paper will hold for Stokes flow as well.

An alternate method to avoid locking is to retain the standard formulation (1.1), but in conjunction with *higher-order* elements. For instance, in [15] it is shown that using the h-version on a class of triangular meshes with polynomial degree $k \geq 4$ completely eliminates locking for both $\boldsymbol{u}$ and p. (We will use k instead of p for the degrees, to avoid confusion with the pressure). For rectangular meshes, however, it has been shown in [2] that the h-version can never be made fully free of locking (as $\lambda \to \infty$), no matter how high k is chosen. The best uniform rate of convergence for $\boldsymbol{u}$ is always at least $O(h^{-1})$ worse than the optimal rate. Note that using a higher degree k not only increases α, the asymptotic order of convergence in the error bound Ch^{α}, but also leads to a reduction in the "constant" $C = C(k)$ (which is generally a decreasing function of k).

The use of high order elements can also be realized in terms of the p-version of the finite element method, where a fixed mesh is used (h constant) and k is increased to obtain accuracy. It was shown in [2, 19] that the p-version eliminates locking in $\boldsymbol{u}$, with an asymptotic rate of convergence which is optimal. However, the results in those papers do not guarantee that the pressures, when calculated by formula (1.3), will be free of locking. In fact, some locking could occur, due to the inf-sup condition for the underlying limit problem (1.5) being dependent on k. Such lack of stability has been investigated in [12] for certain polynomial spaces. To a large extent, however (for problems with smooth input data), this loss in convergence for the pressures is compensated for by the enhanced rates of convergence possible with higher order elements.

Let us mention that several mixed p-type elements have been studied for the Stokes problem in the context of the spectral element method [6]. In this regard, the "$[\mathbb{P}_k]^{\mathrm{N}} \times \mathbb{P}_{k-2}$" element (Method 5 below) is of particular interest since it is optimal for $\boldsymbol{u}$ and quite close to optimal (with the loss of only $O(k^{1/2})$) for p in two dimensions. The three-dimensional version of this element has been analyzed in [14]. (Our analysis here gives alternate proofs of the results in [6, 14] for this element.)

In this paper, we consider families of mixed methods for the elasticity and Stokes problems which are defined for each degree $k \geq 2$. We obtain bounds for the stability and asymptotic convergence of such methods which are uniform in *h and k.* This allows us to precisely characterize the dependence on k of the constant $C(k)$ when the h-version is used, thereby providing a better picture of the possible advantages of using a higher-order element from the family. Similarly, in the case of the p-version, our analysis characterizes the effect of using a more refined mesh.

We will be particularly interested in mixed *hp*-versions using such families of elements. Our estimates will then show the effect of simultaneously increasing the polynomial degree k and decreasing the mesh width h. In terms of locking, both the use of a mixed method and of high-order elements will help in decreasing such effects.

We will restrict our analysis to parallelogram and parallelepiped elements. This is because of the technical difficulty in establishing precise p-stability results over elements like triangles and tetrahedra, which lack suitable tensor product bases. The best result for triangles available in the literature is from 1983 (see [19]) and says that the stability of the p-version over a triangular mesh deteriorates no worse than $k^{-\kappa}$ for some (unknown) κ.

Our goal will be to formulate families of mixed methods which possess a set of desirable properties in the context of *hp*-extensions. More specifically, we will, among various elements defined on parallelograms (and parallelepipeds), characterize those that satisfy a stability condition in terms of both h and k, and that possess the correct approximability for both the velocities and the pressures (again in terms of *h and k*). A further consideration, if minimal degrees of freedom are desired, will be to limit the number of "internal" degrees of freedom. (We will only consider elements that can be defined in terms of external and internal shape functions for *hp*-codes described e.g. in [18].)

Let us note that the *hp*-version can, with proper mesh-degree selection, lead to *exponential* rates of convergence. We have not addressed issues related to such mesh refinement here. However, we do not assume quasiuniform meshes and our results hold in particular for implementations involving "hanging nodes" (see e.g. [10]), by which non-quasiuniform meshes can be constructed (for appropriate domains) using only parallelogram elements. Our results then establish *exponential convergence* for such implementations (see Remark 5.4). (A fully adaptive *hp* implementation has, in fact, been tested with excellent results, using our elements with hanging nodes, by A. Patra and J. T. Oden at TICAM, Austin, TX.)

The plan of our paper is as follows. In the next section we introduce our notation on polynomial subspaces. In Sect. 3 we first define the finite element method. Next, we give a general set of sufficient conditions to be satisfied by the finite element subspaces. Then we discuss various alternatives. Among these, we give an element (Method 1) which is the minimal covered by our analysis. We also discuss why the $[\mathbb{P}_k]^{\mathrm{N}} \times \mathbb{P}_{k-2}$ combination mentioned above (which is a good choice in terms of *p*-refinement), is not as suitable for the *hp*-version. Section 4 is devoted to a projection operator central for our analysis. In the last section we derive stability and convergence results in terms of h and k for the methods satisfying the sufficient conditions introduced earlier. The stability constant is shown to behave no worse than $Ck^{\frac{1-d}{2}}$. The velocities and pressures are shown to converge optimally for all λ, simultaneously in both h and k, except for a possible loss of k^{ϵ} in the velocity and $k^{\frac{d-1}{2}+\epsilon}$ in the pressure, for arbitrary $\epsilon > 0$.

There is a close connection between mixed elements for the Stokes problem and the so-called MITC elements for the Reissner-Mindlin plate equations, cf. [7, 8]. In [17] we utilize this connection and give an *hp*-error analysis for several rectangular MITC plate methods. Our results have been presented in [16].

2. Polynomial spaces

Let us recall that in the h-version, the shape functions used are generally of the nodal (Lagrangian) type, cf. e.g. [9]. However, in the p- and *hp*-versions, it is more advantageous to use *hierarchical* shape functions that are *non-nodal*. Defining the hierarchical shape functions in terms of the integrals of Legendre polynomials (cf. (2.1) below) helps in controlling the accumulation of round-off error. Also, separating the basis into two sets, *internal* shape functions that vanish on the element boundary and *external* shape functions that are non-zero on at least part of it, allows continuity requirements to be imposed in a natural way, purely through the external functions. Moreover, the internal shape functions may be condensed out at the local element level, so that they do not appear in the global stiffness matrix. Note that in h-version terminology the internal shape functions are the so-called "bubble" functions. See [18] for further details regarding basis functions for hp-codes. In this paper the stability of the methods is studied locally on each element, and in this, only the internal shape functions enter.

We denote by $\hat{I}_x$, $\hat{I}_y$ and $\hat{I}_z$ the reference intervals in the x, y and z variables respectively, viz. $\hat{I}_x = \{x \mid -1 \le x \le 1\}$. We will use the same notation $\hat{K}$ for the reference square $\hat{I}_x \times \hat{I}_y$ and cube $\hat{I}_x \times \hat{I}_y \times \hat{I}_z$. As usual, for $S \subset \mathbb{R}^l$, $l = 1, 2, 3$, we let $P_k(S)$ denote the set of polynomials of *total* degree k and $Q_k(S)$ denote the set of polynomials of degree k *in each variable*. Moreover, $Q'_k(S)$ will denote the "trunk" or "serendipity" space [9] of polynomials (defined below).

By $L_i(x)$, $i \ge 0$, we denote the Legendre polynomial of degree i, and for $i \ge 1$, we let

$$U_i(x) = \int_{-1}^{x} L_i(t)dt. \tag{2.1}$$

Let us define for $i \ge 0$, $\gamma_i = (2i+1)^{-1}$. Then we have

$$\int_{-1}^{+1} L_i(x)L_j(x)\,dx = \begin{cases} 2\gamma_i & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

Also,

$$U_i(x) = \gamma_i(L_{i+1}(x) - L_{i-1}(x)) \tag{2.3}$$

from which it follows that

$$U_i(\pm 1) = 0 \tag{2.4}$$

and (using (2.2)),

$$\int_{-1}^{1} U_i(x)U_j(x)\,dx = \begin{cases} \gamma_i^2(2\gamma_{i+1} + 2\gamma_{i-1}) & \text{if } i = j, \\ \gamma_i\gamma_{i-2}(-2\gamma_{i-1}) & \text{if } i = j+2, \\ \gamma_i\gamma_{i+2}(-2\gamma_{i+1}) & \text{if } i = j-2, \\ 0 & \text{otherwise.} \end{cases} \tag{2.5}$$

