ANALYSIS OF MIXED FINITE ELEMENTS FOR ELASTICITY.
I. EXACT STRESS SYMMETRY

PHILIP L. LÉDÉRER AND ROLF STENBERG

Abstract. We consider mixed finite element methods with exact symmetric stress tensors. We derive a new quasi-optimal a priori error estimate uniformly valid with respect to the compressibility. For the a posteriori error analysis we consider the Prager-Synge hypercircle principle and introduce a new estimate uniformly valid in the incompressible limit. All estimates are validated by numerical examples.

1. Introduction

The purpose of this paper is to provide an analysis of mixed finite element methods for linear elasticity. By mixed methods we mean methods based on the principle of minimum of the complementary energy, i.e. methods in which the strain energy is expressed with the stress tensor and the equation of static equilibrium act as a constraint. The Lagrange multiplier connected to this constraint is the displacement vector. By an integration by parts, this formulation is dual to the standard formulation of minimizing the total energy where Dirichlet boundary condition for the displacement now become natural and traction conditions are enforced. A crucial property of this class of methods is that each element is point wise in mechanical equilibrium and along each element edge or face the traction vector is continuous. These conditions are naturally very appealing, and that was the original motivation to develop the methods in the engineering community [15, 32]. The formulation can be traced by various energy methods classically used in elastostatics, cf. [10]. The first mathematical treatment of this principle was done by Friedrichs [16, 13].

Mathematically, the methods (and the corresponding simpler methods for scalar second order problems) have been analyzed thoroughly, and the literature is voluminous as can be seen from the monograph by Boffi, Brezzi and Fortin [11]. The analysis framework usually used is the one laid down in the pioneering work by Raviart and Thomas [27]. More recently, tools of differential exterior calculus has been used to gain insight into the methods [5], which has lead to numerous new methods. These analyses uses the $H(\text{div})$ and $L^2$ norms for the stress and the displacement, respectively. From a mechanical point of view this can be seen as unnatural since the norms do not have any physical meaning and hence blur the origin of the methods, the minimisation of the strain energy. In our analysis we derive the estimates in energy, or closely related, norms. This approach was initiated in a classical paper by Babuška, Osborn and Pitkäranta [6]. For the elasticity problem their technique of mesh dependent norms was first used by Pitkäranta and Stenberg [25]. The norms used are basically the energy norm for the stresses and a broken energy norm for the displacement. This broken $H^1$ norm was first introduced for

This work was supported by the Academy of Finland (Decision 324611).
interior penalty methods by Arnold [1]. During the last decades the interior penalty methods have received considerable attention, now under the name of discontinuous Galerkin methods (DG), cf. [14]. The equilibrium property enables the error in the stress to be decoupled from that of the lower order for the displacement. However, it also leads to a superconvergence estimate for the difference between the finite element solution and the $L^2$ projection of the displacement. Brezzi and Arnold where the first to observe that this can be used in a local postprocessing yielding a more accurate approximation [3]. In Lovadina and Stenberg [23] it was shown that this leads to a simple a posteriori error estimate.

In this paper we first collect the techniques and results mentioned above. In addition, we improve and extend the analysis. For the a priori analysis we use Gudi’s trick for DG methods [18] in order to improve the estimate removing the flaw that the exact stress should be in $H^s$, with $s > 1/2$. In addition, we use a second postprocessing, the so-called Oswald interpolation [14], to obtain a kinematically admissible displacement, i.e. continuous and satisfying the Dirichlet boundary conditions. With this we utilise the classical hypercircle principle of Prager and Synge [26, 24], which states with a kinematically admissible displacement and a statically admissible stress approximation, i.e. one with exact satisfaction of the equilibrium and traction boundary conditions, it is possible to get an a posteriori estimate in the form of an equality. This idea we use to derive an a posteriori estimate that includes oscillation terms in the case when the equilibrium and the traction boundary are satisfied only approximatively.

One attractive feature of mixed methods is that they are robust also in the incompressible limit. For the a priori estimate this is a consequence of the ellipticity in the kernel in Brezzi’s theory of saddle point problems which is valid. The Prager-Synge hypercircle estimate, however, breaks down near incompressibility. For this case we introduce a novel a posteriori estimate.

The plan of the paper is the following: in the next two sections we first recall the elasticity problem and then discuss its mixed approximation. We concretise it by the Watwood-Hartz (or Johnson-Mercier) [32, 21] method, and the families of Arnold-Douglas-Gupta [4] and Arnold-Awanou-Winther [2]. Then we derive the a priori estimates for both the stress and the post processed displacement in Section 4 and Section 5, respectively. In Section 6 we prove both the Prager-Synge and our new a posteriori estimate. In the final section numerical results are given.

We use the established notation for Sobolev spaces and finite element methods. We write $A \lesssim B$ when there exist a positive constant $C$, that is independent of the mesh parameter and in particular of the two Lamé parameters $\mu, \lambda$ (see below) such that $A \leq CB$. Analogously we define $A \gtrsim B$. This means that the dependency of the two Lamé parameters are made explicit in the norms used.

2. The Equations of Elasticity

Let $\Omega \subset \mathbb{R}^d$ be a polygonal or polyhedral domain. The physical unknowns are the displacement vector $u = (u_1, \ldots, u_d)$ and the symmetric stress tensor $\sigma = \{\sigma_{ij}\}$, $\sigma_{ij} = \sigma_{ji}$, $i, j = 1, \ldots, d$. The stress tensor is related to the strain tensor

\begin{equation}
\varepsilon(u), \quad \varepsilon(u)_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right),
\end{equation}

by a linear constitutive law. In order to be explicit, we consider the plain strain ($d = 2$) or 3D ($d = 3$) problem for an isotropic material. We define the compliance
matrix

\[(2.2)\quad C\tau = \frac{1}{2\mu} \left( \tau - \frac{\lambda}{2\mu + n\lambda} \text{tr}(\tau)I \right),\]

with the Lamé parameters \(\mu\) and \(\lambda\). Then it holds

\[(2.3)\quad C\sigma + \varepsilon(u) = 0.\]

The inverse of the compliance matrix, the elasticity matrix, we denote by \(A\), i.e.

\[(2.4)\quad A\tau = C^{-1}\tau = 2\mu\tau + \lambda\text{tr}(\tau)I.\]

In the limit \(\lambda \to \infty\) the material becomes incompressible, i.e.

\[(2.5)\quad \text{div } u = 0.\]

The loading consists of a body load \(f\) and a traction \(g\) on the boundary part \(\Gamma_N\). On the complementary part \(\Gamma_D\) homogeneous Dirichlet conditions for the displacement are given. The equations of elasticity in mixed form are then

\[(2.6)\quad C\sigma - \varepsilon(u) = 0 \quad \text{in } \Omega,
(2.7)\quad \text{div } \sigma + f = 0 \quad \text{in } \Omega,
(2.8)\quad u = 0 \quad \text{on } \Gamma_D,
(2.9)\quad \sigma n = g \quad \text{on } \Gamma_N.\]

For this system we use two variational formulations. The first is: find \(\sigma \in [L^2(\Omega)]^{d \times d}_{\text{sym}}\) and \(u \in [H^1_D(\Omega)]^d\) such that

\[(2.10)\quad B(\sigma, u; \tau, v) = (f, v) + \langle g, v \rangle_{\Gamma_N} \quad \forall (\tau, v) \in [L^2(\Omega)]^{d \times d}_{\text{sym}} \times [H^1_D(\Omega)]^d,\]

with the bilinear form

\[(2.11)\quad B(\sigma, u; \tau, v) = (C\sigma, \tau) - (\varepsilon(u), \tau) - (\varepsilon(v), \sigma).\]

Physically, the natural norms for analysing this problem are

\[(2.12)\quad \|\tau\|_C^2 = (C\sigma, \tau) \quad \text{and} \quad \|\varepsilon(u)\|^2_A = (A\varepsilon(u), \varepsilon(u)),\]

which are the double of the strain energy expressed by the stress and displacement, respectively. The Babuška–Brezzi condition is then simply the identity

\[(2.13)\quad \sup_{\tau \in [L^2(\Omega)]^{d \times d}_{\text{sym}}} \frac{(\tau, \varepsilon(v))}{\|\tau\|_C} = \|\varepsilon(v)\|_A \quad \forall v \in [H^1_D(\Omega)]^d.\]

This gives the following stability estimate (with a known constant, cf. [20])

\[(2.14)\quad \sup_{\eta \in [L^2(\Omega)]^{d \times d}_{\text{sym}}} \sup_{\varepsilon(\tau) \in [H^1_D(\Omega)]^d} \frac{B(\sigma, u; \eta, \varepsilon(\tau))}{\|\eta\|_C^2 + \|\varepsilon(\tau)\|^2_A} \geq \left( \frac{\sqrt{5} - 1}{2} \right)^{1/2} \left( \|\tau\|_C^2 + \|\varepsilon(u)\|^2_A \right) \quad \forall (\tau, v) \in [L^2(\Omega)]^{d \times d}_{\text{sym}} \times [H^1_D(\Omega)]^d.\]

In the incompressible limit \(\|\cdot\|_C\) does not, however, define a norm, and the full stability is a consequence of the ellipticity in the kernel. Instead of using the abstract theory of Brezzi [11], we give an explicit proof of stability.

