VARIATIONAL FORMULATION OF LINEAR ELASTICITY

LONG CHEN

ABSTRACT. This notes presents two variational formulations of linear elasticity: the displacement formulation and the stress-displacement (Hellinger-Reissner) formulation. The displacement formulation focuses on finding the displacement field that minimizes a functional, while the stress-displacement formulation directly computes the stresses. The well-posedness of these formulations relies on Korn's inequalities, inf-sup conditions, and the coercivity of bilinear forms.

We explore variational formulations of linear elasticity. For a comprehensive understanding of linear elasticity, we refer to Introduction to Linear Elasticity. Additionally, we recommend familiarizing oneself with tensor calculus notation and operation rules by reading Tensor Calculus.

1. DISPLACEMENT FORMULATION

Define the Lagrangian

$$\mathcal{I}(\boldsymbol{u}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}) = \int_\Omega \left(\frac{1}{2}\boldsymbol{\varepsilon} : \boldsymbol{\sigma} - \boldsymbol{f} \cdot \boldsymbol{u} \right) \mathrm{d}\boldsymbol{x} + \int_{\Gamma_N} \boldsymbol{t}_N \cdot \boldsymbol{u} \, \mathrm{d}\boldsymbol{x}.$$

Here, the relationships are defined as

$$\begin{aligned} \boldsymbol{\varepsilon} &= \nabla^s \boldsymbol{u}, \\ \boldsymbol{\sigma} &= \mathcal{C}\boldsymbol{\varepsilon} = \lambda \operatorname{tr}(\boldsymbol{\varepsilon}) \boldsymbol{I} + 2\mu\boldsymbol{\varepsilon}, \end{aligned}$$

accompanied by the boundary conditions

$$\boldsymbol{u}|_{\Gamma_D} = \boldsymbol{g}, \quad \boldsymbol{\sigma n}|_{\Gamma_N} = \boldsymbol{t}_N.$$

Here, Γ_N represents an open subset of the boundary $\partial\Omega$, and its complement is denoted by Γ_D, meaning $\Gamma_D \cup \Gamma_N = \partial\Omega$ and Γ_D is closed.

The Lagrangian combines the internal strain energy of the system, represented by the product $\boldsymbol{\varepsilon} : \boldsymbol{\sigma}$, and the work done by external forces, $\boldsymbol{f} \cdot \boldsymbol{u}$, along with the work done by traction forces $\boldsymbol{t}_N \cdot \boldsymbol{u}$ on the Neumann boundary Γ_N.

1.1. **Displacement Formulation.**

We focus on the displacement formulation by eliminating $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$, leading to the optimization problem:

$$\inf_{\boldsymbol{u} \in \boldsymbol{H}^1_{g,D}} \mathcal{I}(\boldsymbol{u}), \tag{1}$$

Date: March 23, 2024.

where $\boldsymbol{H}^1_{g,D} = \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) : \boldsymbol{v} \mid_{\Gamma_D} = \boldsymbol{g}$ in the trace sense $\}$, and

$$\begin{aligned}\mathcal{I}(\boldsymbol{u}) &= \int_\Omega \left(\frac{1}{2}\nabla^s \boldsymbol{u} : \mathcal{C}\nabla^s \boldsymbol{u} - \boldsymbol{f}\cdot\boldsymbol{u}\right) \mathrm{d}\boldsymbol{x} + \int_{\Gamma_N} \boldsymbol{t}_N \cdot \boldsymbol{u} \,\mathrm{d}\boldsymbol{x}\\ &= \int_\Omega \left(\frac{\lambda}{2}(\operatorname{div}\boldsymbol{u})^2 + \mu\nabla^s \boldsymbol{u} : \nabla^s \boldsymbol{u} - \boldsymbol{f}\cdot\boldsymbol{u}\right) \mathrm{d}\boldsymbol{x} + \int_{\Gamma_N} \boldsymbol{t}_N \cdot \boldsymbol{u} \,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

The strong formulation of the Euler-Lagrange equation is

$$\begin{aligned}-2\mu \operatorname{div}\nabla^s \boldsymbol{u} - \lambda \operatorname{grad}\operatorname{div}\boldsymbol{u} &= \boldsymbol{f} \quad \text{in } \Omega,\\ \boldsymbol{u} &= \boldsymbol{g} \quad \text{on } \Gamma_D,\\ \boldsymbol{\sigma}(\boldsymbol{u})\cdot\boldsymbol{n} &= \boldsymbol{t}_N \quad \text{on } \Gamma_N.\end{aligned}$$

The weak formulation seeks $\boldsymbol{u} \in \boldsymbol{H}^1_{g,D}(\Omega)$ such that

$$(\mathcal{C}\nabla^s \boldsymbol{u}, \nabla^s \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) + \langle \boldsymbol{t}_N, \boldsymbol{v}\rangle_{\Gamma_N} \quad \text{for all } \boldsymbol{v} \in \boldsymbol{H}^1_{0,D}(\Omega). \tag{2}$$

This formulation seeks the displacement field $\boldsymbol{u}$ that minimizes the functional $\mathcal{I}(\boldsymbol{u})$, subject to given boundary conditions and force distributions.