That the internal basis functions can be expressed by the integrals of the Legendre polynomials is a consequence of (2.4). Hence, for $k \ge 4$, we let

$$I_k(\hat{K}) = \{v \mid v = \sum_{2 \le i+j \le k-2} a_{ij} U_i(x)U_j(y),\ a_{ij} \in \mathbb{R}\} \tag{2.6}$$

be the internal shape functions of *total* degree k for the reference square, and for $k \ge 6$,

$$I_k(\hat{K}) = \{v \mid v = \sum_{3 \le i+j+l \le k-3} a_{ijl} U_i(x)U_j(y)U_l(z),\ a_{ijl} \in \mathbb{R}\} \tag{2.7}$$

be the internal shape functions of *total* degree k for the reference cube. The internal shape functions of degree k in *each variable* we denote by $J_k(\hat{K})$, i.e, for $k \ge 2$, (2.8) $$J_k(\hat{K}) \;=\; \{v | v = \sum_{i,j=1}^{k-1} a_{ij} U_i(x) U_j(y),\ a_{ij} \in \mathbb{R}\},$$

(2.9) $$J_k(\hat{K}) \;=\; \{v | v = \sum_{i,j,l=1}^{k-1} a_{ijl} U_i(x) U_j(y) U_l(z),\ a_{ijl} \in \mathbb{R}\}.$$

We point out that

(2.10) $$I_k(\hat{K}) \;=\; \{b_{\hat{K}} v | v \in P_{k-2d}(\hat{K})\},$$

(2.11) $$J_k(\hat{K}) \;=\; \{b_{\hat{K}} v | v \in Q_{k-2}(\hat{K})\},$$

where $b_{\hat{K}}$ is the usual basic "bubble function,"

(2.12) $$b_{\hat{K}} = \begin{cases} U_1(x)U_1(y), & \text{for } \hat{K} \subset \mathbb{R}^2, \\ U_1(x)U_1(y)U_1(z), & \text{for } \hat{K} \subset \mathbb{R}^3. \end{cases}$$

Next, let us define the sets of external basis functions we will use. For the reference square, we define these as

(2.13) $$E_k(\hat{K}) = P_1(\hat{I}_x)P_k(\hat{I}_y) \cup P_k(\hat{I}_x)P_1(\hat{I}_y).$$

For the unit square $\hat{K} \in \mathbb{R}^2$, we then have the following decomposition into external and internal shape functions

(2.14) $$Q_k(\hat{K}) = E_k(\hat{K}) \oplus J_k(\hat{K}) \quad \text{and} \quad Q'_k(\hat{K}) = E_k(\hat{K}) \oplus I_k(\hat{K}).$$

In the reference cube, we will use two alternatives for the set of external shape functions, depending on whether the functions are of type Q_k or Q'_k on each face:

(2.15) $$E_k(\hat{K}) = P_1(\hat{I}_x)Q_k(\hat{I}_y \times \hat{I}_z) + P_1(\hat{I}_y)Q_k(\hat{I}_x \times \hat{I}_z) + P_1(\hat{I}_z)Q_k(\hat{I}_x \times \hat{I}_y).$$

(2.16) $$E'_k(\hat{K}) = P_1(\hat{I}_x)Q'_k(\hat{I}_y \times \hat{I}_z) + P_1(\hat{I}_y)Q'_k(\hat{I}_x \times \hat{I}_z) + P_1(\hat{I}_z)Q'_k(\hat{I}_x \times \hat{I}_y),$$

Then, for the unit cube $\hat{K} \subset \mathbb{R}^3$, we have the following decomposition into external and internal shape functions

(2.17) $$Q_k(\hat{K}) = E_k(\hat{K}) \oplus J_k(\hat{K}) \quad \text{and} \quad Q'_k(\hat{K}) = E'_k(\hat{K}) \oplus I_k(\hat{K}).$$

Let us finally remark that external degrees of freedom are grouped into *nodal*, *edge* and *face* (in $\mathbb{R}^3$) shape functions, cf. [18].

3. The finite element methods

The finite element methods are of the form: find $(\boldsymbol{u}_n, p_n) \in \boldsymbol{V}_n \times W_n \subset \boldsymbol{V} \times W$ such that

$$
\begin{aligned}
2\mu(\varepsilon(\boldsymbol{u}_n), \varepsilon(\boldsymbol{v})) - (p_n, \operatorname{div} \boldsymbol{v}) &= (\boldsymbol{f}, \boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{V}_n, \\
\lambda^{-1}(p_n, q) + (\operatorname{div} \boldsymbol{u}_n, q) &= 0 \quad \forall q \in W_n.
\end{aligned} \tag{3.1}
$$

By eliminating the discrete pressure, one can express the above in terms of $\boldsymbol{u}_n$ as the only unknown,

$$
2\mu(\varepsilon(\boldsymbol{u}_n), \varepsilon(\boldsymbol{v})) + \lambda(\Pi_n \operatorname{div} \boldsymbol{u}_n, \Pi_n \operatorname{div} \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{V}_n, \tag{3.2}
$$

where Π_n is the L^2 projection onto W_n. The space W_n consists of functions discontinuous across element boundaries and hence the projection Π_n is calculated locally on each element.

We will define the finite element spaces $\boldsymbol{V}_n$ and W_n as follows. Let $\mathscr{C}_h$ be a parallelogram or parallelepiped mesh on Ω, not necessarily quasiuniform. We assume that $\mathscr{C}_h$ satisfies the usual compatibility conditions and is *regular* in the sense of [9]. For $K \in \mathscr{C}_h$, let $\boldsymbol{F}_K$ be the affine mapping from the reference square or cube $\hat{K}$ onto K.

Let $\boldsymbol{V}_k(\hat{K})$ and $W_k(\hat{K})$ be families of polynomial spaces for the velocity and pressure on $\hat{K}$, with the parameter k related to the degree. On $\mathscr{C}_h$, we then define

$$
\boldsymbol{V}_k(K) = \{ \, \boldsymbol{v} = \hat{\boldsymbol{v}} \circ \boldsymbol{F}_K^{-1} \mid \hat{\boldsymbol{v}} \in \boldsymbol{V}_k(\hat{K}) \, \}, \tag{3.3}
$$

$$
W_k(K) = \{ \, p = \hat{p} \circ \boldsymbol{F}_K^{-1} \mid \hat{p} \in W_k(\hat{K}) \, \}. \tag{3.4}
$$

Then, with $n = n(h, k)$, the finite element spaces are defined as

$$
\boldsymbol{V}_n = \{ \boldsymbol{v} \in \boldsymbol{V} \mid \ \boldsymbol{v}|_K \in \boldsymbol{V}_k(K) \ \ \forall K \in \mathscr{C}_h \}, \tag{3.5}
$$

$$
W_n = \{ p \in W \mid \ p|_K \in W_k(K) \ \ \forall K \in \mathscr{C}_h \}. \tag{3.6}
$$

Note that $\boldsymbol{V}_n$ will consist of functions continuous on Ω.

Let us next state the set of sufficient conditions under which our analysis is valid. We first assume that identical subspaces are used for all components of the velocity, i.e.,

(A1) $\boldsymbol{V}_k(\hat{K}) = [V_k(\hat{K})]^d$ for $d = 2, 3$.

Let $V_k^0(\hat{K}) = V_k(\hat{K}) \cap H_0^1(\hat{K})$ denote the set of internal shape functions used for all components of the displacement, so that $\boldsymbol{V}_k^0(\hat{K}) = \boldsymbol{V}_k(\hat{K}) \cap [H_0^1(\hat{K})]^d = [V_k^0(\hat{K})]^d$. We note that there exists a space $X_k(\hat{K})$ such that

$$
V_k^0(\hat{K}) = \{ v | v = b_{\hat{K}} w, \ w \in X_k(\hat{K}) \}, \tag{3.7}
$$

where $b_{\hat{K}}$ is as in (2.12). Then we assume that the space for the pressure satisfies the following condition.

(A2) $\nabla q \in [X_k(\hat{K})]^d \quad \forall q \in W_k(\hat{K})$.

Due to the definitions of the spaces $V_k^0(\hat{K})$ and $X_k(\hat{K})$, we can define a weighted L^2 projection $\mathscr{T}_k : H_0^1(\hat{K}) \to V_k^0(\hat{K})$ by

$$(v - \mathscr{T}_k v, w)_{\hat{K}} = 0 \quad \forall w \in X_k(\hat{K}). \tag{3.8}$$

We assume the following.

(A3) The projection operator $\mathscr{T}_k$ satisfies

$$\|\mathscr{T}_k u\|_{1,\hat{K}} \le C k^{\frac{d-1}{2}} \|u\|_{1,\hat{K}}.$$

The local stability condition we will prove is a consequence of the above assumptions (A1) - (A3). To have a global condition we need the additional assumption,

(A4) $[E_2(\hat{K})]^d \subset \boldsymbol{V}_k(\hat{K})$.

Finally, we will assume that the spaces $\boldsymbol{V}_k(\hat{K})$ and $W_k(\hat{K})$ contain polynomials of degree k and $k-1$, respectively, in order to get the optimal convergence rate in both h and k. More precisely, we assume the following.

(A5) $[Q_k'(\hat{K})]^d \subset \boldsymbol{V}_k(\hat{K})$.

(A6) $P_{k-1}(\hat{K}) \subset W_k(\hat{K})$.