More precisely, for \(\tau\) we let \(\tau^D\) be the deviatoric part of \(\tau\), defined by the condition \(\text{tr}(\tau^D) = 0\). Hence it holds

\[(2.15)\quad \tau = \tau^D + \frac{1}{d} \text{tr}(\tau)I.\]

By a direct computation we get.
Lemma 1. It holds that
\begin{equation}
(C\tau, \tau) = \frac{1}{2\mu} \|\tau^D\|_0^2 + \frac{1}{2\mu + d\lambda} \|\text{tr}^2\|_0^2.
\end{equation}

From this it is seen that when \(\lambda \to \infty\), \(\|\cdot\|_C\) does not give control over the pressure part (the trace) of the stress tensor.

To derive estimates valid independently of \(\lambda\) we use the norms
\begin{equation}
\mu^{-1/2} \|\tau\|_0 \quad \text{for } \tau \in [L^2(\Omega)]^{d \times d}_{\text{sym}}, \quad \text{and } \mu^{1/2} \|\varepsilon(v)\|_0 \quad \text{for } v \in [H^1_D(\Omega)]^d.
\end{equation}

For the stability estimate the Babuška–Brezzi condition for the Stokes problem \cite{17}
is needed. Using this we prove the following stability estimate.

Theorem 1. It holds that
\begin{equation}
\sup_{v \in [H^1_D(\Omega)]^d} \frac{B(\tau, v; \eta, z)}{\left(\mu^{-1}\|\eta\|_0^2 + \mu\|\varepsilon(z)\|_0^2\right)^{1/2}} \geq \left(\mu^{-1}\|\tau\|_0^2 + \mu\|\varepsilon(v)\|_0^2\right)^{1/2}
\end{equation}
\[\forall (\tau, v) \in \left[L^2(\Omega)\right]^{d \times d}_{\text{sym}} \times \left[H^1_D(\Omega)\right]^d.\]

Proof. Let \((\tau, v)\) be given. By (2.18) there exists \(z \in [H^1_D(\Omega)]^d\) such that
\begin{equation}
\|\text{tr}(\tau)\|_0 = \|\tau\|_0 = \|\varepsilon(v)\|_0 = \|\varepsilon(z)\|_0.
\end{equation}

Let \(\delta > 0\) and \(\gamma > 0\). By the bilinearity we have
\begin{align}
B(\tau, v; \tau - \gamma \varepsilon(v), -v - \delta z) \\
= B(\tau, \varepsilon(v); 0, 0) - \beta B(\tau, v; \varepsilon(v), 0) - \delta B(\tau, v; 0, z).
\end{align}

For the first term we have
\begin{equation}
B(\tau, v; \tau, -v) = (C\tau, \tau) = \frac{1}{2\mu} \|\tau^D\|_0^2 + \frac{1}{2\mu + d\lambda} \|\text{tr}^2\|_0^2.
\end{equation}

From (2.2) we have
\begin{equation}
\|C\tau\|_0 \leq \frac{1}{2\mu} \left(\|\tau\|_0 + \frac{1}{d} \|\text{tr}(\tau)I\|_0\right) \leq \frac{1}{\mu} \|\tau\|_0.
\end{equation}

Using the Schwarz and Young inequalities then gives
\begin{equation}
-\gamma B(\tau, v; \varepsilon(v), 0) = -\gamma (C\tau, \varepsilon(v)) + \gamma \|\varepsilon(v)\|_0^2 \geq -\frac{\gamma}{2} \|C\tau\|_0^2 + \frac{\gamma}{2} \|\varepsilon(v)\|_0^2
\end{equation}
\[\geq -\frac{\gamma}{2\mu^2} \|\tau\|_0^2 + \frac{\gamma}{2} \|\varepsilon(v)\|_0^2.
\]

By (2.20) we have
\begin{equation}
-\delta B(\tau, v; 0, z) = \delta (\tau, \varepsilon(z)) = \frac{\delta}{d} (\text{tr}(\tau), \varepsilon(z)) + \delta (\tau^D, \varepsilon(z))
\end{equation}
\[= \frac{\delta}{d} \|\text{tr}(\tau)\|_0^2 + \delta (\tau^D, \varepsilon(z)) \geq \frac{\delta}{d} \|\text{tr}(\tau)\|_0^2 - \delta \|\tau^D\|_0 \|\varepsilon(z)\|_0
\end{equation}
\[= \frac{\delta}{d} \|\text{tr}(\tau)\|_0^2 - \delta \|\tau^D\|_0 \|\text{tr}(\tau)\|_0
\]
\[\geq \frac{\delta}{2d} \|\text{tr}(\tau)\|_0^2 - \frac{\delta}{2\beta} \|\tau^D\|_0^2.
\]
Collecting the above estimates, we get
\begin{equation}
\mathcal{B}(\tau, v; \tau - \gamma \varepsilon(v), -v - \delta z) \\
\geq \left( \frac{1}{2\mu} - \frac{\delta d}{2\beta} \right) \| \tau^D \|^2_0 + \left( \frac{1}{2\mu + d\lambda} + \frac{\delta^2}{4d^2\mu} \right) \| \text{tr}(\tau) \|^2_0 \\
- \frac{\gamma}{2\mu^2} \| \tau \|^2_0 + \frac{\gamma}{2} \| \varepsilon(v) \|^2_0.
\end{equation}
(2.26)

Now we choose \( \delta = \frac{\beta}{2\mu d} \), which gives
\begin{equation}
\mathcal{B}(\tau, v; \tau - \gamma \varepsilon(v), -v - \delta z) \\
\geq \frac{1}{4\mu} \| \tau^D \|^2_0 + \left( \frac{1}{2\mu + d\lambda} + \frac{\beta^2}{4d^2\mu} \right) \| \text{tr}(\tau) \|^2_0 - \frac{\gamma}{2\mu^2} \| \tau \|^2_0 + \frac{\gamma}{2} \| \varepsilon(v) \|^2_0.
\end{equation}
(2.27)

Let \( C > 0 \) such that
\begin{equation}
\frac{1}{4\mu} \| \tau^D \|^2_0 + \left( \frac{1}{2\mu + d\lambda} + \frac{\beta^2}{4d^2\mu} \right) \| \text{tr}(\tau) \|^2_0 \geq \frac{C}{\mu} \| \tau \|^2_0.
\end{equation}
(2.28)

and choose \( \gamma = C\mu \). This gives
\begin{equation}
\mathcal{B}(\tau, v; \tau - \gamma \varepsilon(v), -v - \delta z) \geq \frac{C}{2} \left( \mu^{-1} \| \tau \|^2_0 + \mu \| \varepsilon(v) \|^2_0 \right).
\end{equation}
(2.29)

It also holds
\begin{equation}
\mu^{-1} \| \tau - \gamma \varepsilon(v) \|^2_0 + \mu \| \varepsilon(v + \delta z) \|^2_0 \leq \mu^{-1} \| \tau \|^2_0 + \mu \| \varepsilon(v) \|^2_0,
\end{equation}
(2.30)

which proves the asserted estimate. \( \square \)

The second variational form is the basis for the mixed finite element method. By dualisation the stress is in
\begin{equation}
H(\text{div} : \Omega) = \{ \tau \in [L^2(\Omega)]^{d \times d} | \text{div} \tau \in [L^2(\Omega)]^d \},
\end{equation}
(2.31)

the displacement in \([L^2(\Omega)]^d\), and the bilinear form used is
\begin{equation}
\mathcal{M}(\sigma, u; \tau, v) = (C\sigma, \tau) + (u, \text{div} \tau) + (v, \text{div} \sigma),
\end{equation}
(2.32)

and the formulation is: find \( \sigma \in H_0(\text{div} : \Omega) \) and \( u \in [L^2(\Omega)]^d \), such that
\begin{equation}
\mathcal{M}(\sigma, u; \tau, v) + (f, v) = 0 \quad \forall (\tau, v) \in H_0(\text{div} : \Omega) \times [L^2(\Omega)]^d,
\end{equation}
(2.33)

with
\begin{equation}
H_0(\text{div} : \Omega) = \{ \tau \in H(\text{div} : \Omega) | \tau n|_{\Gamma_N} = g \},
\end{equation}
(2.34)

and
\begin{equation}
H_0(\text{div} : \Omega) = \{ \tau \in H(\text{div} : \Omega) | \tau n|_{\Gamma_N} = 0 \}.
\end{equation}
(2.35)

