# Finite Elements for Linear Elasticity

Recall that the weak formulation for linear elasticity (in "primary" or "displacement" form) is given by: find $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega) = [H^1(\Omega)]^d$ such that $\boldsymbol{u}|_{\Gamma_D} = \boldsymbol{g}$ and

$$a(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} \, dx + \int_{\Gamma_T} \boldsymbol{t} \cdot \boldsymbol{v} \, ds$$

where $a(\cdot, \cdot)$ is the symmetric bilinear form

$$a(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega \left( 2\mu \nabla^s \boldsymbol{u} : \nabla^s \boldsymbol{v} + \lambda (\nabla \cdot \boldsymbol{u})(\nabla \cdot \boldsymbol{v}) \right) \, dx \tag{1}$$

In order to show that this problem is well-posed (and to then apply the finite element/Galerkin method), we would like to show that this bilinear form is **continuous** and **coercive**. We will assume that the Lamé parameters $\mu$ and $\lambda$ are bounded above and below by constants. (Note that the case of $\lambda \to \infty$ models materials that are *nearly incompressible*. This is an important case, but it causes certain issues with discretizations, called *volumetric locking*. Developing methods that "do not lock" is important to address this problem, but it is beyond the scope of the current discussion.)

Continuity of the bilinear form is fairly straightforward. Recall that continuity is the property that

$$a(\boldsymbol{u}, \boldsymbol{v}) \lesssim \|\boldsymbol{u}\|_{\boldsymbol{H}^1} \|\boldsymbol{v}\|_{\boldsymbol{H}^1}.$$

All the terms in (1) can be written in terms of products of partial derivatives of $\boldsymbol{u}$ and $\boldsymbol{v}$. Such products of partial derivatives are bounded by the $H^1$ norms of $\boldsymbol{u}$ and $\boldsymbol{v}$ by a simple application of the Cauchy–Schwarz inequality.

Coercivity is more involved. We wish to show that

$$a(\boldsymbol{u}, \boldsymbol{u}) \gtrsim \|\boldsymbol{u}\|_{H^1}^2.$$

Since the term with $\lambda(\nabla \cdot \boldsymbol{u})^2$ is semi-definite, it will suffice to prove that

$$\int_\Omega \nabla^s \boldsymbol{u} : \nabla^s \boldsymbol{u} \, dx = \|\nabla^s \boldsymbol{u}\|_{L^2}^2 \gtrsim \|\boldsymbol{u}\|_{H^1}^2.$$

To prove this, we first need a technical lemma.

**Lemma 1.** *There exists a constant $C > 0$ such that for all $p \in L^2(\Omega)$ there exists (a non-unique) $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ such that*

$$\nabla \cdot \boldsymbol{v} = p$$

*and*

$$\|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)} \leq C \|p\|_{L^2(\Omega)}.$$

*Furthermore, if $p$ has zero mean, then we can take $\boldsymbol{v}$ with zero trace.*

*Proof.* Consider the Poisson problem

$$
\begin{aligned}
-\Delta w &= p \quad \text{in } \Omega \\
w &= 0 \quad \text{on } \partial\Omega.
\end{aligned}
$$

By elliptic regularity, the solution to this problem satisfies

$$\|w\|_{H^2(\Omega)} \le C_\Omega \|p\|_{L^2(\Omega)}.$$

Let $\boldsymbol{v} = -\nabla w$. This $\boldsymbol{v}$ satisfies the conclusions of the lemma. (Why?)

Now, we prove the additional statement when $p$ has zero mean, $\boldsymbol{v}$ can be taken to have zero trace. We provide the proof in $\mathbb{R}^2$ for simplicity. Since $p$ has zero mean, the pure Neumann problem

$$
\begin{aligned}
-\Delta w &= p \quad \text{in } \Omega \\
\frac{\partial w}{\partial \boldsymbol{n}} &= 0 \quad \text{on } \partial\Omega.
\end{aligned}
$$

is well posed. As before,

$$\|w\|_{H^2(\Omega)} \le C_\Omega \|p\|_{L^2(\Omega)}.$$

Let $\boldsymbol{v}_1 = -\nabla w$. Then, $\boldsymbol{v}_1$ satisfies

$$\nabla \cdot \boldsymbol{v}_1 = p$$

and

$$\|\boldsymbol{v}_1\|_{\boldsymbol{H}^1(\Omega)} \le C_\Omega \|p\|_{L^2(\Omega)}.$$

However, $\boldsymbol{v}_1$ does not have zero trace — the trace of its normal component is zero, but the trace of its tangential component is in general nonzero.