The method is then defined by specifying the spaces $\boldsymbol{V}_k(\hat{K})$, $W_k(\hat{K})$, and below we will discuss various alternatives. As we shall see, the selection of admissible spaces essentially reduces to ensuring that (A2) is satisfied. For all methods we will assume that $k \ge 2$, i.e. we exclude the case with with piecewise constant pressures. (A low order stable method with this pressure space may be found in [11, p. 134].)

Method 1. We first consider the case when we choose $W_k(\hat{K})$ to be the minimal possible space satisfying (A6), i.e.

$$W_k(\hat{K}) = P_{k-1}(\hat{K}).$$

Since $\nabla q \in [P_{k-2}(\hat{K})]^d$ for all $q \in W_k(\hat{K})$, (A2) is satisfied if we choose $X_k(\hat{K}) = P_{k-2}(\hat{K})$, i.e. $V_k^0(\hat{K}) = \{v | v = b_{\hat{K}} w,\ w \in P_{k-2}(\hat{K})\}$. From (2.10) we then see that $V_k^0(\hat{K}) = I_{k+2d-2}(\hat{K})$. For the two dimensional case ($d = 2$) the assumptions (A4) and (A5) then give the following space for the deflection

$$\boldsymbol{V}_k(\hat{K}) = [E_k(\hat{K}) \oplus I_{k+2}(\hat{K})]^2.$$

(Note that this is the space $[Q_k(\hat{K}) \cap P_{k+2}(\hat{K})]^2$.)
Since our aim is to choose the space $\boldsymbol{V}_k(\hat{K})$ as small as possible it would be natural to choose external degrees of freedom of the type $E_k'(\hat{K})$ in the three dimensional case. For $k = 2, 3$, this is not sufficient to ensure the condition (A4), so we actually take ($d = 3$)

$$\boldsymbol{V}_k(\hat{K}) = \begin{cases} [(E_k'(\hat{K}) \cup E_2(\hat{K})) \oplus I_{k+4}(\hat{K})]^3 & \text{for } k = 2, 3, \\ [E_k'(\hat{K}) \oplus I_{k+4}(\hat{K})]^3 & \text{for } k \ge 4. \end{cases} \tag{3.9}$$

With this choice for $V_k^0(\hat{K})$ and $X_k(\hat{K})$ we denote the projection operator $\mathscr{T}_k$ by $\mathscr{R}_k$. In the next section we prove that (A3) holds for this projection.

Obviously, any combination $\boldsymbol{V}_k(\hat{K}) \times P_{k-1}(\hat{K})$ for which the displacement space $\boldsymbol{V}_k(\hat{K})$ contains the one of Method 1 above is stable, but with more than the minimal number of functions. We could, for instance, choose the following alternatives.

Method 2.

$$\boldsymbol{V}_k(\hat{K}) = [Q_k(\hat{K})]^d, \quad W_k(\hat{K}) = P_{k-1}(\hat{K}). \tag{3.10}$$

This combination is well known to be stable in the h-version [11, pp. 156–157]. Note that for large k this choice will lead to $O(k^d)$ more degrees of freedom for the displacement than for the previous method. However, since the pressure space is unchanged, there is no reason to expect better accuracy.

Method 3.

$$\boldsymbol{V}_k(\hat{K}) = [Q'_{k+2d-2}(\hat{K})]^d, \quad W_k(\hat{K}) = P_{k-1}(\hat{K}). \tag{3.11}$$

A possible advantage of this choice is that the basis functions are of a standard type used in *hp*-codes. The number of degrees of freedom for the displacement is now a fixed number higher than than those of Method 1 (for $d = 2$ this number is 16). The approximability for the displacement (in term of h) is much better now, being $2d - 2$ orders higher. However, since the pressure space is unchanged, this additional approximability will not, in general, translate into better computational results, so again no advantage over Method 1 can be guaranteed (in terms of the asymptotic rate of convergence).

Next, we will discuss alternatives in which the pressure is in Q_{k-1}.

Method 4. We choose

$$W_k(\hat{K}) = Q_{k-1}(\hat{K}). \tag{3.12}$$

The smallest choice for $X_k(\hat{K})$ for which (A2) holds, i.e. for which $\nabla q \in [X_k(\hat{K})]^d$ for all $q \in W_k(\hat{K})$, is now $X_k(\hat{K}) = Q_{k-1}(\hat{K})$. Hence, (3.7) and (2.11) show that we should choose $V_k^0(\hat{K}) = J_{k+1}(\hat{K})$. A natural choice is then to take the external shape functions as $E_k(\hat{K})$. The displacement space we get is then

$$\boldsymbol{V}_k(\hat{K}) = [E_k(\hat{K}) \oplus J_{k+1}(\hat{K})]^d. \tag{3.13}$$

With this choice for $V_k^0(\hat{K})$ and $X_k(\hat{K})$ we denote the projection operator $\mathscr{T}_k$ by $\mathscr{S}_{k+1}$. (This convention for the indices will be convenient in the analysis.) In the next section we will prove (A3) for this projection.

Again, we note that we can increase the displacement space from that of the above method. By this we get the following convenient alternative.

Method 5.

$$\boldsymbol{V}_k(\hat{K}) = [Q_{k+1}(\hat{K})]^d, \quad W_k(\hat{K}) = Q_{k-1}(\hat{K}). \tag{3.14}$$

This choice has previously been analyzed for the case of the pure p-version on a single square element in [6, 14] (the "$[\mathbb{P}_k]^{\mathrm{N}} \times \mathbb{P}_{k-2}$" spectral element). Our general proof leads to the same estimates when applied to this case. Note that in terms of h-approximability the velocity and pressure spaces are not correctly matched, since the $O(h^{k+1})$ approximablity in the H^1 norm for the displacement will be dominated by the error in pressures, which are only $O(h^k)$ accurate in L_2. Hence, for the hp-version, this space is once again not as suitable as Method 1, in terms of approximability.

Our final alternative will be the following.

Method 6.

$$\boldsymbol{V}_k(\hat{K}) = [Q_k(\hat{K})]^d, \quad W_k(\hat{K}) = Q_{k-2}(\hat{K}) \cup P_{k-1}(\hat{K}). \tag{3.15}$$

We note that it now holds that $\nabla q \in [Q_{k-2}(\hat{K})]^d$, and hence the analysis of this method follows from that of the previous one (with a change in the indices).

This alternative appears to be quite useful. It uses standard basis functions for the displacement, the only variable which enters into the calculation after condensing the pressure. Also, with respect to the h-convergence, it has a correct matching of polynomial degrees for the displacement and pressure. The difference from the previous method is that we now have a *maximal* pressure space corresponding to the space selected for the velocities.

Note that this choice gives the *maximal* pressure space satisfying the assumptions, when $\boldsymbol{V}_k(\hat{K})$ is chosen to be $[Q_k(\hat{K})]^d$.

Remark 3.1. If the elements are all *rectangular* (as opposed to parallelogram or parallelepiped), condition (A1) is not necessary. For such purely rectangular (or brick) elements, $V_k(\hat{K})$ can be different for each of the d different components. This allows one to further reduce the number of degrees of freedom for the minimal velocity space given by (3.13) (Method 4) for the case when the pressures are in Q_{k-1}. For $d = 2$ (rectangular elements), we may define

$$\boldsymbol{V}_k(\hat{K}) = [E_k(\hat{K}) \oplus M_k^1(\hat{K})] \times [E_k(\hat{K}) \oplus M_k^2(\hat{K})] \tag{3.16}$$

where

$$\begin{aligned} M_k^1(\hat{K}) &= \{b_{\hat{K}} v | v = x^r y^s,\ 0 \le r \le k-2,\ 0 \le s \le k-1\}, \\ M_k^2(\hat{K}) &= \{b_{\hat{K}} v | v = x^r y^s,\ 0 \le r \le k-1,\ 0 \le s \le k-2\}. \end{aligned}$$

For $d = 3$ (brick elements), (3.16) will now have three components, with

$$M_k^1(\hat{K}) = \{b_{\hat{K}} v | v = x^r y^s z^t,\ 0 \le r \le k-2,\ 0 \le s \le k-1,\ 0 \le t \le k-1\}$$

and M_k^2, M_k^3 defined analogously.

Remark 3.2. Our definition of the methods also covers the case of a general mesh where the mappings $\boldsymbol{F}_K$ are not assumed to be affine. The analysis is, however, only valid for the affine case. We would like to point out here that there is an alternative way of defining elements in which the pressures are assumed to be polynomials in the "global coordinates" and not in the "local coordinates" as above. We could, for example, in Method 2 let

$$W_k(K) = P_{k-1}(K) \tag{3.17}$$

and then define the pressure space by (3.6). By this we get the method that has been traditionally considered, and for which the h-stability has been established for meshes with bi- and trilinear mappings $\boldsymbol{F}_K$, cf. [11, pp. 156-57].