3. Exactly symmetric mixed finite element methods

The mixed finite element method is based on the variational formulation (2.33) in piecewise polynomial subspaces \( V_h \subset [L^2(\Omega)]^d \) and \( S_h \subset H(\text{div} : \Omega) \). We give a unified presentation that covers the following methods:

- The linear triangular method of Johnson–Mercier (JM) [21, 32].
- The triangular family of Arnold–Douglas–Gupta [4].
- The triangular family of Guzman–Neilan [19].
- The tetrahedral family of Arnold–Awanou–Winther [2].
The triangular or tetrahedral mesh is denoted by \( C_h \). The families are indexed by the polynomial degree \( k \geq 2 \) and the displacement space is simply
\[
V_h = \{ v \in [L^2(\Omega)]^d \mid v|_K \in [P_{k-1}(K)]^d \ \forall K \in C_h \}.
\]
In the JM methods it is this with \( k = 2 \), i.e. discontinuous piecewise linear polynomials. The spaces for the stress we defined by
\[
S_h = \{ \tau \in H(\text{div};\Omega) \mid \tau|_K \in S(K) \ \forall K \in C_h \}.
\]
The local space \( S(K) \) are rather involved and here we will not give the explicitly definition. The essential properties are, however, the right approximation order which is ensured by the inclusion
\[
[P_k(K)]_{\text{sym}}^{n \times n} \subset S(K),
\]
and the degrees of freedom needed for the stability; the local degrees of freedom of \( \tau \in S(K) \) contain the moments
\[
\int_K \tau : \varepsilon(v) \ \forall v \in [P_{k-1}(K)]^d,
\]
and
\[
\int_E \tau n \cdot v \ \forall v \in [P_k(E)]^d,
\]
for each edge or face \( E \) of \( K \). For the JM methods (3.3), (3.4) and (3.5) are valid with \( k = 1 \).

For all spaces, except JM, we have
\[
\text{div} \tau \in V_h \ \forall \tau \in S_h,
\]
and hence for these it holds
\[
(\text{div} \tau, v - P_h v) = 0 \ \forall \tau \in S_h,
\]
were \( P_h : [L^2(\Omega)]^d \to V_h \) denotes the \( L_2 \)-projection. For JM (3.6) does not hold, but the divergence is six dimensional, i.e. that of the deflection space. Hence, also for this method there exist a projection with the property (3.7).

For an edge/face \( E \subset \Gamma_N \) we let \( Q_E : [L^2(E)]^d \to [P_k(E)]^d \) denote the \( L^2 \) projection and define \( Q_h|_{\Gamma} = Q_E \).

The trial and test finite element spaces are then defined as
\[
S_h^0 = \{ \tau \in S_h \mid \tau n = Q_h g \ \text{on} \ \Gamma_N \},
\]
and
\[
S_h^0 = \{ \tau \in S_h \mid \tau n = 0 \ \text{on} \ \Gamma_N \},
\]
The mixed finite element method is: find \( (\sigma_h, u_h) \in S_h^0 \times V_h \) such that
\[
\mathcal{M}(\sigma_h, u_h; \tau, v) + (f, v) = 0 \ \forall (\tau, v) \in S_h^0 \times V_h.
\]

4. Stability and a Priori Error Analysis

In this section we will derive a priori error estimates. For the displacement we use the following broken energy norm, where \( \Gamma_h \) denotes the edges/faces in the interior of \( \Omega \).
\[
\|v\|_h^2 = \sum_{K \in C_h} \|\varepsilon(v)\|_{0,K}^2 + \sum_{E \in \Gamma_h} h^{-1}_E \|v\|_{0,E}^2 + \sum_{E \in \Gamma_D} h^{-1}_E \|v\|_{0,E}^2 \quad v \in V_h.
\]
The stability of the method is proven by the following two conditions.
Lemma 2. It holds that
\begin{equation}
\sup_{\tau \in S_0^h} \frac{\langle \text{div} \, \tau, v \rangle}{\|\tau\|_0} \gtrsim \|v\|_h \quad \forall v \in V_h.
\end{equation}

Proof. Let \( v \in [H^1_D(\Omega)]^d \) be given. \( \tau \in S_0^h \) we choose such that all degrees of freedom vanish, except (3.4) and (3.5) which are chosen such that
\begin{equation}
\int_K \tau : \varepsilon(v) = \int_K |\varepsilon(v)|^2 \quad \forall K \in C_h,
\end{equation}
\begin{equation}
\int_E \tau \cdot v = h_E^{-1} \int_E |[v]|^2 \quad \forall E \in \Gamma_h,
\end{equation}
and
\begin{equation}
\int_E \tau \cdot v = h_E^{-1} \int_E |v|^2 \quad \forall E \subset \Gamma_D.
\end{equation}
Hence
\begin{equation}
\langle \text{div} \, \tau, v \rangle = \|v\|_h^2.
\end{equation}
By scaling it holds
\begin{equation}
\|\tau\|_0 \lesssim \|v\|_h,
\end{equation}
which proves the claim. □

Lemma 3. It holds that
\begin{equation}
\sup_{v \in V_h} \frac{\langle v, \text{div} \, \tau \rangle}{\|v\|_h} \geq C_1 \|\text{tr}(\tau)\|_0 - C_2 \|\tau^D\|_0 \quad \forall \tau \in S_0^h.
\end{equation}

Proof. Given \( \tau \in S_0^h \), (2.18) implies that there exists \( v \in [H^1_D(\Omega)]^d \) such that
\begin{equation}
\langle v, \text{div} \, \tau \rangle = -\beta \|\text{tr}(\tau)\|_0^2 \text{ and } \|\varepsilon(v)\|_0 = \|\text{tr}(\tau)\|_0.
\end{equation}
Let \( P_h v \in V_h \) be the projection in (3.7). It holds
\begin{equation}
\langle P_h v, \text{div} \, \tau \rangle = \langle v, \text{div} \, \tau \rangle = -\langle \varepsilon(v), \tau \rangle
\end{equation}
\begin{equation}
\geq -\langle \varepsilon(v), \tau \rangle - \langle \varepsilon(v), \tau^D \rangle \geq \beta \|\text{tr}(\tau)\|_0^2 - \|\varepsilon(v)\|_0 \|\tau^D\|_0
\end{equation}
\begin{equation}
= \|\varepsilon(v)\|_0 \left( \beta \|\text{tr}(\tau)\|_0 - \|\tau^D\|_0 \right).
\end{equation}
By scaling we have
\begin{equation}
\|P_h v\|_h \lesssim \|\varepsilon(v)\|_0.
\end{equation}
Combining the two estimates above proves the claim. □

In analogy with the proof of Theorem 1 we then obtain the stability of the mixed method.