Let $\boldsymbol{t}$ denote the unit tangent vector. Then, regularity of the fourth-order problem implies that there exists some $\psi \in H^2(\Omega)$ such that

$$
\begin{aligned}
\psi|_{\partial\Omega} &= 0 \\
\left.\frac{\partial \psi}{\partial \boldsymbol{n}}\right|_{\partial\Omega} &= \boldsymbol{v}_1|_{\partial\Omega} \cdot \boldsymbol{t}
\end{aligned}
$$

In other words, $\psi$ has zero trace and prescribed normal derivative. $\psi$ satisfies the bound

$$\|\psi\|_{H^2(\Omega)} \le C_t \|\boldsymbol{v}_1\|_{\boldsymbol{H}^1(\Omega)}.$$

Define $\boldsymbol{v}_2$ by

$$\boldsymbol{v}_2 = \operatorname{curl}(\psi) = (\partial\psi/\partial y, -\partial\psi/\partial x)^T.$$

The **normal trace** of $\boldsymbol{v}_2$ is equal to the **tangential trace** of $\nabla\psi$, which vanishes because $\psi$ is identically zero along the boundary. The **tangential trace** of $\boldsymbol{v}_2$ is equal to the (negative) **normal trace** of $\nabla\psi$, so

$$\boldsymbol{v}_2|_{\partial\Omega} \cdot \boldsymbol{t} = -\boldsymbol{v}_1|_{\partial\Omega} \cdot \boldsymbol{t}.$$

Summarizing:

- $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ both have zero normal trace on the boundary
- $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ have equal magnitude but opposite sign tangential traces

Therefore, $\boldsymbol{v}_1|_{\partial\Omega} = -\boldsymbol{v}_2|_{\partial\Omega}$ Let $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$. It is clear that $\boldsymbol{v}$ has zero trace on the boundary. Additionally, $\nabla \cdot \boldsymbol{v}_2 = 0$, so $\nabla \cdot \boldsymbol{v} = \nabla \cdot \boldsymbol{v}_1 = p$, and

$$\begin{aligned}\|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)} &\leq \|\boldsymbol{v}_1\|_{\boldsymbol{H}^1(\Omega)} + \|\boldsymbol{v}_2\|_{\boldsymbol{H}^1(\Omega)} \\ &\leq \|\boldsymbol{v}_1\|_{\boldsymbol{H}^1(\Omega)} + \|\psi\|_{H^2(\Omega)} \\ &\leq (1 + C_t)\|\boldsymbol{v}_1\|_{\boldsymbol{H}^1(\Omega)} \\ &\leq C_\Omega(1 + C_t)\|p\|_{L^2(\Omega)}.\end{aligned}$$

□

Now, we need another technical result in the space

$$\widehat{\boldsymbol{H}}^k(\Omega) = \left\{\boldsymbol{v} \in \boldsymbol{H}^k(\Omega) : \int_\Omega \boldsymbol{v}\, dx = 0 \text{ and } \int_\Omega \operatorname{rot} \boldsymbol{v}\, dx = 0\right\}$$

where

$$\operatorname{rot} \boldsymbol{v} = -\partial v_1/\partial y + \partial v_2/\partial x$$

We collect some easy calculus definitions and results:

$$\nabla^s \boldsymbol{v} = \frac{1}{2}\left(\nabla \boldsymbol{v} + (\nabla \boldsymbol{v})^T\right) = \nabla \boldsymbol{v} - \frac{1}{2}(\operatorname{rot} \boldsymbol{v})\chi$$

where

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

$$\chi : \nabla \boldsymbol{v} = \operatorname{rot} \boldsymbol{v}$$

$$\chi : \nabla \times \boldsymbol{v} = \nabla \cdot \boldsymbol{v}$$

where

$$\nabla \times \boldsymbol{v} = \begin{pmatrix} \partial v_1/\partial y & -\partial v_1/\partial x \\ \partial v_2/\partial y & -\partial v_2/\partial x \end{pmatrix}$$

For any $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ and $\boldsymbol{w} \in \boldsymbol{H}^1_0(\Omega)$, it holds that

$$\int_\Omega \nabla \boldsymbol{v} : \nabla \times \boldsymbol{w}\, dx = 0$$

(what calculus identity is this related to?)