4. The projection operator

In this section we will analyze the projection operator $\mathscr{T}_k$, defined by (3.8), for the two different cases:

$$\mathscr{T}_k = \mathscr{R}_k \quad \text{for} \quad V_k^0(\hat{K}) = I_{k+2d-2}(\hat{K}) \quad \text{and} \quad X_k(\hat{K}) = P_{k-2}(\hat{K}),$$

and

$$\mathscr{T}_k = \mathscr{S}_{k+1} \quad \text{for} \quad V_k^0(\hat{K}) = J_{k+1}(\hat{K}) \quad \text{and} \quad X_k(\hat{K}) = Q_{k-1}(\hat{K}).$$

For the second case this means that $\mathscr{S}_k : H_0^1(\hat{K}) \to J_k(\hat{K})$ is defined through

$$(v - \mathscr{S}_k v, w)_{\hat{K}} = 0 \quad \forall w \in Q_{k-2}(\hat{K}). \tag{4.1}$$

By $I(\hat{K})$ we denote the set of all polynomials that vanish on $\partial\hat{K}$. Then it is easy to see that for d=2 (d=3) the functions $U_i(x)U_j(y)$ ($U_i(x)U_j(y)U_l(z)$) form a basis for $I(\hat{K})$. We therefore have for any $u \in I(\hat{K})$,

$$u(x,y) = \sum_{i,j=1}^{\infty} a_{ij} U_i(x) U_j(y) \tag{4.2}$$

and

$$u(x,y,z) = \sum_{i,j,l=1}^{\infty} a_{ijl} U_i(x) U_j(y) U_l(z), \tag{4.3}$$

in $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively, where $a_{ij} \neq 0$ and $a_{ijl} \neq 0$ for only a finite number of indices i, j and l.

Next, we observe that when the projection operators are restricted to act in $I(\hat{K})$, they have the following simple characterizations.

Lemma 4.1. *For the projection operators* $\mathscr{R}_k : I(\hat{K}) \to I_{k+2d-2}(\hat{K})$ *and* $\mathscr{S}_k : I(\hat{K}) \to J_k(\hat{K})$ *it holds that:*

$$\mathscr{R}_k u(x,y) = \sum_{2\le i+j\le k} a_{ij} U_i(x) U_j(y), \tag{4.4}$$

$$\mathscr{S}_k u(x,y) = \sum_{i,j=1}^{k-1} a_{ij} U_i(x) U_j(y), \tag{4.5}$$

for $\hat{K} \subset \mathbb{R}^2$*, and*

$$\mathscr{R}_k u(x,y,z) = \sum_{3\le i+j+l\le k+1} a_{ijl} U_i(x) U_j(y) U_l(z), \tag{4.6}$$

$$\mathscr{S}_k u(x,y,z) = \sum_{i,j,l=1}^{k-1} a_{ijl} U_i(x) U_j(y) U_l(z), \tag{4.7}$$

for $\hat{K} \subset \mathbb{R}^3$*, where* $u \in I(\hat{K})$ *is given by (4.2) and (4.3), respectively.*

Proof. Consider the case $\mathscr{R}_k$ for $d = 2$. Let $u \in I(\hat{K})$ be given by (4.2), and let

$$\mathscr{R}_k u(x,y) = \sum_{2\le i+j\le k} b_{ij} U_i(x) U_j(y).$$

Due to the linear independence of the derivatives of the Legendre polynomials, we have the characterization

$$P_{k-2}(\hat{K}) = \{v | v = \sum_{2\le l+m\le k} c_{lm} U_l''(x) U_m''(y),\ c_{lm} \in \mathbb{R}\}, \quad \text{for} \quad d = 2. \tag{4.8}$$

Let w be one of the basis functions of $P_{k-2}(\hat{K})$ from the above characterization, i.e., $w = U_l''(x)U_m''(y)$, with $2 \le l+m \le k$. Integrating by parts and using (2.1) - (2.4) we get

$$(u,w)_{\hat{K}} = \sum_{i,j=1}^{\infty} a_{ij} \int_{-1}^{1} U_i'(x) U_l'(x)\, dx \int_{-1}^{1} U_j'(y) U_m'(y) dy = 4\gamma_l \gamma_m a_{lm} . \tag{4.9}$$

A similar calculation shows that

$$(\mathscr{R}_k u, w)_{\hat{K}} = 4\gamma_l \gamma_m b_{lm} \tag{4.10}$$

and hence by the definition of $\mathscr{R}_k$,

$$b_{lm} = a_{lm}, \quad \text{for} \quad 2 \le l+m \le k. \tag{4.11}$$

The other cases are treated in exactly the same manner by using the fact that

$$P_{k-2}(\hat{K}) = \{v | v = \sum_{3\le i+j+l\le k} c_{ijl} U_i''(x) U_j''(y) U_l''(z),\ c_{ijl} \in \mathbb{R}\}, \quad \text{for} \quad d = 3, \tag{4.12}$$

and

$$Q_{k-2}(\hat{K}) = \{v | v = \sum_{i,j=1}^{k-1} c_{ij} U_i''(x) U_j''(y),\ c_{ij} \in \mathbb{R}\}, \quad \text{for} \quad d = 2, \tag{4.13}$$

$$Q_{k-2}(\hat{K}) = \{v | v = \sum_{i,j,l=1}^{k-1} c_{ijl} U_i''(x) U_j''(y) U_l''(z),\ c_{ijl} \in \mathbb{R}\}, \quad \text{for} \quad d = 3. \quad \square \tag{4.14}$$

Let us define $a_{ij}, a_{ijl} = 0$, for i, j or $l = 0, -1$. Then we have the following result in $\mathbb{R}^2$ (with $u_x = \partial u/\partial x$ and $u_y = \partial u/\partial y$).

Lemma 4.2. *For $u \in I(\hat{K})$ given by (4.2),*

$$\|u_x\|^2_{0,\hat{K}} = \sum_{j=0}^{\infty} \sum_{i=1}^{\infty} 4\gamma_i \gamma_j (\gamma_{j-1} a_{i,j-1} - \gamma_{j+1} a_{i,j+1})^2 \tag{4.15}$$

and

$$\|u_y\|^2_{0,\hat{K}} = \sum_{j=1}^{\infty} \sum_{i=0}^{\infty} 4\gamma_i \gamma_j (\gamma_{i-1} a_{i-1,j} - \gamma_{i+1} a_{i+1,j})^2. \tag{4.16}$$

(The summations in (4.15) and (4.16) are obviously finite.)

Proof. Using (4.2) and the definition of $U_i(x)$, we have

$$u_x = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} L_i(x) U_j(y),$$

so that, using (2.2), (2.5),

$$\begin{aligned}
\|u_x\|^2_{0,\hat{K}} &= \sum_{i=1}^{\infty} 2\gamma_i \int_{-1}^{1} (\sum_{j=1}^{\infty} a_{ij} U_j(y))^2 dy \\
&= \sum_{i=1}^{\infty} 2\gamma_i \sum_{j=1}^{\infty} (a_{ij}^2 \int_{-1}^{+1} U_j^2(y) dy + 2a_{ij} a_{i,j+2} \int_{-1}^{+1} U_j(y) U_{j+2}(y) dy) \\
&= \sum_{i=1}^{\infty} 2\gamma_i (\sum_{j=1}^{\infty} a_{ij}^2 \gamma_j^2 (2\gamma_{j-1} + 2\gamma_{j+1}) + 2a_{ij} a_{i,j+2} \gamma_j \gamma_{j+2} (-2\gamma_{j+1})).
\end{aligned}$$

Rearranging the terms in the inner sum, (4.15) follows from this. The expression for $\|u_y\|^2_{0,\hat{K}}$ is obtained analogously. $\square$

The three-dimensional analog of Lemma 4.2, given below, may be established by similar methods.

Lemma 4.3. *For $u \in I(\hat{K})$ given by (4.3),*

$$\|u_x\|^2_{0,\hat{K}} = \sum_{l=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=1}^{\infty} 8\gamma_i\gamma_j\gamma_l((\gamma_{j-1}\gamma_{l-1}a_{i,j-1,l-1} - \gamma_{j-1}\gamma_{l+1}a_{i,j-1,l+1}) - (\gamma_{j+1}\gamma_{l-1}a_{i,j+1,l-1} - \gamma_{j+1}\gamma_{l+1}a_{i,j+1,l+1}))^2 \,. \tag{4.17}$$

Also, $\|u_y\|^2_{0,\hat{K}}$, $\|u_z\|^2_{0,\hat{K}}$ are given by corresponding permutations of (4.17).

Then we have the following result.