Theorem 2. It holds that
\begin{equation}
\sup_{(\eta, z) \in S_0^h \times V_h} \frac{\mathcal{M}(\tau, v; \eta, z)}{(\mu^{-1} \|\eta\|_0^2 + \mu \|\varepsilon(v)\|_0^2)^{1/2}} \geq \left( \mu^{-1} \|\tau\|_0^2 + \mu \|v\|_h^2 \right)^{1/2} \quad \forall (\tau, v) \in S_0^h \times V_h.
\end{equation}
We then get the a priori estimate. Here \( f_h \) is any piecewise polynomial approximation of \( f \).
Theorem 3. It holds that
\[
\mu^{-1/2}\|\sigma - \sigma_h\|_0 + \mu^{1/2}\|P_h u - u_h\|_h \\
\lesssim \mu^{-1/2} (\inf_{\tau \in S^0_h} \|\sigma - \tau\|_0 + (\sum_{K \in \mathcal{C}_h} h_K^{-2}\|f - f_h\|_{0,K})^{1/2}).
\]  

Proof. By the stability there exist \((\eta, v) \in S^0_h \times V_h\) with
\[
\mu^{-1/2}\|\sigma_h - \tau\|_0 + \mu^{1/2}\|P_h u - u_h\|_h \lesssim \mathcal{M}(\sigma_h - \tau, u_h, \eta, v).
\]
By the consistency we have
\[
\mathcal{M}(\sigma - \tau, u - P_h u; \eta, v) = \mathcal{M}(\sigma - \tau, u - P_h u; \eta, v).
\]
Writing out we have
\[
\mathcal{M}(\sigma - \tau, u - P_h u; \eta, v) = (\mathcal{C}(\sigma - \tau), \eta) + (u - P_h u, \text{div } \eta) + (\text{div } (\sigma - \tau), v).
\]
From (2.16) and (4.14) it follows
\[
(\mathcal{C}(\sigma - \tau), \eta) \lesssim \mu^{-1/2}\|\sigma - \tau\|_0 \mu^{-1/2}\|\eta\|_0 \lesssim \mu^{-1/2}\|\sigma - \tau\|_0.
\]
By the property (3.7) the second term vanish
\[
(u - P_h u, \text{div } \eta) = 0.
\]
Let \(I_h^n v\) be the so called Oswald interpolant to \(v\), for wich it holds [14]
\[
\|\nabla I_h^n v\|_0 + \left(\sum_{K \in \mathcal{C}_h} h_K^{-2}\|v - I_h^n v\|_{0,K}^2\right)^{1/2} \lesssim \|v\|_h.
\]
Using this we obtain
\[
(\text{div } (\sigma - \tau), v) = (\text{div } (\sigma - \tau), v - I_h^n v) + (\text{div } (\sigma - \tau), I_h^n v).
\]
The first term above we treat as follows. First, (4.20) and (4.14) yield
\[
(\text{div } (\sigma - \tau), v - I_h^n v) = \sum_{K \in \mathcal{C}_h} (\text{div } (\sigma - \tau), v - I_h^n v)_K
\]
\[
\leq \sum_{K \in \mathcal{C}_h} \|\text{div } (\sigma - \tau)\|_{0,K} \|v - I_h^n v\|_{0,K}
\]
\[
\leq \left(\mu^{-1} \sum_{K \in \mathcal{C}_h} h_K^2\|\text{div } (\sigma - \tau)\|_{0,K}^2\right)^{1/2} \left(\mu \sum_{K \in \mathcal{C}_h} h_K^{-2}\|v - I_h^n v\|_{0,K}^2\right)^{1/2}
\]
\[
\lesssim \left(\mu^{-1} \sum_{K \in \mathcal{C}_h} h_K^2\|\text{div } (\sigma - \tau)\|_{0,K}^2\right)^{1/2}.
\]
By a posteriori error analysis techniques [31] we have
\[
h_K\|\text{div } (\sigma - \tau)\|_{0,K} \lesssim (\|\sigma - \tau\|_{0,K} + h_K\|f - f_h\|_{0,K}),
\]
and hence
\[
(\text{div } (\sigma - \tau), v - I_h^n v) \lesssim \mu^{-1/2}(\|\sigma - \tau\|_0 + (\sum_{K \in \mathcal{C}_h} h_K^2\|f - f_h\|_{0,K}^2)^{1/2}),
\]
Finally, an integration by parts, and (4.20) and (4.14), yield
\[
(\text{div} (\sigma - \tau), P_h^* v) = - (\sigma - \tau, \nabla P_h^* v) \leq \|\sigma - \tau\|_0 \|\nabla P_h^* v\|_0
\]
(4.25)
\[
\lesssim \|\sigma - \tau\|_0 \|v\|_h \lesssim \mu^{1/2} \|\sigma - \tau\|_0.
\]
Collecting the estimates proves the claim. □

Above result shows that for an exact, sufficiently smooth, solution we get (by standard arguments) the expected convergence result
\[
\|\sigma - \sigma_h\|_0 = O(h^{k+1}).
\]
(4.26)
The estimate for \(\|P_h u - u_h\|_h\) is a superconvergence result, i.e. we have (again for a sufficiently smooth exact solution)
\[
\|P_h u - u_h\|_h = O(h^{k+1}) \quad \text{wheras} \quad \|u - u_h\|_h = O(h^{k-1}),
\]
(4.27)
except for JM for which we have \(\|u - u_h\|_h = O(h)\). This property enables the postprocessing of the solution in the next section.

5. POSTPROCESSING OF THE DISPLACEMENT

In the following we provide a simple technique to derive a post processed continuous displacement field with enhanced convergence properties. To this end let \(k\) be the approximation order of the stress space, see (3.3), then we define the following two spaces
\[
V_h^* = \{ v \in [L^2(\Omega)]^d \mid v|_K \in [P_{k+1}(K)]^d \quad \forall K \in \mathcal{C}_h \},
\]
(5.1)
\[
V_h^0 = V_h^* \cap [H^1_0(\Omega)]^d.
\]
(5.2)
Further let \(P_h^* : L^2(\Omega) \rightarrow V_h^*\) denote the \(L^2\) projection on \(V_h^*\).

Postprocessing. Step I: The first step is used to derive a discontinuous displacement with an enhanced accuracy. To this end we state the problem: find \(u_h^* \in V_h^*\) such that
\[
P_h u_h^* = u_h
\]
(5.3)
\[
(\varepsilon(u_h^*), \varepsilon(v))_K = (C\sigma_h, \varepsilon(v))_K \quad \forall v \in (I - P_h)V_h^*|_K.
\]

Lemma 4. It holds that
\[
\|u - u_h^*\|_h \lesssim \|u - P_h^* u\|_h + \mu^{-1} \left( \|\sigma - \sigma_h\|_0 + \left( \sum_{K \in \mathcal{C}_h} h_K^2 \|f - f_k\|_{0,K}^2 \right)^{1/2} \right).
\]
(5.4)
Proof. By the triangle inequality we have
\[
\|u - u_h^*\|_h \leq \|P_h^* u - u\|_h + \|P_h^* u - u_h^*\|_h.
\]
(5.5)
Next, we write
\[
P_h^* u - u_h^* = P_h^* (u - u_h^*) = (P_h^* - P_h) (u - u_h^*) + P_h (u - u_h^*)
\]
(5.6)
\[
= (P_h^* - P_h) (u - u_h^*) + (P_h u - u_h).
\]
For convenience we denote
\[
v = (P_h^* - P_h) (u - u_h^*) \in (I - P_h)V_h^*|_K.
\]
(5.7)
From (5.3) we get
\[
\| \varepsilon ((P_h^* - P_h)(u - u_h^*)) \|_{0,K}^2 = (\varepsilon((P_h^* - P_h)(u - u_h^*)), \varepsilon(v))_{0,K} = (\varepsilon(P_h u) - \varepsilon(u_h^*), \varepsilon(v))_{0,K}.
\]
(5.8)

Combining, we get
\[
\| (P_h^* - P_h)(u - u_h^*) \|_{0,K} \lesssim (\| u - P_h^* u \|_{0,K} + \mu^{-1}\| \sigma - \sigma_h \|_{0,K}).
\]
(5.10)

Combining the estimates gives the claim. □

Postpostprocessing. Step II: The second step is used to derive the final continuous displacement approximation (used for the hypercircle technique below) by applying an averaging operator \( I_h^* : V_h^* \rightarrow V_h^a \). Now let \( u_h^a = I_h^* u_h^* \), then we have the following error estimate.

**Theorem 4.** It holds that
\[
\| u_h^a - u_h^* \|_h \lesssim \left( \sum_{K \in T_h} \left( \sum_{E \in \Gamma_h} h_E^{-1} \| u_h^* \|_{0,E}^2 \right)^{1/2} \right)^{1/2}.
\]
(5.13)

\[
\| u_h^a - u_h^* \|_h \leq \| u - u_h^* \|_h.
\]

By the triangle inequality we have
\[
\| \varepsilon(u - u_h^2) \|_0 \leq \| u - u_h^* \|_h + \| u_h^* - u_h^2 \|_h
\]
(5.14)

and the claim follows from Lemma 4. □

For a sufficiently smooth solution we now have
\[
\| \varepsilon(u - u_h^2) \|_0 = O(h^{k+1}),
\]
(5.15)

which should be compared to (4.27).
6. A posteriori error estimates by the Hypercircle theorem

First we recall the Hypercircle method [26] and include its proof.

**Theorem 5.** (The Prager-Synge hypercircle theorem) Suppose that:
- The stress $\Sigma \in H(\text{div} : \Omega)$ is statically admissible; $\text{div} \Sigma + f = 0$ in $\Omega$, and $\Sigma n = g$ on $\Gamma_N$.
- The displacement $U \in [H^1(\Omega)]^d$ is kinematically admissible; $U|_{\Gamma_D} = 0$.