**Theorem 1.** *There exists a constant $C > 0$ such that*

$$\|\nabla^s \boldsymbol{v}\|_{L^2(\Omega)} \geq C\|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)}$$

*for all $\boldsymbol{v} \in \widehat{\boldsymbol{H}}^1(\Omega)$.*

*Proof.* Let $\boldsymbol{v} \in \widehat{\boldsymbol{H}}^1(\Omega)$. Since the rotation of $\boldsymbol{v}$ has zero mean, by the previous result, there exists $\boldsymbol{w}$ (in $\boldsymbol{H}^1(\Omega)$ with zero trace) such that

$$\nabla \cdot \boldsymbol{w} = \operatorname{rot} \boldsymbol{v}$$
$$\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)} \leq C_1 \|\boldsymbol{v}\|_{H^1(\Omega)}$$

Then,

$$\begin{aligned}
\int_\Omega \nabla^s \boldsymbol{v} : (\nabla \boldsymbol{v} - \nabla \times \boldsymbol{w})\, dx &= \int_\Omega \left( \nabla \boldsymbol{v} - \frac{1}{2} (\operatorname{rot} \boldsymbol{v}) \chi \right) : (\nabla \boldsymbol{v} - \nabla \times \boldsymbol{w})\, dx \\
&= \|\nabla \boldsymbol{v}\|^2_{L^2(\Omega)} - \int_\Omega \nabla \boldsymbol{v} : \nabla \times \boldsymbol{w}\, dx \\
&\qquad - \frac{1}{2} \int_\Omega (\operatorname{rot} \boldsymbol{v}) \left( \chi : \nabla \boldsymbol{v} - \chi : \nabla \times \boldsymbol{w} \right)\, dx \\
&= \|\nabla \boldsymbol{v}\|^2_{L^2(\Omega)} - \frac{1}{2} \int_\Omega (\operatorname{rot} \boldsymbol{v}) \left( \operatorname{rot} \boldsymbol{v} - \nabla \cdot w \right)\, dx \\
&= \|\nabla \boldsymbol{v}\|^2_{L^2(\Omega)}
\end{aligned}$$

The Cauchy-Schwarz inequality then implies that

$$\|\nabla \boldsymbol{v}\|^2_{L^2(\Omega)} \leq \|\nabla^s \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \boldsymbol{v} - \nabla \times \boldsymbol{w}\|_{L^2} \leq C \|\nabla^s \boldsymbol{v}\|_{L^2(\Omega)} \|\boldsymbol{v}\|_{H^1(\Omega)}.$$

Since $\boldsymbol{v}$ has zero mean, we have that

$$\|\nabla \boldsymbol{v}\|^2_{L^2(\Omega)} \geq C \|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)},$$

(why?) and the result follows. □

**Theorem 2.** *There exists a constant $\alpha > 0$ such that*

$$\|\nabla^s \boldsymbol{v}\|_{L^2(\Omega)} + \|\boldsymbol{v}\|_{L^2(\Omega)} \geq \alpha \|\boldsymbol{v}\|_{H^1(\Omega)}$$

*for all $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$.*

*Proof.* Any $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ can be written as the sum

$$\boldsymbol{v} = \boldsymbol{z} + \boldsymbol{w},$$

where $\boldsymbol{z} \in \widehat{\boldsymbol{H}}^1(\Omega)$ and $\boldsymbol{w}$ is an infinitesimal rigid motion (i.e. $\nabla^s \boldsymbol{w} = 0$, $\boldsymbol{w} = \boldsymbol{c} + b(y, -x)$).

$\boldsymbol{w}$ can take the specific form

$$b = \frac{-1}{2|\Omega|} \int_\Omega \operatorname{rot} \boldsymbol{v}\, dx$$
$$c = \frac{1}{|\Omega|} \int_\Omega (\boldsymbol{v} - b(y, -x)^T)\, dx$$

This implies that

$$\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)} \leq C \|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)}$$

and so

$$\|\boldsymbol{z}\|_{\boldsymbol{H}^1(\Omega)} \leq C\|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)}$$

also. (Why? Does a more general result hold?).

Now, we argue by contradiction. Suppose that the conclusion of the theorem does not hold. Then, there exists some sequence $\{\boldsymbol{v}_n\} \subseteq \boldsymbol{H}^1(\Omega)$ such that

$$\|\boldsymbol{v}_n\|_{\boldsymbol{H}^1(\Omega)} = 1$$

but

$$\|\nabla^s \boldsymbol{v}_n\|_{L^2(\Omega)} + \|\boldsymbol{v}_n\|_{L^2(\Omega)} < \frac{1}{n}$$

For each $n$, decompose $\boldsymbol{v}_n = \boldsymbol{z}_n + \boldsymbol{w}_n$ as above. Then,

$$\|\nabla^s \boldsymbol{z}_n\|_{L^2(\Omega)} = \|\nabla^s \boldsymbol{v}_n\|_{L^2(\Omega)} < \frac{1}{n}.$$

By the previous theorem, $\boldsymbol{z}_n \to 0$ in $\boldsymbol{H}^1(\Omega)$.