Theorem 4.1. *There exists a constant C independent of k such that for any $u \in H^1_0(\hat{K})$,*

$$\|\mathscr{T}_k u\|_{1,\hat{K}} \le Ck^{\frac{d-1}{2}}\|u\|_{1,\hat{K}}, \quad \text{for} \quad \mathscr{T}_k = \mathscr{R}_k, \mathscr{S}_{k+1}. \tag{4.18}$$

Proof. Suppose that we first prove the estimate for the operator $\mathscr{T}_k$ defined only in $I(\hat{K})$. Since $I(\hat{K})$ is dense in $H^1_0(\hat{K})$, an elementary theorem (cf. e.g. [1, pp. 13–14]) states that this operator has a unique norm-preserving extension to $H^1_0(\hat{K})$. It is now easily seen that this extension coincides with $\mathscr{T}_k$ as defined by (3.8) over the whole of $H^1_0(\hat{K})$. Hence, it is sufficient to prove the estimate for $u \in I(\hat{K})$.

For this, we first consider $\mathscr{S}_k$ in the two-dimensional case, with u given by (4.2). $\mathscr{S}_k u$ is then given by (4.5). We write the sum in (4.15) as

$$\sum_{j=0}^{\infty}\sum_{i=1}^{\infty} = \sum_{j=0}^{k-2}\sum_{i=1}^{k-1} + \sum_{j=k-1}^{k}\sum_{i=1}^{k-1}$$

(all other terms being zero). From this,

$$\begin{aligned}\|\mathscr{S}_k u_x\|^2_{0,\hat{K}} &= \sum_{j=0}^{k-2}\sum_{i=1}^{k-1} 4\gamma_i\gamma_j(\gamma_{j-1}a_{i,j-1} - \gamma_{j+1}a_{i,j+1})^2 \\ &\quad + \sum_{i=1}^{k-1}(4\gamma_i\gamma_{k-1}(\gamma_{k-2}a_{i,k-2})^2 + 4\gamma_i\gamma_k(\gamma_{k-1}a_{i,k-1})^2) \\ &\le \|u_x\|^2_{0,\hat{K}} + B \,.\end{aligned} \tag{4.19}$$

To bound B, let k be odd (say). Then for each $i = 1, \dots, k-1$,

$$\gamma_{k-2}a_{i,k-2} = -\sum_{m=0}^{\frac{k-3}{2}}(\gamma_{2m-1}a_{i,2m-1} - \gamma_{2m+1}a_{i,2m+1}) \tag{4.20}$$

(where $a_{i,-1} \equiv 0$). Hence,

$$(\gamma_{k-2}a_{i,k-2})^2 \le \frac{k-1}{2}\sum_{m=0}^{\frac{k-3}{2}}(\gamma_{2m-1}a_{i,2m-1} - \gamma_{2m+1}a_{i,2m+1})^2$$

so that, using the fact that $\gamma_{k-1} \le \gamma_j$ for $j \le k-1$,

$$(4.21)\quad 4\gamma_i\gamma_{k-1}(\gamma_{k-2}a_{i,k-2})^2 \le \frac{k-1}{2}\sum_{m=0}^{\frac{k-3}{3}} 4\gamma_i\gamma_{2m}(\gamma_{2m-1}a_{i,2m-1}-\gamma_{2m+1}a_{i,2m+1})^2.$$

Using a similar argument for $4\gamma_i\gamma_k(\gamma_{k-1}a_{i,k-1})^2$, we have

$$\begin{aligned} B &\le Ck\sum_{i=1}^{k-1}\sum_{j=0}^{k-2} 4\gamma_i\gamma_j(\gamma_{j-1}a_{i,j-1}-\gamma_{j+1}a_{i,j+1})^2 \\ &\le Ck\|u_x\|^2_{0,\hat{K}}, \qquad (4.22)\end{aligned}$$

by (4.15). Hence, combining (4.19), (4.22), we have

$$(4.23)\qquad \|\mathscr{S}_k u_x\|^2_{0,\hat{K}} \le Ck\|u_x\|^2_{0,\hat{K}}.$$

An analogous bound holds for $\|\mathscr{S}_k u_y\|_{0,\hat{K}}$, from which the theorem follows for $\mathscr{S}_k$.

For the other case $\mathscr{R}_k$, we may again obtain an estimate of the form

$$(4.24)\qquad \|\mathscr{R}_k u_x\|^2_{0,\hat{K}} \le \|u_x\|^2_{0,\hat{K}} + B.$$

Here, B is now the sum of $O(k)$ terms of the form $4\gamma_i\gamma_{k-i}(\gamma_{k-i-1}a_{i,k-i-1})^2$ and $4\gamma_i\gamma_{k-i+1}(\gamma_{k-i}a_{i,k-i})^2$. It can be seen that each of these terms will again satisfy the bound in (4.21). Hence, (4.23) will hold for $\mathscr{R}_k$ as well, and the theorem follows for $d = 2$.

Let us now prove the theorem for $\mathscr{S}_k$, for the case $d = 3$. Analogously to the preceding proof, we first note that the sum in (4.17) for $\|\mathscr{S}_k u_x\|^2_{0,\hat{K}}$ may be decomposed as

$$\begin{aligned}\sum_{l=0}^{k}\sum_{j=0}^{k}\sum_{i=1}^{k-1} &= \left(\sum_{l=0}^{k-2}\sum_{j=0}^{k-2}\sum_{i=1}^{k-1}\right) + \left(\sum_{l=k-1}^{k}\sum_{j=0}^{k-2}\sum_{i=1}^{k-1} + \sum_{l=0}^{k-2}\sum_{j=k-1}^{k}\sum_{i=1}^{k-1}\right) \\ &\quad + \left(\sum_{l=k-1}^{k}\sum_{j=k-1}^{k}\sum_{i=1}^{k-1}\right) = A + B + D\end{aligned}$$

where, using (4.17), it is easy to see that

$$A \le \|u_x\|^2_{0,\hat{K}}.$$

Let us now bound a typical term from B, say

$$\sum_{j=0}^{k-2}\sum_{i=1}^{k-1} 8\gamma_i\gamma_j\gamma_{k-1}((\gamma_{j-1}\gamma_{k-2}a_{i,j-1,k-2}) - (\gamma_{j+1}\gamma_{k-2}a_{i,j+1,k-2}))^2.$$

Then, analogously to (4.21), we have (for k odd e.g.), for each $i = 1,\ldots,k-1$, $j = 0,1,\ldots,k-2$,

$$\gamma_{j-1}\gamma_{k-2}a_{i,j-1,k-2} - \gamma_{j+1}\gamma_{k-2}a_{i,j+1,k-2} = \\ -\sum_{m=0}^{\frac{k-3}{2}}((\gamma_{j-1}\gamma_{2m-1}a_{i,j-1,2m-1} - \gamma_{j-1}\gamma_{2m+1}a_{i,j-1,2m+1}) \\ -(\gamma_{j+1}\gamma_{2m-1}a_{i,j+1,2m-1} - \gamma_{j+1}\gamma_{2m+1}a_{i,j+1,2m+1})) \tag{4.25}$$

from which, using the same arguments as those leading to (4.22), we may again establish

$$B \le Ck\|u_x\|^2_{0,\hat{K}}.$$

Finally, consider a term from D, e.g.

$$\sum_{i=1}^{k-1} 8\gamma_i\gamma_{k-1}\gamma_{k-1}(\gamma_{k-2}\gamma_{k-2}a_{i,k-2,k-2})^2.$$

We have for $i = 1, 2, \ldots, k-1$, k odd,

$$\gamma_{k-2}\gamma_{k-2}a_{i,k-2,k-2} = \\ -\sum_{n=0}^{\frac{k-3}{2}}\sum_{m=0}^{\frac{k-3}{2}}((\gamma_{2n-1}\gamma_{2m-1}a_{i,2n-1,2m-1} - \gamma_{2n-1}\gamma_{2m+1}a_{i,2n-1,2m+1}) \\ -(\gamma_{2n+1}\gamma_{2m-1}a_{i,2n+1,2m-1} - \gamma_{2n+1}\gamma_{2m+1}a_{i,2n+1,2m+1}))\ .$$

Note that unlike (4.21) or (4.25), there are $O(k^2)$ (not $O(k)$) terms in the above sum. We may now use a similar argument as that leading to (4.22), except that due to the $O(k^2)$ terms, we now get

$$D \le Ck^2\|u_x\|^2_{0,\hat{K}}.$$

The result follows for $\mathscr{S}_k$, and, by similar arguments, for the case $\mathscr{R}_k$ as well. □

Remark 4.1. For $\mathscr{S}_k$, the estimate (4.18) is *sharp*, i.e. there exists a sequence $u^{(k)} \in H^1_0(\hat{K})$, $\|u^{(k)}\|_{1,\hat{K}} = 1$, such that

$$\|\mathscr{S}_k u^{(k)}\|_{1,\hat{K}} \ge Ck^{\frac{d-1}{2}}.$$

This follows from the results in [6, 14], as explained in Remark 5.1 ahead.

Remark 4.2. It may be easily seen that defining one-dimensional analog of $\mathscr{T}_k$, the estimate (4.18) in Theorem 4.1 will hold for the case $d = 1$ as well.