Then it holds

\[
\|\sigma - \Sigma\|^2 + \|\sigma - A\varepsilon(U)\|^2 = \|\Sigma - A\varepsilon(U)\|^2
\]

and

\[
\|\sigma - \frac{1}{2}(\Sigma + A\varepsilon(U))\|_C = \frac{1}{2}\|\Sigma - A\varepsilon(U)\|_C.
\]

**Proof.** We have

\[
\|\Sigma - A\varepsilon(U)\|^2 = \|\Sigma - \sigma + \sigma - A\varepsilon(U)\|^2
\]

\[
= \|\Sigma - \sigma\|^2 + \|\sigma - A\varepsilon(U)\|^2 + (C(\Sigma - \sigma), \sigma - A\varepsilon(U)).
\]

Next, the symmetry of $C$, $A = C^{-1}$, and $\varepsilon(u) = C\sigma$, give

\[
(C(\Sigma - \sigma), \sigma - A\varepsilon(U)) = (\Sigma - \sigma, C(\sigma - A\varepsilon(U)))
\]

\[
= (\Sigma - \sigma, \varepsilon(u) - \varepsilon(U)).
\]

An integration by parts yields

\[
(\Sigma - \sigma, \varepsilon(u) - \varepsilon(U)) = -(\text{div} (\Sigma - \sigma), u - U) + ((\Sigma - \sigma)n, u - U)_{\Gamma_N}
\]

\[
= - (\text{div} \Sigma + f, u - U) + ((\Sigma n - g, u - U)_{\Gamma_N} = 0,
\]

which proves the first identity.

Next, the orthogonality (6.5) also yields

\[
\|\sigma - \frac{1}{2}(\Sigma + A\varepsilon(U))\|^2
\]

\[
= \|A\varepsilon(u) - \frac{1}{2}(\Sigma + A\varepsilon(U))\|^2
\]

\[
= \frac{1}{4}\|2(A\varepsilon(u) - A\varepsilon(U)) + (A\varepsilon(U)) - \Sigma\|^2
\]

\[
= \|A\varepsilon(u) - A\varepsilon(U)\|^2 + (C(A\varepsilon(u) - A\varepsilon(U)), A\varepsilon(U) - \Sigma)
\]

\[+
\frac{1}{4}\|A\varepsilon(U) - \Sigma\|^2
\]

\[
= (C(A\varepsilon(u) - A\varepsilon(U)), A\varepsilon(u) - \Sigma) + \frac{1}{4}\|A\varepsilon(U) - \Sigma\|^2
\]

\[
= (\varepsilon(u) - \varepsilon(U), A\varepsilon(u) - \Sigma) + \frac{1}{4}\|A\varepsilon(U) - \Sigma\|^2
\]

\[
= \frac{1}{4}\|A\varepsilon(U) - \Sigma\|^2.
\]

\[\square\]

Let $P_K = P_h|_K$, where $P_h$ is the projection (3.7).

The a posteriori estimate is now.
Theorem 6. It holds that
\begin{equation}
\|\sigma - \sigma_h\|_C^2 + \|\sigma - \mathcal{A}\varepsilon(u_h^\sigma)\|_C^2 \leq \|\sigma_h - \mathcal{A}\varepsilon(u_h^\sigma)\|_C^2 + \text{osc}(f)^2 + \text{osc}(g)^2.
\end{equation}
and
\begin{equation}
\|\sigma - \frac{1}{2}(\sigma_h + \mathcal{A}\varepsilon(u_h^\sigma))\|_C \leq \frac{1}{2}\|\sigma_h - \mathcal{A}\varepsilon(u_h^\sigma)\|_C + \text{osc}(f) + \text{osc}(g)
\end{equation}
with
\begin{equation}
\text{osc}(f) = C\left(\sum_{K \in \mathcal{C}_h} h_K^2 \|f - P_K f\|_{0,K}^2\right)^{1/2}
\end{equation}
and
\begin{equation}
\text{osc}(g) = C\left(\sum_{E \in \Gamma_N} h_E \|g - Q_E g\|_{0,E}^2\right)^{1/2}.
\end{equation}

Proof. Let \((\bar{\sigma}, \bar{u}) \in [L^2(\Omega)]^{d \times d}_{\text{sym}} \times [H^1_0(\Omega)]^d\) be the solution to
\begin{equation}
\mathcal{B}(\bar{\sigma}, \bar{u}; \tau, v) = (P_h f, v) + \langle Q_h g, v \rangle_{\Gamma_N} \quad \forall (\tau, v) \in [L^2(\Omega)]^{d \times d}_{\text{sym}} \times [H^1_0(\Omega)]^d.
\end{equation}
Now \((\sigma_h, u_h)\) are admissible approximations to \((\bar{\sigma}, \bar{u})\) and the preceding theorem yields
\begin{equation}
s\|\bar{\sigma} - \sigma_h\|_C^2 + \|\bar{\sigma} - \mathcal{A}\varepsilon(u_h^\sigma)\|_C^2 \leq \|\sigma_h - \mathcal{A}\varepsilon(u_h^\sigma)\|_C^2
\end{equation}
and
\begin{equation}
\|\bar{\sigma} - \frac{1}{2}(\sigma_h + \mathcal{A}\varepsilon(u_h^\sigma))\|_C \leq \frac{1}{2}\|\sigma_h - \mathcal{A}\varepsilon(u_h^\sigma)\|_C.
\end{equation}
For the difference \((\sigma - \bar{\sigma}, u - \bar{u})\) we get
\begin{equation}
\mathcal{B}(\sigma - \bar{\sigma}, u - \bar{u}; \tau, v) = (f - P_h f, v) + \langle g - Q_h g, v \rangle_{\Gamma_N}.
\end{equation}
The stability then yield
\begin{equation}
\|\sigma - \bar{\sigma}\|_C + \|\varepsilon(u - \bar{u})\|_A \lesssim \sup_{\|\varepsilon(v)\|_A = 1} \left(\langle f - P_h f, v \rangle + \langle g - Q_h g, v \rangle_{\Gamma_N}\right).
\end{equation}
By the properties of the projection operators and Korn’s inequality we have
\begin{equation}
(f - P_h f, v) = \sum_{K \in \mathcal{C}_h} (f - P_K f, v - P_K v)_K \leq \sum_{K \in \mathcal{C}_h} \|f - P_K f\|_{0,K} \|v - P_K v\|_{0,K} \lesssim \sum_{K \in \mathcal{C}_h} h_K \|f - P_K f\|_{0,K} \|\nabla v\|_{0,K} \lesssim \left(\sum_{K \in \mathcal{C}_h} h_K^2 \|f - P_K f\|_{0,K}^2\right)^{1/2} \|\nabla v\|_{0,K} \lesssim \left(\sum_{K \in \mathcal{C}_h} h_K^2 \|f - P_K f\|_{0,K}^2\right)^{1/2} \|\varepsilon(v)\|_A.
\end{equation}
Using the trace theorem, a similar estimate also gives
\begin{equation}
\langle (g - Q_h g, v)_{\Gamma_N} \rangle \lesssim \left(\sum_{E \in \Gamma_N} h_E \|g - Q_E g\|_{0,E}^2\right)^{1/2} \|\varepsilon(v)\|_A.
\end{equation}
The assertion then follows by combining the above estimates. □
Remark 1. For \( f \) and \( g \) smooth, it holds

\[
\text{osc}(f) = \mathcal{O}(h^{k+1}) \quad \text{and} \quad \text{osc}(g) = \mathcal{O}(h^{k+3/2})
\]

and only \( \text{osc}(g) \) is a higher order term. However, in real problems the loading \( f \) is a constant and \( \text{osc}(f) \) vanish.

Note also that when the oscillation terms vanish, the estimates become equalities.

7. An a posteriori estimator uniformly valid in the incompressible limit

The drawback of the estimate by the hypercircle argument is that it is formulated in terms of \( \| \cdot \|_C \) which, unfortunately, ceases to be a norm in the incompressible limit \( \lambda \to \infty \) and that the stress computed from the displacement, i.e.

\[
\mathcal{A}(u_h^a) = 2\mu \varepsilon(u_h^a) + \lambda \text{div } u_h^a I,
\]
grows without limit unless \( \text{div } u_h^a \) will vanish identically. It is well known [7, 29] that this requires piecewise polynomials of degree four or higher.

