FINITE ELEMENT METHOD FOR LINEAR ELASTICITY

HONGTAO CHEN

1. Mathematical Background

Here we discuss elasticity, and derive the weak formulation of the pure displacement form of the linear elasticity equations with mixed boundary conditions.

1.1. Linear Elasticity.

1.1.1. *Introduction to elasticity.* We study the displacement field $u : \Omega \to \mathbb{R}^d$ resulting from letting a medium originally at equilibrium reach a second equilibrium after applying an external load. We call the external load applied to our medium $f : \Omega \to \mathbb{R}^d$. Thus set up, elasticity theory is, in general, non-linear, and satisfies Hookes Law

$$\sigma = C : \epsilon$$

and the equilibrium equation $\mathrm{div}(\sigma) + f = 0$, where σ is the Cauchy stress tensor, C is the fourth-order stiffness tensor, and ϵ is the strain tensor.

This course focuses on the more specific case of homogeneous linear isotropic materials. Under a force, isotropic materials deform the same way (relative to the direction of the force) regardless of the relative orientations of the force and the material.

In order to model linear elasticity, we need two parameters characterizing the material. We primarily use the Lame parameters λ and μ, which together describe the compressibility of the material. As is usually done with linear elasticity, we take λ and μ to be constants.

We then write the stress tensor of the deformed medium as $\sigma(u) : \Omega \to \mathbb{R}^{d,d}$. Specifically, we get the linear elasticity version of Hookes Law:

$$\sigma(u) = \lambda(\mathrm{div}\, u)I + 2\mu\epsilon(u)$$

where the strain rate tensor $\epsilon(u) : \Omega \to \mathbb{R}^{d,d}$ for our deformed medium is the symmetric part of the gradient tensor

$$\epsilon(u) = \frac{1}{2}(\nabla u + (\nabla u)^T).$$

1.1.2. *Strong form of the linear elasticity equations.* Our equations and assumptions above then give us the following strong form of the linear elasticity equations: Find $u : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\begin{gathered}
\mathrm{div}(\sigma(u)) + f = 0, \quad \text{in} \quad \Omega, \\
\sigma(u) = \lambda(\mathrm{div}\, u)I + 2\mu\epsilon(u), \quad \text{in} \quad \Omega, \\
u = 0, \text{on} \quad \partial\Omega_D, \quad \sigma(u) \cdot n = g, \text{on} \quad \partial\Omega_N
\end{gathered}$$

where f is the volume force applied on Ω and g is the normal force applied on $\partial\Omega_N$.

1.1.3. *Derivation of the pure displacement weak form of the linear elasticity equations.* We begin with a few lemmas that we will need to derive and simplify our weak form.

Lemma 1.1. *Let $\sigma(u)$ be defined as above, and let v be a vector-valued function. Then $\sigma(u) : \nabla v = \sigma(u) : \epsilon(v)$.*

Proof. Firstly, we have

$$\begin{aligned}\sigma(u) &= \lambda(\operatorname{div} u)I + 2\mu\epsilon(u) \\ &= \begin{pmatrix} (\lambda+2\mu)u_{1x} + \lambda u_{2y} & \mu(u_{1y}+u_{2x}) \\ \mu(u_{1y}+u_{2x}) & \lambda u_{1x} + (\lambda+2\mu)u_{2y} \end{pmatrix}\end{aligned}$$

which leads to

$$\begin{aligned}\sigma(u) : \nabla v &= ((\lambda+2\mu)u_{1x} + \lambda u_{2y})v_{1x} + \mu(u_{1y}+u_{2x})v_{1y} \\ &\quad + \mu(u_{1y}+u_{2x})v_{2x} + (\lambda u_{1x} + (\lambda+2\mu)u_{2y})v_{2y},\end{aligned}$$

and

$$\begin{aligned}\sigma(u) : \epsilon(v) &= ((\lambda+2\mu)u_{1x} + \lambda u_{2y})v_{1x} + \mu(u_{1y}+u_{2x}) \cdot \frac{1}{2}(v_{1y}+v_{2x}) \\ &\quad + \mu(u_{1y}+u_{2x})\frac{1}{2}(v_{1y}+v_{2x}) + (\lambda u_{1x} + (\lambda+2\mu)u_{2y})v_{2y}.\end{aligned}$$