5. The stability and error analysis

5.1. The inf-sup condition

As is well known [7], the convergence of mixed methods depends not only on the approximability of the spaces $\boldsymbol{V}_n$, W_n, but also on the satisfaction of an inf-sup or stability condition between them. In this section, we establish such an inf-sup condition in terms of both h and k, for a general FEM that satisfies the conditions (A1) - (A6). The main theorem is the following.

Theorem 5.1. *Let the spaces $\boldsymbol{V}_n$, W_n satisfy assumptions (A1)–(A4). Then for $d = 2, 3$*

$$\sup_{\boldsymbol{v}\in\boldsymbol{V}_n\setminus\{0\}} \frac{(\operatorname{div}\boldsymbol{v}, q)}{\|\boldsymbol{v}\|_1} \geq Ck^{-(\frac{d-1}{2})}\|q\|_0 \quad \forall q \in W_n, \tag{5.1}$$

where the constant C is independent of h, k and q.

This stability estimate follows from the corresponding local estimate on each element:

Lemma 5.1. *Let the spaces $\boldsymbol{V}_k(\hat{K})$, $W_k(\hat{K})$ satisfy assumptions (A1)–(A3). Then for every $q^* \in W_k(K) \cap L^2_0(K)$ there exists $\boldsymbol{v}^* \in \boldsymbol{V}_k(K) \cap [H^1_0(K)]^d$ such that*

$$(\operatorname{div}\boldsymbol{v}^*, q^*)_K \geq C_1 k^{-(\frac{d-1}{2})}\|q^*\|^2_{0,K} \quad \textit{and} \quad |\boldsymbol{v}^*|_{1,K} \leq C_2\|q^*\|_{0,K} \tag{5.2}$$

where C_1, C_2 are positive constants independent of K, h, k and q^*.

Proof. Let $q^* \in W_k(K) \cap L^2_0(K)$ be arbitrary and define $\hat{q}^* \in W_k(\hat{K}) \cap L^2_0(\hat{K})$ by $\hat{q}^* = q^* \circ \boldsymbol{F}_K$ ($\hat{q}^*$ is in $L^2_0(\hat{K})$ since $\boldsymbol{F}_K$ is affine).

On the reference element $\hat{K}$, the continuous inf-sup condition holds. Hence, there is $\hat{\boldsymbol{v}} \in [H^1_0(\hat{K})]^d$ such that

$$(\operatorname{div}\hat{\boldsymbol{v}}, \hat{q}^*)_{\hat{K}} \geq C\|\hat{q}^*\|^2_{0,\hat{K}} \quad \text{and} \quad |\hat{\boldsymbol{v}}|_{1,\hat{K}} \leq \|\hat{q}^*\|_{0,\hat{K}}, \tag{5.3}$$

with C independent of $\hat{q}^*$ and $\hat{\boldsymbol{v}}$.

Define $\hat{\boldsymbol{v}}^* \in \boldsymbol{V}^0_k(\hat{K})$ by $\hat{\boldsymbol{v}}^* = \mathscr{T}_k\hat{\boldsymbol{v}}$. By (A2), the definition of $\mathscr{T}_k$, (5.3) and (A3), we then obtain (integrating by parts)

$$\begin{aligned}(\operatorname{div}\hat{\boldsymbol{v}}^*, \hat{q}^*)_{\hat{K}} &= -(\hat{\boldsymbol{v}}^*, \nabla\hat{q}^*)_{\hat{K}} = -(\mathscr{T}_k\hat{\boldsymbol{v}}, \nabla\hat{q}^*)_{\hat{K}}\\ &= -(\hat{\boldsymbol{v}}, \nabla\hat{q}^*)_{\hat{K}} = (\operatorname{div}\hat{\boldsymbol{v}}, \hat{q}^*)_{\hat{K}}\\ &\geq C\|\hat{q}^*\|^2_{0,\hat{K}}\end{aligned} \tag{5.4}$$

and

$$|\hat{\boldsymbol{v}}^*|_{1,\hat{K}} = |\mathscr{T}_k\hat{\boldsymbol{v}}|_{1,\hat{K}} \leq Ck^{\frac{d-1}{2}}\|\hat{\boldsymbol{v}}\|_{1,\hat{K}} \leq Ck^{\frac{d-1}{2}}\|\hat{q}^*\|_{0,\hat{K}}. \tag{5.5}$$

This proves the discrete inf-sup condition on the reference element.

Next, we let

$$\boldsymbol{v}^* = \boldsymbol{P}_K(\hat{\boldsymbol{v}}^*)k^{-(d-1)/2}h^d_K = |\boldsymbol{J}_K|^{-1}\boldsymbol{J}_K\hat{\boldsymbol{v}}^* \circ \boldsymbol{F}^{-1}_K k^{-(d-1)/2}h^d_K, \tag{5.6}$$

where $\boldsymbol{J}_K$ is the Jacobian matrix of $\boldsymbol{F}_K$ and $|\boldsymbol{J}_K|$ is the determinant of $\boldsymbol{J}_K$. $\boldsymbol{P}_K$ is the Piola transformation, and by assumption (A1), $\boldsymbol{P}_K\hat{\boldsymbol{v}}^* \in \boldsymbol{V}_k(K)\cap[H_0^1(K)]^d$. Using the basic property of the Piola transform (c.f. e.g. [7, pp. 97–98]), we have

$$(\operatorname{div}\boldsymbol{v}^*, q^*)_K = (\operatorname{div}\hat{\boldsymbol{v}}^*, \hat{q}^*)_{\hat{K}}\, k^{-(d-1)/2} h_K^d. \tag{5.7}$$

By the assumption of the regularity of the elements, we have

$$\|q^*\|_{0,K} \approx C h_K^{d/2}\|\hat{q}^*\|_{0,\hat{K}} \text{ and } |\boldsymbol{v}^*|_{1,K} \le C k^{-(d-1)/2} h_K^{d/2}|\hat{\boldsymbol{v}}^*|_{1,\hat{K}}. \tag{5.8}$$

The assertion then follows from (5.7) and (5.8), and the local conditions (5.4) and (5.5). □

We now use the local stability result above to prove the main theorem, using a standard argument, cf. e.g. [11].

Proof of Theorem 5.1. Let $q \in W_n$ be arbitrary and write $q = \overline{q} + q^*$ with $\overline{q}$ being the L^2 projection of q onto the space of piecewise constants:

$$\overline{W}_n = \{q \in W_n \mid q_{|K} \in P_0(K)\ \forall K \in \mathscr{C}_h\}.$$

It is well known (cf. e.g. [11]) that the pair $(\boldsymbol{V}_n, \overline{W}_n)$ is stable, provided (A4) is satisfied. Hence, there is $\overline{\boldsymbol{v}} \in \boldsymbol{V}_n$ such that

$$(\operatorname{div}\overline{\boldsymbol{v}}, \overline{q}) \ge C_3\|\overline{q}\|_0^2 \text{ and } |\overline{\boldsymbol{v}}|_1 \le C_4\|\overline{q}\|_0, \tag{5.9}$$

with the positive constants C_3 and C_4 independent of $\overline{\boldsymbol{v}}$ and $\overline{q}$.

For each $K \in \mathscr{C}_h$, we have $q^*{}_{|K} \in W_k(K)\cap L_0^2(K)$ and hence by Lemma 5.1 we can find $\boldsymbol{v}^* \in \boldsymbol{V}_n$ such that $\boldsymbol{v}^*{}_{|K} \in [H_0^1(K)]^d$ and

$$(\operatorname{div}\boldsymbol{v}^*, q^*) \ge C_1 k^{-(d-1)/2}\|q^*\|_0^2 \quad \text{with} \quad |\boldsymbol{v}^*|_1 \le C_2\|q^*\|_0. \tag{5.10}$$

Since $\boldsymbol{v}^*{}_{|K} \in [H_0^1(K)]^d$, it holds that

$$(\operatorname{div}\boldsymbol{v}^*, \overline{q}) = 0. \tag{5.11}$$

Let now $\boldsymbol{v} = \delta\overline{\boldsymbol{v}} + \boldsymbol{v}^*$. Using (5.9)-(5.11), the Schwarz inequality and the arithmetic-geometric mean inequality we get

$$\begin{aligned}
(\operatorname{div}\boldsymbol{v}, q) &= \delta(\operatorname{div}\overline{\boldsymbol{v}}, \overline{q}) + \delta(\operatorname{div}\overline{\boldsymbol{v}}, q^*) + (\operatorname{div}\boldsymbol{v}^*, \overline{q}) + (\operatorname{div}\boldsymbol{v}^*, q^*)\\
&\ge \delta C_3\|\overline{q}\|_0^2 - \delta|\overline{\boldsymbol{v}}|_1\|q^*\|_0 + C_1 k^{-(d-1)/2}\|q^*\|_0^2\\
&\ge \delta C_3\|\overline{q}\|_0^2 - (\delta/2\epsilon)\|q^*\|_0^2 - (\delta\epsilon/2)|\overline{\boldsymbol{v}}|_1^2 + C_1 k^{-(d-1)/2}\|q^*\|_0^2\\
&\ge \delta(C_3 - (C_4^2\epsilon/2))\|\overline{q}\|_0^2 + (C_1 k^{-(d-1)/2} - (\delta/2\epsilon))\|q^*\|_0^2\\
&\ge C_5 k^{-(d-1)/2}(\|\overline{q}\|_0^2 + \|q^*\|_0^2)\\
&= C_5 k^{-(d-1)/2}\|q\|_0^2,
\end{aligned}$$

where we first choose $\epsilon = C_3/C_4^2$ and then $\delta = \epsilon C_1 k^{-(d-1)/2}$. Since we then also have

$$|\boldsymbol{v}|_1 \le \delta|\overline{\boldsymbol{v}}|_1 + |\boldsymbol{v}^*|_1 \le \delta C_4\|\overline{q}\|_0 + C_2\|q^*\|_0 \le C\,\|q\|_0,$$

the assertion is proved.