In this section we will therefore derive the following a posteriori estimate uniformly valid with respect to the second Lamé parameter.

Theorem 7. It holds

\[
\mu^{-1/2} \| \sigma - \sigma_h \|_0 + \mu^{1/2} \| \varepsilon(u - u_h^a) \|_0 \\
\lesssim \mu^{1/2} \| \mathcal{C} \sigma_h - \varepsilon(u_h^a) \|_0 + \text{osc}(f) + \text{osc}(g).
\]

Proof. By Theorem 1 there exists \( (\tau, v) \in [L^2(\Omega)]^{d \times d} \times [H^1(\Omega)]^d \) with \( \mu^{-1/2} \| \tau \|_0 + \mu^{1/2} \| \varepsilon(v) \|_0 = 1 \), such that

\[
\min \left( \| \sigma - \sigma_h \|_0 + \| \varepsilon(u - u_h^a) \|_0 \right) < B(\sigma - \sigma_h, u - u_h^a; \tau, v)
\]

\[
= (\mathcal{C}(\sigma - \sigma_h), \tau) - (\varepsilon(u - u_h^a), \tau) - (\varepsilon(v), \sigma - \sigma_h).
\]

Since \( \mathcal{C} \sigma - \varepsilon(u) = 0 \) we have

\[
(\mathcal{C}(\sigma - \sigma_h), \tau) - (\varepsilon(u - u_h^a), \tau) = (\mathcal{C} \sigma_h - \varepsilon(u_h^a), \tau) \leq \| \mathcal{C} \sigma_h - \varepsilon(u_h^a) \|_0 \| \tau \|_0.
\]

Since \( \text{div } \sigma_h = P_h f \) and \( \sigma_h n = Q_h g \) we get

\[
-\varepsilon(v), \sigma - \sigma_h = (\text{div } (\sigma - \sigma_h), v) + ((\sigma - \sigma_h) n, v)_{\Gamma_N}
\]

\[
= (P_h f - f, v) + (g - Q_h g, v)_{\Gamma_N},
\]

and (6.16), (6.17) give

\[
-\varepsilon(v), \sigma - \sigma_h \lesssim \text{osc}(f) + \text{osc}(g).
\]

The assertion follows by collecting the above estimates. \( \square \)

8. Numerical examples

In this section we validate our theoretical findings with various numerical examples. All numerical examples were implemented in the finite element library Netgen/NGSolve, see [28]. For simplicity we only consider the two dimensional case and we use the following two methods:

- The Johnson–Mercier method (JM) from [21] considers linear displacements and linear stresses.
- The Arnold–Douglas–Gupta method (ADG) from [4] where we use the choice of linear displacements and quadratic stresses.
For both methods we hence have
\begin{equation}
V_h = \{ v \in [L^2(\Omega)]^d \mid v|_K \in [P_1(K)]^d \quad \forall K \in \mathcal{C}_h \}.
\end{equation}
As mentioned in Section 3 we need to specify the local stress space $S(K)$ which then defines the global stress space by (3.2). Both, the JM and ADG method use a similar construction. Each triangle $K \in \mathcal{C}_h$ is divided into three sub triangles $K_i$ with $i = 1, 2, 3$, by connecting the barycentre with the three vertices. $S(K)$ is given by
\begin{equation}
S(K) = \{ \tau \in H(\text{div}, K) \mid \tau|_{K_i} \in [P_k(K_i)]^{2 \times 2}, i = 1, 2, 3 \},
\end{equation}
with $k = 1$ for JM and $k = 2$ for ADG. For the postprocessing
\begin{equation}
V_h^* = \{ v \in [L^2(\Omega)]^d \mid v|_K \in [P_{k+1}(K)]^d \quad \forall K \in \mathcal{C}_h \}
\end{equation}
is used. Our first example contains a cured boundary and for that we use curved elements in order to retrieve the convergence rates of the analysis. To illustrate this in more details an example is given in Figure 1. Here we consider an element $K \in \mathcal{C}_h$ with the vertices $V_0, V_1$ and $V_2$ (the triangle filled with gray color). Now let $\Psi_K \in P^{k+2}(K)$ with $\Psi_K(K) = \tilde{K}$ be a polynomial mapping from $K$ to the curved triangle $\tilde{K}$ (filled with orange color), where we have chosen the order $k + 2$ as suggested in [9]. Then, in order to guarantee normal continuity, see (3.2), the stress finite elements are mapped by a Piola transformation, see [11], which includes the mapping $\Psi_K$. Thus if $\tilde{\tau}$ is a basis function on a given reference element $\tilde{K}$ and $\Phi_K$ denotes the linear mapping from $\tilde{K}$ to $K$, then the mapped basis function on $K$ is given by
\begin{equation}
\tau = \frac{1}{\det(D(\Psi_K \circ \Phi_K))} D(\Psi_K \circ \Phi_K) \tilde{\tau},
\end{equation}
where $D(\cdot)$ denotes the Jacobian. For more details we refer to [8, 9]. Note that the mapping $\Psi_K \circ \Phi_K$ is applied for all sub triangles as illustrated in Figure 1.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Barycentric refined element at the boundary of the domain (left) and corresponding curved element (right).}
\end{figure}

We have chosen two classical examples with known exact solutions [30]. In them the material parameters used are the Young modulus $E$ and the Poisson ratio $\nu$. They are related to the Lamé parameters by
\begin{equation}
\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \quad \mu = \frac{E}{2(1+\nu)}.
\end{equation}
8.1. *Circular hole in an infinite plate.* The problem is that of the unstressed circular hole with radius $a$ in an infinite plate subject to the unidirectional tension $\sigma_\infty$ as discussed in Section 19.3.1 in [30]. The exact displacement is

$$u_x = \frac{\sigma_\infty a}{8\mu} \left( \frac{r}{a} (\kappa + 1) \cos(\theta) + 2a \frac{\kappa}{r} ((1 + \kappa) \cos(\theta) + \cos(3\theta)) - 2a^3 \frac{\kappa}{r^3} \cos(3\theta) \right),$$

$$u_y = \frac{\sigma_\infty a}{8\mu} \left( \frac{r}{a} (\kappa - 3) \sin(\theta) + 2a \frac{\kappa}{r} ((1 - \kappa) \sin(\theta) + \sin(3\theta)) - 2a^3 \frac{\kappa}{r^3} \sin(3\theta) \right),$$

and the exact stress components are given by

$$\sigma_{xx} = \sigma_\infty \left( 1 - \frac{a^2}{r^2} \left( \frac{3}{2} \cos(2\theta) + \cos(4\theta) \right) + \frac{3}{2} \frac{a^4}{r^4} \cos(4\theta) \right),$$

$$\sigma_{yy} = \sigma_\infty \left( -\frac{a^2}{r^2} \frac{1}{2} \cos(2\theta) - \cos(4\theta) \right) - \frac{3}{2} \frac{a^4}{r^4} \cos(4\theta),$$

$$\sigma_{xy} = \sigma_\infty \left( -\frac{a^2}{r^2} \frac{1}{2} \sin(2\theta) + \sin(4\theta) \right) + \frac{3}{2} \frac{a^4}{r^4} \sin(4\theta).$$

The computations we do for the domain $\Omega = (-b,b) \times (-w,w) \setminus \Omega_0$ with the hole given by $\Omega_0 = \{(x,y) \in \mathbb{R}^2 : |(x,y)| \leq a\}$, see the left picture of Figure 2.

We choose $a = 1$ and $b = w = 4a$, and use the material parameters $E = 1$ and $\nu = 0.3, 0.4, 0.49, 0.49999$. On the outer boundary we assign the traction obtained from (8.4) with $\sigma_\infty = 1$. The displacement is fixed to be orthogonal to the rigid body motions.