Therefore, we obtain $\sigma(u) : \nabla v = \sigma(u) : \epsilon(v)$. □

Lemma 1.2. *Let $\sigma(u)$ be defined as above, and let v be a vector-valued function. Then $\sigma(u) : \epsilon(v) = \lambda(\operatorname{div} u)(\operatorname{div} v) + 2\mu\epsilon(u) : \epsilon(v)$.*

Proof. It suffices to note that $I : \epsilon(v) = \operatorname{div} v$. □

By testing v with $v = 0$ on $\partial\Omega_D$, the two lemmas yields:

$$\begin{aligned}(f, v) &= (-\operatorname{div}(\sigma(u)), v) = (\sigma(u), \nabla v) - (\sigma(u) \cdot n, v)_{\partial\Omega_N} \\ &= (\sigma(u), \epsilon(v)) - (g, v)_{\partial\Omega_N} \\ &= \lambda(\operatorname{div} u, \operatorname{div} v) + 2\mu(\epsilon(u), \epsilon(v)) - (g, v)_{\partial\Omega_N}.\end{aligned}$$

If we define $H^1_{0D} := \{u \in (H^1(\Omega))^d \mid u = 0 \text{ on } \partial\Omega_D\}$, this leads us to the following pure displacement weak formulation of the linear elasticity equations: Find $u \in H^1_{0D}$ such that

$$a(u, v) = (f, v) + (g, v)_{\partial\Omega_N}, \quad \forall v \in H^1_{0D}, \tag{1.1}$$

where $a(u, v) = \lambda(\operatorname{div} u, \operatorname{div} v) + 2\mu(\epsilon(u), \epsilon(v))$.

1.1.4. *Well-posedness of the pure displacement weak form.* To argue that our weak form is well-posed, we must first establish some preliminary theorems.

Theorem 1.3. *(Korns Inequality.) Let $\Omega \subset \mathbb{R}^d$. There exists c such that for all $v \in [H^1(\Omega)]^d$, $c\|v\|_{[H^1(\Omega)]^d} \leq \|\epsilon(v)\|_{[L^2(\Omega)]^d} + \|v\|_{L^2(\Omega)}$.*

This result and its proof can be found in [5].

Theorem 1.4. *There exists a positive constant c such that $\|\epsilon(v)\|_{[L^2(\Omega)]^d} \geq c\|v\|_{[H^1(\Omega)]^d}, \forall v \in H^1_0(\Omega)$.*

Proof.

$$\begin{aligned}\|\epsilon(u)\|^2 &= \frac{1}{4}(\nabla u + \nabla u^T, \nabla u + \nabla u^T)\\ &= \frac{1}{2}\|\nabla u\|^2 + \frac{1}{2}(\nabla u, \nabla u^T).\end{aligned}$$

Now if $u \in H_0^1 \cap H^2$ we can integrate by parts to find that

$$(\nabla u, \nabla u^T) = -(u, \operatorname{div}(\nabla u^T)) = -(u, \nabla(\operatorname{div} u)) = \|\operatorname{div} u\|^2.$$

By density this holds for all $u \in H_0^1$, without requiring also $u \in H^2$ .Therefore, it follows that $\|\epsilon(v)\|_{[L^2(\Omega)]^d} \geq c\|v\|_{[H^1(\Omega)]^d}$. □

Remark 1.5. The same inequality can also holds for $v \in H_{0D}^1(\Omega)$, although the proof is more difficult [5].

By careful calculation, we have

Lemma 1.6.

$$\epsilon(u) : \epsilon(v) \leq \nabla u : \nabla v$$

.

Proof. We give the proof in two dimensions for the sake of clarity, which can be proved from $\nabla u : \nabla v - \epsilon(u) : \epsilon(v) = \frac{1}{2}(u_{1y} - u_{2x})^2 \geq 0$. □

Lemma 1.7. $\operatorname{div} u \cdot \operatorname{div} u \leq d\nabla u \cdot \nabla u$.