Remark 5.1. It is clear that an improved version of (A3) would imply an improved stability estimate in Theorem 5.1. In [6] and Sect. 3 of [14], it is proved by means of counterexamples that the stability estimate of Theorem 5.1 is *sharp* when Method 5 is used over a single square. This implies that (A3) is the best estimate that can hold for $\mathscr{T}_k = \mathscr{S}_{k+1}$ in this case (Remark 4.1). We remark though, that numerical experiments in [14] indicate that the effect of the loss of $O(k^{\frac{d-1}{2}})$ in the stability may not be fully apparent till k is quite high ($k \geq 20$ in two dimensions).

5.2. Convergence estimates

We now use Theorem 5.1 to derive estimates that give asymptotic rates of convergence in terms of both h and k. Suppose the regularity of the solution is expressed in terms of Sobolev spaces, then both $\boldsymbol{u}$ and p converge at an optimal rate in terms of h, while the rate in terms of k is optimal up to $O(k^{\epsilon})$ for $\|\boldsymbol{u} - \boldsymbol{u}_n\|_1$ and up to $O(k^{\frac{d-1}{2}+\epsilon})$ for $\|p - p_n\|_0$. Similar optimality holds for the L^2 estimate $\|\boldsymbol{u} - \boldsymbol{u}_n\|_0$ provided the following shift theorem holds,

$$\|\boldsymbol{u}\|_2 + \|p\|_1 \leq C\|\boldsymbol{f}\|_0. \tag{5.12}$$

We will require the following interpolation result.

Lemma 5.2. *For any $r \geq 1$, let $Y^r = X^r \times Z^r$, where $X^r = [H^r(\Omega) \cap H_0^1(\Omega)]^d$ and $Z^r = H^{r-1}(\Omega) \cap L_0^2(\Omega)$. Then using the K-method of interpolation [5], for $q = r_1 + \theta(r_2 - r_1)$, $r_2 \geq r_1 \geq 1$, $0 \leq \theta \leq 1$,*

$$[Y^{r_1}, Y^{r_2}]_\theta = Y^q.$$

Proof. We note that

$$Z^r = \{p \mid p \in H^{r-1}(\Omega),\ I(p) \in \{0\}\}$$

where $I(p) = \int_\Omega p\,dx$ satisfies $I \in \mathscr{L}(H^{r-1}(\Omega), \mathbb{R})$ for $r \geq 1$. This is easy then to verify the conditions of Theorem 14.3 of [13] and thereby deduce that

$$[Z^{r_1}, Z^{r_2}]_\theta = Z^q.$$

Using the fact that a similar relation holds for the X^r spaces, we may then use a standard result on the interpolation of products of spaces (equation (6.42), Chapter 2 of [13]) to get the lemma. □

Theorem 5.2. *Let assumptions (A1) - (A6) be valid and suppose that the solution to (1.1) satisfies $(\boldsymbol{u}, p) \in [H^m(\Omega)]^d \times H^{m-1}(\Omega)$. For every $\epsilon > 0$ there is a positive constant C_ϵ such that*

$$\|\boldsymbol{u} - \boldsymbol{u}_n\|_1 + k^{-(d-1)/2}\|p - p_n\|_0 \leq C_\epsilon h^l k^{-m+1+\epsilon}(\|\boldsymbol{u}\|_m + \|p\|_{m-1}), \tag{5.13}$$

with $l = \min\{m-1, k\}$. When the shift theorem (5.12) is valid, we additionally have

$$\|\boldsymbol{u} - \boldsymbol{u}_n\|_0 \leq C_\epsilon h^{l+1} k^{-m+\epsilon}(\|\boldsymbol{u}\|_m + \|p\|_{m-1}). \tag{5.14}$$

Proof. It is well known that it is sufficient to perform the error analysis for the completely incompressible case, i.e. for the approximation of (1.5) (see [7]). This case is covered by the classical theory for the approximation of saddle point problems (cf. [7, 11]). Using this theory, the Korn inequality and the stability condition (5.1) give the two estimates

$$\begin{aligned}&\|\boldsymbol{u}-\boldsymbol{u}_n\|_1 + k^{-(d-1)/2}\|p-p_n\|_0 \\ &\leq Ck^{(d-1)/2}\{\inf_{\boldsymbol{v}\in\boldsymbol{V}_n}\|\boldsymbol{u}-\boldsymbol{v}\|_1 + \inf_{q\in W_n}\|p-q\|_0\}\end{aligned} \tag{5.15}$$

and

$$\|\boldsymbol{u}-\boldsymbol{u}_n\|_1 + k^{-(d-1)/2}\|p-p_n\|_0 \leq C\{\inf_{\boldsymbol{v}\in\mathbf{Z}_n}\|\boldsymbol{u}-\boldsymbol{v}\|_1 + \inf_{q\in W_n}\|p-q\|_0\} \tag{5.16}$$

where $\mathbf{Z}_n = \{\boldsymbol{v}\in\boldsymbol{V}_n \,|\, (\operatorname{div}\boldsymbol{v}, q) = 0 \ \forall q\in W_n\}$. From the second estimate (5.16), we directly get (by taking $\boldsymbol{v}=\mathbf{0}$ and $q=0$)

$$\|\boldsymbol{u}-\boldsymbol{u}_n\|_1 + k^{-(d-1)/2}\|p-p_n\|_0 \leq C(\|\boldsymbol{u}\|_1 + \|p\|_0). \tag{5.17}$$

On the other hand, for any $s>1$, the first estimate (5.15) gives, using (A5) and (A6) (for $\boldsymbol{u}, p$ smooth enough),

$$\begin{aligned}&\|\boldsymbol{u}-\boldsymbol{u}_n\|_1 + k^{(d-1)/2}\|p-p_n\|_0 \\ &\leq C(s)h^{\min(s-1,k)}k^{-s+1+(d-1)/2}(\|\boldsymbol{u}\|_s + \|p\|_{s-1}).\end{aligned} \tag{5.18}$$

Let $s-1 > (d-1)(m-1)/2\epsilon$. For $k \geq s-1$, we may now interpolate between (5.17) and (5.18), using Lemma 5.2 with $\theta = \frac{m-1}{s-1}$, to get (5.13) (see Theorem 4.2 of [3] for details). For $k < s-1$, we may assume $k \leq (d-1)(m-1)/2\epsilon$. Then choosing C_ϵ to be $(\frac{(d-1)(m-1)}{2\epsilon})^{(d-1)/2}$, (5.13) follows by taking $s=m$ in (5.18).

To prove the L^2-estimate for the deflection we have to slightly modify the usual duality argument. As usual, we first consider the solution $(\boldsymbol{z}, r)\in\boldsymbol{V}\times W$ to

$$\begin{aligned}2\mu(\varepsilon(\boldsymbol{z}),\varepsilon(\boldsymbol{v})) - (r,\operatorname{div}\boldsymbol{v}) &= (\boldsymbol{u}-\boldsymbol{u}_n,\boldsymbol{v}) \quad \forall\boldsymbol{v}\in\boldsymbol{V},\\ (\operatorname{div}\boldsymbol{z}, q) &= 0 \quad \forall q\in W.\end{aligned} \tag{5.19}$$

With the regularity assumption (5.12) we have

$$\|\boldsymbol{z}\|_2 + \|r\|_1 \leq C\|\boldsymbol{u}-\boldsymbol{u}_n\|_0. \tag{5.20}$$

In the usual way it now follows that

$$\begin{aligned}\|\boldsymbol{u}-\boldsymbol{u}_n\|_0^2 &= 2\mu(\varepsilon(\boldsymbol{z}-\widetilde{\boldsymbol{z}}),\varepsilon(\boldsymbol{u}-\boldsymbol{u}_n))\\ &\quad -(r-\widetilde{r},\operatorname{div}(\boldsymbol{u}-\boldsymbol{u}_n)) - (\operatorname{div}(\boldsymbol{z}-\widetilde{\boldsymbol{z}}), p-p_n)\end{aligned} \tag{5.21}$$

for any $\widetilde{\boldsymbol{z}}\in\boldsymbol{V}_n$ and $\widetilde{r}\in W_n$.