To validate the theoretical findings we introduce the relative $L^2$-error of the stress and strain

$$(8.5) \quad e_0^\sigma = \frac{\|\sigma - \sigma_h\|_0}{\|\sigma\|_0}, \quad e_0^\varepsilon = \frac{\|\varepsilon(u) - \varepsilon(u_h)\|_0}{\|\varepsilon(u)\|_0},$$

for which the a priori estimates of Theorems 3 and 4, and the a posteriori estimate of Theorem 7 hold.
The second set of error quantities are the relative errors in strain energy for the stress directly obtained from the method, and computed from the displacement $u_h^a$, through (7.1)

$$e^C_{\sigma} = \frac{||\sigma - \sigma_h||_C}{||\sigma||_C}, \quad e^{A\varepsilon(u)}_{C} = \frac{||\sigma - A\varepsilon(u_h^a)||_C}{||\sigma||_C}. $$

The last set is those given by the Hypercircle estimate

$$e^{\text{mean}}_{C} = \frac{||\sigma - \frac{1}{2}(\sigma_h + A\varepsilon(u_h^a))||_C}{||\sigma||_C}, \quad c_{\text{eff}} = \frac{||\sigma - \frac{1}{2}(\sigma_h + A\varepsilon(u_h^a))||_C}{\frac{1}{2}||\sigma_h - A\varepsilon(u_h^a)||_C},$$

where $c_{\text{eff}}$ measures the efficiency of estimate (6.8). The oscillation is a higher order term, and we expect that $c_{\text{eff}} \to 1$ when $h \to 0$, which means that the error estimator is asymptotically exact. Further, we introduce the symbol $N$ for the number of elements in $C_h$. For uniform refinements we have $h \sim N^{-1/2}$.

In Table 1 and Table 2 the errors and the order of convergence (oc) for the JM and the ADG method are given for varying Poisson ratios for a uniform mesh refinement. As predicted by the analysis all errors converge with optimal order $O(N^{-1})$ and $O(N^{-3/2})$, for JM and ADG, respectively, and the constant $c_{\text{eff}}$ converges to 1. Further, the quantities $e^\sigma_0$, $e^u_0$ and $e^C_{\sigma}$ stay constant for all values $\nu$. Since in the incompressible limit $\nu \to \frac{1}{2}$, the Lamé parameter $\lambda \to \infty$, see (8.3), we expect...
that the errors, when computed from (7.1), should deteriorate. Indeed, although

converging with optimal order, the errors $e^\text{mean}_C$ and $e^A(u)_C$ start to grow significantly in the incompressible limit. To give more insight on this behaviour we have plotted in Figure 5 the error $e^A(u)_C$ for both methods with respect to the Poisson ratio. As we can see the blow up occurs continuously and can be made arbitrarily big if one approaches the incompressible limit.

We finally note that in all cases the stress $\sigma_h$ is much more accurate than that computed by $A(u_h)$. Furthermore, the latter dominates in $e^\text{mean}_C$, thus it holds $e^\text{mean}_C \sim \frac{1}{2} e^A(u)_C$, for all values of the Poisson ratio.

\subsection{L-shape example.} We employ an adaptive mesh refinement for the L-shape example from Section 10.3.2 of [30]. To this end let $\Omega$ be given by

\[\Omega = \{(x, y) : |x| + |y| \leq 2^{1/2}a \} \setminus \{(x, y) : |x - 2^{-1/2}a| + |y| \leq 2^{-1/2}a\},\]

as illustrated in the right picture of Figure 2. For our test case we choose $a = 1$ and use $E = 1$ and different values of $\nu$. Further, we define the constants $\alpha = 0.544483737$ and $Q = 0.543075579$. The exact displacement field, up to rigid-body displacements and rotations, is given by

\[u_x = \frac{1}{2\mu}r^\alpha((\kappa - Q(\alpha + 1)) \cos(\alpha \theta) - \alpha \cos((\alpha - 2)\theta))\]

\[u_y = \frac{1}{2\mu}r^\alpha((\kappa + Q(\alpha + 1)) \sin(\alpha \theta) + \alpha \sin((\alpha - 2)\theta)),\]
\[ e^0 = 9.9 \times 10^{-3} \] for the ADG method and varying 
\( \alpha \), \( \nu \), and \( e_{\text{rel}} \) in an infinite plate Example 8.1 for the ADG method and varying 

| \( N \) | \( e^0_{\text{ADG}} \) | \( e^0_{\text{JM}} \) | \( e^0_{\text{JM}} \) | \( e^0_{\text{ADG}} \) | \( e^0_{\text{JM}} \) | \( e^0_{\text{JM}} \) | \( e^0_{\text{ADG}} \) | \( e^0_{\text{JM}} \) | \( e^0_{\text{JM}} \) |
|---|---|---|---|---|---|---|---|---|---|
| 202 | 9.9 \times 10^{-3} | 2.5 \times 10^{-2} | 9.9 \times 10^{-3} | 2.8 \times 10^{-2} | 1.5 \times 10^{-2} | 0.94 |
| 808 | 2.7 \times 10^{-3} | 1.1 \times 10^{-2} | 2.7 \times 10^{-3} | 1.3 \times 10^{-2} | 6.7 \times 10^{-3} | 0.98 |
| 3232 | 3.4 \times 10^{-4} | 3.0 \times 10^{-4} | 3.4 \times 10^{-4} | 1.7 \times 10^{-3} | 8.5 \times 10^{-4} | 0.98 |
| 12928 | 4.0 \times 10^{-5} | 3.2 \times 10^{-5} | 4.0 \times 10^{-5} | 3.1 \times 10^{-4} | 9.5 \times 10^{-5} | 0.98 |

**Table 2.** Errors and estimated order of convergence for the hole in an infinite plate Example 8.1 for the ADG method and varying Poisson ratios \( \nu \).

\[ \nu = 0.49 \]

- **Figure 5.** The error \( e^0_{\text{ADG}} \) of the JM and ADG method for the hole in an infinite plate example 8.1 for varying Poisson ratios on a fixed mesh with \( N = 606 \).

and the stress components are

\[
\sigma_{xx} = \alpha \alpha^{-1} ((2 - Q(\alpha + 1)) \cos((\alpha - 1)\theta) - (\alpha - 1) \cos((\alpha - 3)\theta)),
\]

\[
\sigma_{yy} = \alpha \alpha^{-1} ((2 + Q(\alpha + 1)) \cos((\alpha - 1)\theta) + (\alpha - 1) \cos((\alpha - 3)\theta)),
\]

\[
\sigma_{xy} = \alpha \alpha^{-1} ((\alpha - 1) \sin((\alpha - 3)\theta) + Q(\alpha + 1) \sin((\alpha - 1)\theta)).
\]
To solve the problem we again enforce traction boundary conditions on the whole boundary $\partial \Omega$. We use a uniform refinement and adaptive refinements where we use the a posteriori error estimator of Theorem 6 and for the incompressible limit the estimator given in Theorem 7. Now let $K \in C_h$ be an arbitrary element, then we define the local contributions

$$
\eta(K)^2 = \frac{1}{4} \| \sigma_h - A \varepsilon(u_h^\alpha) \|_{C,K}^2 \quad \text{and} \quad \eta^{inc}(K)^2 = \mu^{1/2} \| C \sigma_h - \varepsilon(u_h^\alpha) \|_{0,K}^2,
$$

where $\| \cdot \|_{C,K}$ reads as the norm $\| \cdot \|_C$ restricted on the element $K$. The adaptive mesh refinement loop is defined as usual by

SOLVE $\rightarrow$ ESTIMATE $\rightarrow$ MARK $\rightarrow$ REFINE $\rightarrow$ SOLVE $\rightarrow$ ...

In the marking step we mark an element $K \in C_h$ for refinement if $\eta(K) \geq \frac{1}{4} \max_{K \in C_h} \eta(K)$ or $\eta^{inc}(K) \geq \frac{1}{4} \max_{K \in C_h} \eta^{inc}(K)$. After that, the mesh refinement algorithm refines the marked elements plus further elements to guarantee a regular triangulation. Beside the error quantities introduced above we further define the (relative) estimators

$$
\eta = \frac{\frac{1}{2} \| \sigma_h - A \varepsilon(u_h^\alpha) \|_C}{\| \sigma \|_C}, \quad \eta^{inc} = \frac{\mu^{1/2} \| C \sigma_h - \varepsilon(u_h^\alpha) \|_0}{\mu^{-1/2} \| \sigma \|_0},
$$

and the scaled error

$$
e_{\text{eff}}^{inc} = \frac{\mu^{1/2} \| \varepsilon(u) - \varepsilon(u_h^\alpha) \|_0}{\mu^{-1/2} \| \sigma \|_0}.
$$

In Figure 6 we plot the error history of $\varepsilon^\alpha_c$, $\varepsilon^A_c(u)$ and $e_{\text{eff}}^\text{mean}$ for the JM and the ADG method using an adaptive refinement based on the estimator $\eta$ for a moderate Poisson ratio $\nu = 0.3$. From the coarsest to the finest mesh the measure of efficiency $c_{\text{eff}}$ varies in in the range $0.99 - 1.00$, and hence the error estimator $\eta$ is not plotted as it would be indistinguishable from $e_{\text{eff}}^\text{mean}$. From the figure we see that all errors converge with optimal order $O(N^{-(k+1)/2})$.