Proof. Proved by Cauchy-Schwarz inequality. □

Since,

$$a(u, u) = \lambda(\operatorname{div} u, \operatorname{div} u) + 2\mu(\epsilon(u), \epsilon(v)) \leq 2(\lambda + \mu)\|u\|_1^2$$

it immediately leads to

$$a(u, v) \leq (a(u, u))^{1/2}(a(v, v))^{1/2} \leq 2(\lambda + \mu)\|u\|_1\|v\|_1.$$

Therefore, $a(u, v)$ is bounded.

From the Korn's inequality, $a(u, u) \geq 2\mu c\|u\|_1^2$. Thus, $a(u, v)$ is coercive.

Now that we have satisfied all the assumptions of the Lax-Milgram Theorem, we can apply it to our problem to conclude that

Theorem 1.8. *There exists a unique solution $u \in H_{0D}^1$ to the equation $a(u, v) = (f, v) + (g, v)_{\partial\Omega_N}$. Moreover, $\|u\|_{[H^1(\Omega)]^d} \leq C(\|f\|_{[H^{-1}(\Omega)]^d} + \|g\|_{[H^{-1/2}(\partial\Omega)]})$.*

Thus our pure displacement weak formulation of the linear elasticity equations is well-posed.

1.1.5. *Displacement-pressure form of the linear elasticity equations.* We introduce in this section the displacement-pressure formulation of the linear elasticity equations. Substituting a scalar unknown p (which can be identified with pressure) for $-\lambda \operatorname{div} u$ into the pure-displacement strong form yields the following displacement-pressure strong form of the linear elasticity equations.

$$-2\mu(\operatorname{div} \epsilon(u)) + \nabla p = f, \tag{1.2}$$

$$\operatorname{div} u + \frac{1}{\lambda} p = 0. \tag{1.3}$$

This leads to the following displacement-pressure weak form of the linear elasticity equations.

$$2\mu(\epsilon(u), \epsilon(v)) - (\operatorname{div} v, p) = (f, v), \quad \forall v \in H_0^1, \tag{1.4}$$

$$(\operatorname{div} u, q) + \frac{1}{\lambda}(p, q) = 0, \quad \forall q \in L^2. \tag{1.5}$$

The displacement-pressure formulation comes directly from the displacement formulation, which we have already proved it's well-posed.

1.2. **Finite element method.** Now we apply the Galerkin method to discretize our weak form into a solvable form. If we definie

$$A_{ij} = \lambda(\operatorname{div} \phi_i, \operatorname{div} \phi_i) + 2\mu(\epsilon(\phi_i), \epsilon(\phi_j)), \quad F_i = (f, \phi_i) + (g, \phi_i)_{\partial\Omega_N}$$

then the unknowns can be found from $AU = F$. Moreover, from Ceas Lemma we have the follow error estimates.

$$\|u - u_h\|_{H^1} \leq C \inf_{v \in V_h} \|u - v\|_{H^1}.$$

If k refers to the degree of our polynomial space, then we also have

$$\|u - u_h\|_{H^1} \leq Ch^k \|u\|_{k+1}.$$

Thus, we have a well-posed discretized system such that we expect the solution to the discretized problem to converge to the true solution as we refine our mesh.

2. Locking for Nearly Incompressible Materials

2.1. **Nearly incompressible materials.** We say a solid material is incompressible if applying pressure to it does not change its shape. Thus, a nearly incompressible solids shape does not change very much under pressure. Examples of nearly incompressible elastic materials include natural rubber and solid propellants. Mathematically speaking, nearly incompressible materials are often characterized by Lames first parameter λ large or by Poissons parameter ν close to 1/2 which is expressed by $\nu = \frac{\lambda}{2(\lambda+\mu)}$. Thus for a fixed , these two characterizations of incompressibility are equivalent.

Locking is a phenomenon that occurs in finite element approximations of the pure displacement formulation of linear elasticity for certain nearly incompressible materials. That is, locking is inaccuracy of the previous displacement solution u_h. More specifically, locking occurs when the finite element computation produces unrealistically small displacements [6].