Next, we denote by $(\boldsymbol{z}_n, r_n)\in\boldsymbol{V}_n\times W_n$ the "Stokes projection" of $(\boldsymbol{z}, r)$, i.e. the solution to

$$(5.22)\qquad \begin{aligned} 2\mu(\varepsilon(\boldsymbol{z}-\boldsymbol{z}_n),\varepsilon(\boldsymbol{v}))-(r-r_n,\operatorname{div}\boldsymbol{v}) &= 0 \quad \forall \boldsymbol{v}\in \boldsymbol{V}_n,\\ (\operatorname{div}(\boldsymbol{z}-\boldsymbol{z}_n),q) &= 0 \quad \forall q\in W_n. \end{aligned}$$

From the estimate (5.13) already proven we have

$$(5.23)\qquad \|\boldsymbol{z}-\boldsymbol{z}_n\|_1 \le C_\epsilon h k^{-1+\epsilon}(\|\boldsymbol{z}\|_2+\|r\|_1).$$

We now choose

$$(5.24)\qquad \widetilde{\boldsymbol{z}} = \boldsymbol{z}_n \quad\text{and}\quad \widetilde{r} = \Pi_n r,$$

where Π_n denotes the L^2 projection onto W_n. From the second equation in (5.22) we then have

$$(5.25)\qquad (\operatorname{div}(\boldsymbol{z}-\widetilde{\boldsymbol{z}}),p_n) = (\operatorname{div}(\boldsymbol{z}-\boldsymbol{z}_n),\Pi_n p)\quad (=0),$$

and thus (5.21) and (5.23) give

$$\begin{aligned} &\|\boldsymbol{u}-\boldsymbol{u}_n\|_0^2 = 2\mu(\varepsilon(\boldsymbol{z}-\boldsymbol{z}_n),\varepsilon(\boldsymbol{u}-\boldsymbol{u}_n))\\ &-(r-\Pi_n r,\operatorname{div}(\boldsymbol{u}-\boldsymbol{u}_n))-(\operatorname{div}(\boldsymbol{z}-\boldsymbol{z}_n),p-\Pi_n p)\\ &\le C\{\|\boldsymbol{z}-\boldsymbol{z}_n\|_1\|\boldsymbol{u}-\boldsymbol{u}_n\|_1+\|r-\Pi_n r\|_0\|\boldsymbol{u}-\boldsymbol{u}_n\|_1+\|\boldsymbol{z}-\boldsymbol{z}_n\|_1\|p-\Pi_n p\|_0\}\\ &\le \left\{C_\epsilon h k^{-1+\epsilon}\|\boldsymbol{z}\|_2(\|\boldsymbol{u}-\boldsymbol{u}_n\|_1+\|p-\Pi_n p\|_0)+Chk^{-1}\|r\|_1\|\boldsymbol{u}-\boldsymbol{u}_n\|_1\right\}. \end{aligned}$$

Hence, the asserted estimate follows from (5.20), (5.13) and the interpolation estimate (cf. [3]) for $\|p-\Pi_n p\|_0$. □

Remark 5.2. For the approximation of the stress tensor $\boldsymbol{\sigma}_n = 2\mu\varepsilon(\boldsymbol{u}_n)+p_n\boldsymbol{I}$ we thus get the estimate

$$(5.26)\qquad \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_n\|_0 \le C_\epsilon h^l k^{-m+(d+1)/2+\epsilon}(\|\boldsymbol{u}\|_m+\|p\|_{m-1}).$$

Remark 5.3. For the p-version of the two dimensional problem it was shown in [2] that for $\boldsymbol{u}$ with $\operatorname{div}\boldsymbol{u}=0$ there is an approximation $\boldsymbol{u}^I\in\boldsymbol{V}_n$ such that

$$\operatorname{div}\boldsymbol{u}^I = 0 \quad\text{and}\quad \|\boldsymbol{u}-\boldsymbol{u}^I\|_1 \le Ck^{-m+1}\|\boldsymbol{u}\|_m.$$

By using this in (5.16), one gets the estimate

$$(5.27)\qquad \|\boldsymbol{u}-\boldsymbol{u}_n\|_1+k^{-1/2}\|p-p_n\|_0 \le Ck^{-m+1}(\|\boldsymbol{u}\|_m+\|p\|_{m-1}).$$

However, since we have a parallelogram mesh, the results in Sect. 5 of [2] show that with respect to the mesh size h, $\|\boldsymbol{u}-\boldsymbol{u}^I\|_1 \ge Ch^{l-1}$ for any $\boldsymbol{u}^I$ which satisfies $\operatorname{div}\boldsymbol{u}^I=0$ exactly. Based on (5.27), though, we expect that ϵ may be taken to be 0 (and $C_\epsilon = C$) in Theorem 5.2.

Remark 5.4. The factor $h^l k^{-m+1}$ in (5.13) represents the asymptotic rate of best approximation when *hp* spaces over quasiuniform meshes are used and the solution $(\boldsymbol{u}, p)$ is in $[H^m(\Omega)]^d \times H^{m-1}(\Omega)$.

Note that in general (5.15) above guarantees that the rate of convergence for the finite element solution will be no worse than the best approximation rate modulo a maximum possible loss of $O(k^{d-1})$ for p and $O(k^{\frac{d-1}{2}})$ for $\boldsymbol{u}$. It is well-known that for properly constructed *hp* spaces, the asymptotic rate of best approximation is *exponential* (see e.g. [3]). The meshes required are highly refined near the corners of the domain (non-quasiuniform). Such meshes can still be constructed over various domains of interest using only the parallelogram elements we have analyzed here (by the use of hanging nodes — see e.g. [10]). Then, since the loss due to the lack of stability is at most *algebraic*, (5.15) establishes *exponential* convergence for such *hp* methods.

References

1. Aubin, J.-P. (1979): Applied Functional Analysis. Wiley, New York
2. Babuška I., Suri, M. (1992): Locking effects in the finite element approximation of elasticity problems. Numer. Math. **62**, 439–463
3. Babuška I., Suri, M. (1994): The p and h-p versions of the finite element method: basic principles and properties. SIAM Review **36**, 578–632
4. Babuška, I, Szabó, B.A. (1982): On the rates of convergence of the finite element method. Internat. J. Numer. Math. Engng. **18**, 323–341
5. Bergh, I., Løfstrøm, J. (1976): Interpolation Spaces. Springer, Berlin Heidelberg New York
6. Bernardi, C., Maday, Y. (1992): Approximations Spectrales de Problèmes aux Limites Elliptiques. Springer
7. Brezzi, F., Fortin, M. (1991): Mixed and Hybrid Finite Element Methods. Springer
8. Brezzi, F., Fortin, M., Stenberg, R. (1991): Error analysis of mixed-interpolated elements for Reissner-Mindlin plates. Mathematical Models and Methods in Applied Sciences **1**, 125–151
9. Ciarlet, P.G. (1978): The Finite Element Method for Elliptic Problems. North Holland, Amsterdam New York Oxford
10. Demkowicz, L., Oden, T.J., Rachowicz, K., Hardy, D. (1989): Toward a universal $h-p$ adaptive element strategy. Part I: Constrained approximation and data structure. Comput. Meth. Appl. Mech. Engng. **77**, 79–113
11. Girault, V., Raviart, P.A. (1986): Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms. Springer
12. Jensen, S., Vogelius, M. (1990): Divergence stability in connection with the p-version of the finite element method. RAIRO Math. Mod. Num. Anal. **24**, 737–764
13. Lions, J.L., Magenes, E. (1972): Non-homogeneous Boundary Value Problems and Applications I. Springer, Berlin Heidelberg New York
14. Maday, Y., Patera, A.T., Rønquist, E.M.: The $[\mathbb{P}_k]^{\mathrm{N}} \times \mathbb{P}_{k-2}$ method for the approximation of the Stokes problem. Preprint. To appear in Numer. Math.
15. Scott, L.R., Vogelius, M. Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials. Math. Modeling Num. Anal. **19**, 111–143
16. Stenberg, R., Suri, M. (1994): Mixed *hp* elements for plates. In: Papadrakakis, M., Topping, B.H.V. (eds.) Advances in Finite Element Techniques, pp. 21–24. Civil-Comp Ltd., Edinburgh, Scotland
17. Stenberg, R., Suri, M.: An *hp* error analysis of MITC plate elements. Helsinki University of Technology, Laboratory for Strength of Materials. Report 23. To appear in SIAM J. Num. Anal.

18. Szabó B.A, Babuška, I. (1991): Finite Element Analysis. Wiley, New York
19. Vogelius, M. (1983): An analysis of the *p*-version of the finite element method for nearly incompressible materials. Numer. Math. **41**, 39–53

This article was processed by the author using the LaTeX style file *pljour1m* from Springer-Verlag.