To show the drastic decrease of the errors when using an adaptive algorithm, we also include the error $e_c^\alpha$ for a uniform refinement. Since the exact solution is in the Sobolev space $H^s$, with $s < 1.54$, a uniform mesh only yields a convergence rate of $O(h^{0.54})$, i.e. $O(N^{-0.27})$.

In Figure 7 we plot the same quantities but using an incompressible setting with $\nu = 0.49999$. Although all quantities still converge with an optimal order, we observe the same error deterioration as in the previous example. Thus, while $e_c^\alpha$ (and also $e_c^\varepsilon$, $e_c^{\text{eff}}$ which are not plotted) is not affected by the choice of the Poisson ratio $\nu$, the errors $e_{\text{eff}}^\text{mean}$ and $e_{\text{eff}}^{A_c(u)}$ and the estimator $\eta$ are much bigger and should not be used in practice.

To this end we follow the theory of Theorem 7, i.e. we employ the estimator $\eta^{inc}$. In Figure 8 the corresponding relative errors and the estimator are plotted and we observe that (up to an unknown constant $O(1)$) the error estimator gives a good prediction of the errors $e_c^\varepsilon$ and $e_{\text{eff}}^{\text{inc}}$. Further, all errors converge with optimal error $O(N^{-(k+1)/2})$, $k = 1, 2$. Again we include the errors when using an uniform refinement which shows the drastic decrease when using an adaptive algorithm.

References

[1] Douglas N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19(4):742–760, 1982.
Errors for JM for $\nu = 0.3$

Errors for ADG for $\nu = 0.3$

$\mathcal{O}(N^{-0.27}) -\cdots \mathcal{O}(N^{-3/2}) \cdots \mathcal{O}(N^{-1})$

$e^\sigma_C, e^{A(u)}_C, e^{\text{mean}}_C, \eta$

**Figure 6.** Error of the JM and ADG for the L-shape example 8.2 with an adaptive refinement using estimator $\eta$ and a constant Poisson ratio $\nu = 0.3$.

Errors for JM for $\nu = 0.49999$

Errors for ADG for $\nu = 0.49999$

$\mathcal{O}(N^{-0.27}) -\cdots \mathcal{O}(N^{-3/2}) \cdots \mathcal{O}(N^{-1})$

$e^\sigma_C, e^{A(u)}_C, e^{\text{mean}}_C, \eta$

**Figure 7.** Error of the JM and ADG for the L-shape example 8.2 with an adaptive refinement using estimator $\eta$ and a constant Poisson ratio $\nu = 0.49999$.

[2] Douglas N. Arnold, Gerard Awanou, and Ragnar Winther. Finite elements for symmetric tensors in three dimensions. *Math. Comp.*, 77(263):1229–1251, 2008.

[3] Douglas N. Arnold and Franco Brezzi. Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. *RAIRO Modél. Math. Anal. Numér.*, 19(1):7–32, 1985.

[4] Douglas N. Arnold, Jim Douglas, Jr., and Chaitan P. Gupta. A family of higher order mixed finite element methods for plane elasticity. *Numer. Math.*, 45(1):1–22, 1984.

[5] Douglas N. Arnold, Richard S. Falk, and Ragnar Winther. Finite element exterior calculus, homological techniques, and applications. *Acta Numer.*, 15:1–155, 2006.
ANALYSIS OF MIXED FINITE ELEMENTS FOR ELASTICITY

Figure 8. Error of the JM and ADG for the L-shape example 8.2 with an adaptive refinement using estimator \( \eta^{\text{inc}} \) and a constant Poisson ratio \( \nu = 0.49999 \).

- \( \mathcal{O}(N^{-0.27}) \)
- \( \mathcal{O}(N^{-3/2}) \)
- \( \mathcal{O}(N^{-1}) \)

\[
\begin{align*}
e^0 & \quad e^0_{\text{inc}} \quad e^0_{\eta^{\text{inc}}} \quad e^0_{\text{uni}}
\end{align*}
\]

[6] Ivo Babuška, John Osborn, and Juhani Pitkäranta. Analysis of mixed methods using mesh dependent norms. *Math. Comp.*, 35(152):1039–1062, 1980.

[7] Ivo Babuška and Manil Suri. Locking effects in the finite element approximation of elasticity problems. *Numer. Math.*, 62(4):439–463, 1992.

[8] Christine Bernardi. Optimal finite-element interpolation on curved domains. *SIAM Journal on Numerical Analysis*, 26(5):1212–1240, 1989.

[9] Fleurianne Bertrand and Gerhard Starke. Parametric raviart–thomas elements for mixed methods on domains with curved surfaces. *SIAM Journal on Numerical Analysis*, 54(6):3648–3667, 2016.

[10] Raymond L. Bisplinghoff, James W. Mar, and Theodore H. H. Pian. *Statics of deformable solids*. Dover Publications, Inc., New York, 1990. Corrected reprint of the 1965 original.

[11] Daniele Boffi, Franco Brezzi, and Michel Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.

[12] Susanne C. Brenner. Korn’s inequalities for piecewise \( H^1 \) vector fields. *Math. Comp.*, 73(247):1067–1087, 2004.

[13] Richard Courant and David Hilbert. *Methods of mathematical physics. Vol. I*. Interscience Publishers, Inc., New York, N.Y., 1953.

[14] Daniele Antonio Di Pietro and Alexandre Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications (Berlin)* [Mathematics & Applications]. Springer, Heidelberg, 2012.

[15] Baudouin M. Fraejs de Veubeke. Displacement and equilibrium models in the finite element method. In O.C Zienkiewics and G.S. Holister, editors, *Stress analysis*, pages 145–197. Wiley, 1965.

[16] Kurt O. Friedrichs. Ein Verfahren der Variationsrechnung, das Minimum eines Integrals als das Maximum eines anderen Ausdruckes darzustellen. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, pages 13–20, 1929.

[17] Vivette Girault and Pierre-Arnaud Raviart. *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1986. Theory and algorithms.

[18] Thirupathi Gudi. A new error analysis for discontinuous finite element methods for linear elliptic problems. *Math. Comp.*, 79(272):2169–2189, 2010.

[19] Johnny Guzmán and Michael Neilan. Symmetric and conforming mixed finite elements for plane elasticity using rational bubble functions. *Numer. Math.*, 126(1):159–171, 2014.
[20] Antti Hannukainen, Rolf Stenberg, and Martin Vohralík. A unified framework for a posteriori error estimation for the Stokes problem. Numer. Math., 122(4):725–769, 2012.
[21] Claes Johnson and Bertrand Mercier. Some equilibrium finite element methods for two-dimensional elasticity problems. Numer. Math., 30(1):103–116, 1978.
[22] Ohannes A. Karakashian and Frederic Pascal. A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. SIAM J. Numer. Anal., 41(6):2374–2399, 2003.
[23] Carlo Lovadina and Rolf Stenberg. Energy norm a posteriori error estimates for mixed finite element methods. Math. Comp., 75(256):1659–1674 (electronic), 2006.
[24] Jindřich Nečas and Ivan Hlaváček. Mathematical theory of elastic and elasto-plastic bodies: an introduction, volume 3 of Studies in Applied Mechanics. Elsevier Scientific Publishing Co., Amsterdam-New York, 1980.
[25] J. Pitkäranta and R. Stenberg. Analysis of some mixed finite element methods for plane elasticity equations. Math. Comp., 41(164):399–423, 1983.
[26] W. Prager and J. L. Synge. Approximations in elasticity based on the concept of function space. Quart. Appl. Math., 5:241–269, 1947.
[27] P.-A. Raviart and J. M. Thomas. A mixed finite element method for 2nd order elliptic problems. In Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975), pages 292–315. Lecture Notes in Math., Vol. 606. Springer, Berlin, 1977.
[28] J. Schöberl. NETGEN An advancing front 2D/3D-mesh generator based on abstract rules. Computing and Visualization in Science, 1(1):41–52, 1997.
[29] L. R. Scott and M. Vogelius. Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials. RAIRO Modél. Math. Anal. Numér., 19(1):111–143, 1985.
[30] B. A. Szabó and I. Babuška. Finite Element Analysis. A Wiley-Interscience publication. Wiley, 1991.
[31] Rüdiger Verfürth. A posteriori error estimation techniques for finite element methods. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.
[32] V.B. Watwood and B.J. Hartz. An equilibrium stress field model for finite element solution of two-dimensional elastostatic problems. Internat. Jour. Solids and Structures, 4:857–873, 1968.