We can analyze from a mathematical perspective why locking occurs. Firstly since,

$$2\mu c\|u\|_1^2 \leq a(u, u) \leq 2(\lambda + \mu)\|u\|_1^2,$$

the coefficient constant of the error bound is $(\lambda + \mu)/c\mu$ and rendered ineffective for λ large. Secondly, $\|u_h\|_1$ is too small, even though that is not necessarily what we want. We know from [6] that locking occurs when $\operatorname{div} v_h \geq C_h\|v_h\|_{H^1}$. Since $\|\epsilon(v_h)\| \geq c\|v_h\|_1$, we can derive that

$$a(u_h, u_h) \geq (\lambda C_h^2 + 2\mu c)\|u_h\|_{H^1}^2.$$

Then our discretized displacement is bounded, since $\|u_h\|_{H^1} \leq \frac{1}{(\lambda C_h^2 + 2\mu c)}\|f\| \leq \frac{1}{\lambda C_h^2}\|f\|$.As we can see from this bound, for λ large, $\|u_h\|_{H^1}$ is forced to be small for a fixed h. However, as h gets smaller, it can begin to compensate for the large value of λ. Thus, we expect $\|u_h\|_{H^1}$ to eventually rise to a reasonable value for a fixed large λ as we decrease h. In other words, for a large value of λ we expect a discretized solution to be too small on a coarse mesh, but more accurate on a fine mesh. In order to examine locking more precisely, we also show a table with

	$\mathbf{Q_1}$			$\mathbf{Q_2}$		
$(\mu = 100)$ h	$\lambda = 100$, $\nu = 0.25$	$\lambda = 5\times10^4$, $\nu = 0.499$	$\lambda = 5\times10^5$, $\nu = 0.4999$	$\lambda = 100$, $\nu = 0.25$	$\lambda = 5\times10^4$, $\nu = 0.499$	$\lambda = 5\times10^5$, $\nu = 0.4999$
0.707107	-1.93324	-0.200963	-0.0933249	-2.00141	-1.28297	-1.27746
0.353553	-1.98402	-0.446974	-0.131043	-2.00278	-1.30861	-1.30472
0.176777	-1.99807	-0.842845	-0.259832	-2.00326	-1.31904	-1.3158
0.0883883	-2.00196	-1.14013	-0.572642	-2.00344	-1.32336	-1.3205
0.0441942	-2.00307	-1.26541	-0.962982	-2.00351	-1.32521	-1.32254
Extrapolated	-2.00357	-1.32685	-1.32437	-2.00357	-1.32685	-1.32437
Relative difference	0.03510	0.84854	0.92953	0.00108	0.03307	0.03542

the y-displacements of the point $(8, -1)$ for the Q^1 and Q^2 discretizations. Simply comparing the coarse- and fine-mesh displacements in the incompressible cases demonstrates the locking we previously mentioned. The coarse-mesh displacements for Q^1 are dramatically smaller than those on the fine mesh, which is precisely the locking effect we described. This effect can also be seen for the Q^2 displacements, although it is less drastic. For a material that is certainly not nearly incompressible ($\lambda = 100$ or $\nu = 0.25$), however, the difference between the coarse-mesh and fine-mesh solutions is much smaller.We also calculate for comparison the relative difference between the coarse-mesh solutions and extrapolated solutions. The extrapolated solutions were determined using Richardson extrapolation with the last three Q^2 displacements for each value of λ, since Q^2 discretization is more accurate than Q^1 discretization.

2.2. **Displacement-pressure formulation.** To overcome the locking, we use the displacement-pressure formulation of the linear elasticity questions in Sec. 1. We can find the constants of the error estimates does not depend on λ.

3. Stress-displacement formulation

In the isotropic case, since $\epsilon(u) = \nabla u - \gamma\chi$ where $\gamma = \frac{1}{2}rot(u) = -\frac{1}{2}(\frac{\partial v_1}{\partial y} - \frac{\partial v_2}{\partial x})$ and

$$\chi = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

the mapping $\sigma \to A\sigma$ has the form

$$A\sigma = (1/2\mu)(\sigma - \frac{\lambda}{2\mu + d\lambda} tr(\sigma)I) = \epsilon(u).$$

We now turn to other types of weak formulations involving both σ and u. One of these is to seek $\sigma \in H(\mathrm{div}, \Omega; \mathbb{S})$, the space of square-integrable symmetric matrix fields with square-integrable divergence, and $u \in L^2(\Omega; V)$, satisfying

$$(A\sigma, \tau) + (\mathrm{div}\,\tau, u) = 0, \quad \forall \tau \in H(\mathrm{div}, \Omega; \mathbb{S}), \tag{3.1}$$

$$(\mathrm{div}\,\sigma, v) = (f, v), \quad \forall q \in L^2(\Omega; V). \tag{3.2}$$

where $(A\sigma,\tau) := \frac{1}{2\mu}\left[(\sigma,\tau) - \frac{\lambda}{2\mu+d\lambda}(tr(\sigma), tr(\tau))\right]$. The wellposedness of the above formulation is deduced by C. Kenig.