So, both $\boldsymbol{v}_n$ and $\boldsymbol{z}_n$ are bounded in $\boldsymbol{H}^1(\Omega)$, and so $\boldsymbol{w}_n$ is a bounded sequence with a convergent subsequence (the sequence $\boldsymbol{w}_n$ lies in the space of infinitesimal rigid body motions, which is a three-dimensional space, and so the Bolzano–Weierstrass theorem applies). Since $\boldsymbol{z}_n \to 0$, the corresponding subsequence of $\boldsymbol{v}_n$ is also convergent, and it must converge to a infinitesimal rigid motion $\boldsymbol{v}$ with

$$\|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)} = 1 \quad \text{and} \quad \|\boldsymbol{v}\|_{L^2(\Omega)} = 0$$

which is a contradiction (why?). □

**Corollary 1.** *If the problem is* ***not*** *pure traction (i.e. $|\Gamma_D| > 0$) then*

$$\|\nabla^s \boldsymbol{v}\|_{L^2(\Omega)} \geq C\|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)}.$$

*Proof.* The same proof implies that there is a rigid motion $\boldsymbol{v}$ in $\boldsymbol{V}$ such that

$$\|\boldsymbol{v}\|_{H^1(\Omega)} = 1.$$

But the only rigid motion satisfying the homogeneous displacement boundary conditions is identically zero. □

**Theorem 3.** *Excluding the case of pure traction boundary conditions, the linear elasticity variational problem has a unique solution.*

**Theorem 4.** *Let $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and $\boldsymbol{t} \in \boldsymbol{L}^2(\Gamma_T)$. Then, the variational problem: find $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ such that, for all $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$,*

$$a(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) + \int_{\Gamma_T} \boldsymbol{t} \cdot \boldsymbol{v}\, ds$$

*is solvable if and only if*

$$\int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\, dx + \int_{\Gamma_T} \boldsymbol{t} \cdot \boldsymbol{v}\, ds = 0$$

*for all $\boldsymbol{v} \in \mathbf{RM}$. When the equation is solvable, there is a unique solution in $\widehat{\boldsymbol{H}}^1(\Omega)$.*

*Proof.* If the variational formulation is solvable, then taking the test function to be a rigid motion proves that the compatibility condition is necessary.

We now prove that the condition is sufficient. By Theorem 1, the bilinear form is coercive on $\widehat{\boldsymbol{H}}^1(\Omega)$, and so a unique solution $\boldsymbol{u}^*$ exists for all test functions $\boldsymbol{v} \in \widehat{\boldsymbol{H}}^1(\Omega)$. Now assume that the compatibility condition holds. Then, the $\boldsymbol{u}^*$ is also a solution for any test function $\boldsymbol{v} \in \mathbf{RM}$. Since any $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ can be (uniquely) written as a sum of $\boldsymbol{z} \in \widehat{\boldsymbol{H}}^1(\Omega)$ and $\boldsymbol{w} \in \mathbf{RM}$, we see that $\boldsymbol{u}^*$ is a solution of the variational problem. □

The above results show that the problem is well-posed, and so a unique solution exists, and stability holds with respect to the given right-hand side. We can apply the standard Galerkin/finite element methodology, by choosing finite dimensional subspaces $\boldsymbol{V}_h \subseteq \boldsymbol{H}^1(\Omega)$. For example, we can choose $\boldsymbol{V}_h = [V_h]^d$; this is simply the vector-valued version of the standard finite element space $V_h$. For $\boldsymbol{v}_h \in \boldsymbol{V}_h$, $\boldsymbol{v}_h = (v_1, \ldots, v_d)$, and $v_i \in V_h$. The solution will be optimal in the energy norm, and using the results above, we can prove error estimates of the form

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{H^1(\Omega)} \leq Ch\|\boldsymbol{u}\|_{H^2(\Omega)}$$

where $\boldsymbol{V}_h$ is constructed using linear finite elements. Note that the constant $C$ **depends on the Lamé parameters** $\mu$ **and** $\lambda$. If $\lambda$ is very large, then $C$ will be large; no matter how small $h$ is (how fine the mesh is), there is some $\lambda$ for which the error will be large. This is the problem of volumetric locking.