A second weak formulation, that enforces the symmetry weakly, seeks $\sigma \in H(\mathrm{div},\Omega;\mathbb{M})$, $u \in L^2(\Omega;V)$, and $p \in L^2(\Omega)$ satisfying

$$(A\sigma,\tau) + (\operatorname{div}\tau, u) + (as(\tau), p) = 0, \quad \forall \tau \in H(\mathrm{div},\Omega;\mathbb{M}), \tag{3.3}$$

$$(\operatorname{div}\sigma, v) = (f, v), \quad \forall q \in L^2(\Omega;V), \tag{3.4}$$

$$(as(\sigma), q) = 0, \quad \forall q \in L^2(\Omega), \tag{3.5}$$

where $\mathbb{M}$ is the space of $n \times n$ matrices and $as(\tau) := \tau : \chi$. Stable finite element discretizations with reasonable computational complexity based on the variational formulation (3.1) have proved very difficult to construct. One successful approach has been to use composite elements, in which the approximate displacement space consists of piecewise polynomials with respect to one triangulation of the domain, while the approximate stress space consists of piecewise polynomials with respect to a different, more refined, triangulation. The simplest and lowest order spaces in the family of spaces constructed consist of discontinuous piecewise linear vector fields for displacements and a stress space which is locally the span of piecewise quadratic matrix fields and the cubic matrix fields that are divergence-free. Hence, it takes 24 stress and six displacement degrees of freedom to determine an element on a given triangle. A simpler first-order element pair with 21 stress and three displacement degrees of freedom per triangle is also constructed in [4]. In three dimensions, a piecewise quartic stress space is constructed with 162 degrees of freedom on each tetrahedron in [1].

Because of the lack of suitable mixed elasticity elements that strongly impose the symmetry of the stresses, a number of authors have developed approximation schemes based on the weak symmetry formulation (4.1). Although (3.1) and (4.1) are equivalent on the continuous level, an approximation scheme based on (4.1) may not produce a symmetric approximation to the stress tensor, depending on the choices of finite element spaces.

The remainder of these notes will be devoted to the development and analysis of mixed finite element methods based on the formulation (4.1) of the equations of elasticity with weak symmetry. An important advantage of such an approach is that it allows us to approximate the stress matrix by the copies of standard finite element approximations of $H(\mathrm{div},\Omega)$ used to discretize scalar second order elliptic problems.

4. Mixed formulation of the equations of elasticity with weak symmetry

Mixed formulation of the equations of elasticity with weak symmetry is to seek $\sigma \in H(\mathrm{div},\Omega;\mathbb{M})$, $u \in L^2(\Omega;V)$, and $p \in L^2(\Omega)$ satisfying

$$(A\sigma,\tau) + (\operatorname{div}\tau, u) + (as(\tau), p) = 0, \quad \forall \tau \in H(\mathrm{div},\Omega;\mathbb{M}), \tag{4.1}$$

$$(\operatorname{div}\sigma, v) = (f, v), \quad \forall q \in L^2(\Omega;V), \tag{4.2}$$

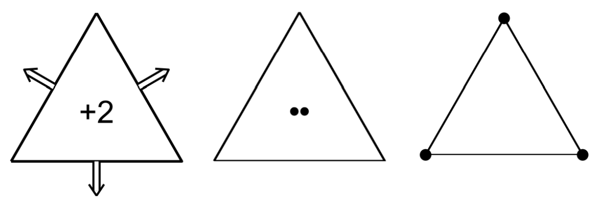
$$(as(\sigma), q) = 0, \quad \forall q \in L^2(\Omega). \tag{4.3}$$

This is wellposed which follow from the general theory of such saddle point problems once we establish two conditions: (1) $\|\tau\|^2_{H(\mathrm{div})} \le c_1(A\tau,\tau)$, whenever τ satisfies $(\operatorname{div}\tau, v) = 0$ and $(as(\tau), q) = 0$, (2) for all nonzero (v, q), there exists nonzero τ, such that $(\operatorname{div}\tau, v)+(as(\tau), q) \geq c_2\|\tau\|_{H(\operatorname{div})}(\|v\|+\|q\|)$, for some positive constants c_1 and c_2.

In this section we restrict the domain $\Omega \subset \mathbb{R}^2$. We also define for each T the bubble function b_T and normalized by $\int_T b_T = 1$. Define B to be the span of $\{rot(b_T)\}$ and $\mathbb{B} = \{\tau|(\tau_{i1}, \tau_{i2}) \in B, i = 1, 2\}$. Similarly, we define $\mathbb{R}T_0 = \{\tau|(\tau_{i1}, \tau_{i2}) \in RT_0, i = 1, 2\}$. The PEERS space [3] for the approximation of the stress is given by

$$V_h = \mathbb{R}T_0 + \mathbb{B}.$$

The displacement and rotation will be approximated in vector P_0 and continuous P_1 (CP_1).



The numerical solution is to seek $(\sigma_h, u_h, p_h) \in V_h \times P_0 \times CP_1$ satisfying

$$(A\sigma_h, \tau) + (\operatorname{div}\tau, u_h) + (as(\tau), p_h) = 0, \quad \forall \tau \in V_h, \tag{4.4}$$

$$(\operatorname{div}\sigma, v) = (f, v), \quad \forall q \in P_0, \tag{4.5}$$

$$(as(\sigma), q) = 0, \quad \forall q \in CP_1. \tag{4.6}$$

If define $\Pi_h : H^1 \to \mathbb{R}T_0$ as the Cartesian product of this operator and infer that

$$(\operatorname{div}(\tau - \Pi_h\tau), q) = 0, \ \forall q \in P_0, \quad \|\tau - \Pi_h\tau\| \leq ch\|\tau\|_1, \ \tau \in H^1.$$

Moreover, $\operatorname{div}\mathbb{B} = 0$ so that $\operatorname{div} V_h = \operatorname{div}\mathbb{R}T_0 = P_0$. Therefore, letting P_h be L^2 projection, we have

$$(\operatorname{div}\tau, v - P_h v) = 0, \ \tau \in V_h, v \in L^2.$$

It follows that

$$P_h \operatorname{div} = \operatorname{div}\Pi_h, \ \tau \in H^1.$$

For any space $X \in L^2$, define

$$\hat{X} = \{x \in X \mid \int tr(x) = 0\}$$

Then $V_h = \hat{V}_h \oplus \mathbb{R}I$. Putting $\tau = I$ into (4.4), we derive $\sigma_h \in \hat{V}_h$ satisfying

$$(A\sigma_h, \tau) + (\operatorname{div}\tau, u_h) + (as(\tau), p_h) = 0, \forall \tau \in \hat{V}_h.$$

Theorem 4.1. *There exists a unique element* (σ_h, u_h, p_h) *satisfying* (4.4)-(4.6). *Moreover, there exists a constant* C *such that*

$$\|\sigma - \sigma_h\|_{H(\operatorname{div})} + \|u - u_h\| + \|p - p_h\| \leq C\inf(\|\sigma - \tau\|_{H(\operatorname{div})} + \|u - v\| + \|p - q\|)$$

for $(\tau, v, q) \in$ *PEERS.*

The proof of Theorem 4.1 will be based on the abstract stability theory for mixed methods of [5] applied to the alternate characterization of the discrete solution using (4.4)-(4.6). It suffices to prove the following two lemmas.

Lemma 4.2. *There exists a constant C such that $C(A\tau,\tau) \geq \|\tau\|^2_{H(\mathrm{div})}$ for all $\tau \in Z_h = \{\tau \in \hat{V}_h \mid (\mathrm{div}\,\tau, v) + (as(\tau), q) = 0, v \in P_0, q \in CP_1\}$. In fact, the inequality hold for all divergence-free $\tau \in H(\hat{\mathrm{div}})$.*

Lemma 4.3. *There exists a constant C such that, for all pairs $v \in P_0, \eta \in CP_1$, there exists a nonzero $\tau \in \hat{V}_h$ such that $C(\mathrm{div}\,\tau, v) + (as(\tau),\eta) \geq \|\tau\|_{H(\mathrm{div})}(\|v\| + \|\eta\|)$. Moreover, τ can be chosen so that $\mathrm{div}\,\tau = v$ and $(as(\tau) - \eta, z) = 0, \quad \forall z \in CP_1$.*

Before proving these two lemmas, let us note that Lemma 4.2 establishes 1st Brezzi condition when $\hat{V}_h$ is considered in place of V_h. The second Brezzi condition follows from Lemma 4.3, so that the error estimate of Theorem 4.1 is valid, except that the range of τ in the infimum must be restricted to $\hat{V}_h$, since $\sigma \in H(\hat{\mathrm{div}})$,

$$\inf_{\tau \in V_h} \|\sigma - \tau\|_{H(\mathrm{div})} = \inf_{\tau \in \hat{V}_h} \|\sigma - \tau\|_{H(\mathrm{div})}$$

and Theorem 4.1 follows from Lemma 4.2 and 4.3.

Proof of Lemma 4.2: For $\tau \in V_h$, $\mathrm{div}\,\tau \in P_0$; hence, the condition $(\mathrm{div}\,\tau, v) = 0, \forall v \in P_0$ implies that

$$\mathrm{div}\,\tau = 0.$$

Consequently, it suffices to demonstrate the inequality for divergence-free tensors in $H(\hat{\mathrm{div}})$ i.e., $\tau \in H(\mathrm{div})$ such that

$$\mathrm{div}\,\tau = 0, \qquad \int tr(\tau) = 0$$

Take $v \in H^1_0$, such that $\mathrm{div}\, v = tr(\tau)$ and $\|v\|_1 \leq C\|tr(\tau)\|$. Now, setting $\tilde{\tau} = \tau - tr(\tau)I/2$, we have

$$\begin{aligned} \|tr(\tau)\|^2 &= (tr(\tau), I : \nabla v) = -2(\tilde{\tau}, \nabla v) - 2(\mathrm{div}\,\tau, v) \\ &= -2(\tilde{\tau}, \nabla v) \leq 2\|\tilde{\tau}\|\|v\|_1, \end{aligned}$$

so that $\|tr(\tau)\| \leq C\|\tilde{\tau}\|$. Further,

$$\|\tau\|^2_{H(\mathrm{div})} = \|\tau\|^2 = \|\tilde{\tau}\|^2 + \frac{1}{2}\|tr(\tau)\|^2 \leq C\|\tilde{\tau}\|^2.$$

Hence,

$$(A\tau,\tau) = \frac{1}{2\mu}\|\tilde{\tau}\|^2 + \frac{1}{4(\mu+\lambda)}\|tr(\tau)\|^2 \geq (C\mu)^{-1}\|\tau\|^2_{H(\mathrm{div})}.$$

Proof of Lemma 4.3: Given $v \in P_0, q \in CP_1$, take $\rho \in H^1$ such that $\mathrm{div}\,\rho = v$ and $\|\rho\|_1 \leq C\|v\|$. Let $\tau^1 = \Pi_h \rho$, then we have

$$\mathrm{div}\,\tau^1 = v, \quad \|\tau^1\|_{H(\mathrm{div})} \leq C\|v\|.$$

Set s equal to the mean value of $\eta - as(\tau^1)$ on Ω; so, $|s| \leq C(\|\tau^1\| + \|\eta\|) \leq C(\|v\| + \|\eta\|)$, and $\beta = \eta - as(\tau^1) - s$ has mean value zero. Thus, we can find $q \in H^1_0$ such that

$$\mathrm{div}\, q = \beta, \quad \|q\|_1 \leq C\|\beta\| \leq C(\|v\| + \|\eta\|)$$

By standard interpolation results [5] we can approximate q by $q_h \in CP_1 \cap H^1_0$ such that

$$\|q_h\|_1 \leq C\|q\|_1, \qquad \sum_{T \in \mathcal{T}_h} h_T^{-2}\|q - q_h\|^2_{0,T} \leq C\|q\|^2_1.$$

Set

$$a_T = \int_T (q - q_h)dx \in \mathbb{R}^2, \quad T \in \mathcal{T}_h$$

and note that

$$|a_T| \le Ch_T \|q - q_h\|_{0,T}$$

Next, recall that b_T denotes the normalized bubble functions on T and set

$$r = q_h + \sum_{T\in\mathcal{T}_h} a_T b_T \in H_0^1$$

so,

$$\int_T r = \int_T q_h + a_T = \int_T q, \quad T \in \mathcal{T}_h.$$

Since $\|\nabla b_T\|_{0,T} \le Ch_T^{-2}$, it follows that

$$\|\nabla a_T b_T\|_{0,T} \le Ch_T^{-1}\|q - q_h\|_{0,T}$$

and so,

$$\|\nabla a_T b_T\|^2 = \sum_{T\in\mathcal{T}_h} \|\nabla a_T b_T\|_{0,T}^2 \le C\|q\|_1^2.$$

Thus, the bound follows that

$$\|r\|_1 \le C\|\nabla r\| \le C(\|v\| + \|\eta\|).$$

Now set $\tau^2 = \tau^1 + \begin{pmatrix} rot(r_1) \\ rot(r_2) \end{pmatrix} + \frac{s}{2}\chi$. One easily verifies that $\tau^2 \in V_h$ and that

$$\|\tau^2\| \le C(\|\tau^1\| + \|r\|_1 + |s|) \le C(\|v\| + \|\eta\|).$$

Moreover,

$$\operatorname{div}\tau^1 = \operatorname{div}\tau^2 = v$$

and for any $z \in CP_1$,

$$(as(\tau^2), z) = (as(\tau^1) + \operatorname{div} r + s, z) = (\eta - \beta, z) - (r, \nabla z).$$

But, $\nabla z \in P_0$, so that

$$(r, \nabla z) = (q, \nabla z) = -(\operatorname{div} q, z) = -(\beta, z)$$

Combining the above two gives

$$(as(\tau^2), z) = (\eta, z).$$

Finally, let t be the mean value of $tr(\tau^2)$ and set $\tau = \tau^2 - (t/2)I \in \hat{V}_h$. Then, $\|\tau\|_{H(\mathrm{div})} \le \|\tau^2\|_{H(\mathrm{div})} \le C(\|v\| + \|\eta\|)$, $\operatorname{div}\tau = \operatorname{div}\tau^2 = v$, and $(as(\tau), z) = (\eta, z)$. We thus have

$$(\operatorname{div}\tau, v) + (as(\tau), \eta) = \|v\|^2 + \|\eta\|^2 \ge C^{-1}\|\tau\|_{H(\mathrm{div})}(\|v\| + \|\eta\|)$$

and the proof of Lemma 4.3 has been accomplished.

Corollary 4.4. *Suppose $f \in H^1$. Then,*

$$\|\sigma - \sigma_h\|_{H(\mathrm{div})} + \|u - u_h\| + \|p - p_h\| \le Ch\|f\|_1.$$

References

1. S. Adams and B. Cockburn, A mixed finite element method for elasticity in three dimensions, J. Sci. Comput. 25 (2005), 515-521.
2. Douglas N. Arnold, Richard S. Falk and Ragnar Winther Defferential Complexes and Stability of Finite Element Methods II: The Elasticity Complex
3. Douglas N. Arnold, Franco Brezzi, and Jim Douglas, Jr., PEERS: a new mixed finite element for plane elasticity, Japan J. Appl. Math. 1 (1984), no. 2, 347-367.
4. D. N. Arnold and R. Winther, Mixed finite elements for elasticity, Numer. Math. 92 (2002), 401-419
5. S. Brenner and L. Scott, The Mathematical Theory of Finite Element Methods. Springer, 2002.
6. D. Braess, Finite Elements: Theory, fast solvers, and applications in solid mechanics. Cambridge University Press, 2007.
7. F. Brezzi and M. Fortin, Mixed and Hybrid Finite Element Methods, Springer Series in Computational Mathematics, vol. 15, Springer, New York, 1991.