1.2. **Korn Inequalities.** To ensure the well-posedness of the weak formulation, Korn's inequality is crucial:

$$\|D\boldsymbol{u}\| \leq C\|\nabla^s \boldsymbol{u}\|, \quad \forall \boldsymbol{u} \in \boldsymbol{H}^1_{0,D}(\Omega). \tag{3}$$

Since

$$\|D\boldsymbol{u}\|^2 = \|\operatorname{sym}(D\boldsymbol{u})\|^2 + \|\operatorname{skw}(D\boldsymbol{u})\|^2 = \|\nabla^s \boldsymbol{u}\|^2 + 2\|\nabla\times\boldsymbol{u}\|^2,$$

Korn's inequality is not straightforward because it suggests that $\|\nabla\times\boldsymbol{u}\|$ should be bounded by $\|\nabla^s \boldsymbol{u}\|$. Moreover, without certain conditions on $\boldsymbol{u}$, like $\boldsymbol{u} \in \ker(\nabla^s)$ but $\boldsymbol{u} \notin \ker(D)$, the right side becomes zero while the left does not. In the given inequality, the nonzero Lebesgue measure $|\Gamma_D| \neq 0$ ensures that $\boldsymbol{H}^1_{0,D}(\Omega) \cap \ker(\nabla^s) = \varnothing$.

For the special case where $\Gamma_D = \partial\Omega$, we have $\boldsymbol{u} \in \boldsymbol{H}^1_0(\Omega)$, allowing integration by parts without boundary terms. We have the identity

$$2\operatorname{div}\nabla^s \boldsymbol{u} = \Delta\boldsymbol{u} + \operatorname{grad}\operatorname{div}\boldsymbol{u}, \tag{4}$$

which can be proved as follows: for $k = 1, 2, \ldots, d$,

$$(\operatorname{div} 2\nabla^s \boldsymbol{u})_k = \sum_{i=1}^d \partial_i(\partial_i u_k + \partial_k u_i) = \Delta\boldsymbol{u}_k + \partial_k(\operatorname{div}\boldsymbol{u}).$$

Multiplying (4) by $\boldsymbol{u}$ and integrating by parts to get the identity

$$2\|\nabla^s \boldsymbol{u}\|^2 = \|D\boldsymbol{u}\|^2 + \|\operatorname{div}\boldsymbol{u}\|^2,$$

which leads to the first Korn inequality.

Lemma 1.1 (First Korn Inequality)**.**

$$\|D\boldsymbol{u}\| \leq \sqrt{2}\|\nabla^s \boldsymbol{u}\|, \quad \boldsymbol{u} \in \boldsymbol{H}^1_0(\Omega). \tag{5}$$

We now present the second Korn's inequality which necessitates conditions to exclude $\ker(\nabla^s)$. Recall the characterization of the kernel of the symmetric gradient:

$$\ker(\nabla^s) = \{\boldsymbol{\omega}\times\boldsymbol{x} + \boldsymbol{c},\ \boldsymbol{\omega}, \boldsymbol{c} \in \mathbb{R}^3\}.$$

The constant vector $\boldsymbol{c} \in \ker(\nabla^s)$ can be excluded by imposing the condition $\int_\Omega \boldsymbol{u} = 0$. When $\boldsymbol{u} = \boldsymbol{\omega} \times \boldsymbol{x} = [\boldsymbol{\omega}]_\times \boldsymbol{x}$, it follows that $D\boldsymbol{u} = [\boldsymbol{\omega}]_\times$ and

$$[\nabla \times \boldsymbol{u}]_\times = \operatorname{skw}(D\boldsymbol{u}) = [\boldsymbol{\omega}]_\times.$$

Hence, $\boldsymbol{\omega} = \nabla \times \boldsymbol{u}$ if $\boldsymbol{u} = \boldsymbol{\omega} \times \boldsymbol{x} \in \ker(\nabla^s)$, a condition that can be removed by requiring $\int_\Omega \boldsymbol{\omega} = 0$. Therefore, we define the subspace

$$\widehat{\boldsymbol{H}}^1(\Omega) := \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) : \int_\Omega \boldsymbol{v} \,\mathrm{d}\boldsymbol{x} = \int_\Omega \nabla \times \boldsymbol{v} \,\mathrm{d}\boldsymbol{x} = 0\}. \tag{6}$$

This helps in excluding the rigid body motions from the consideration by ensuring that both the constant translation and constant rotation components are nullified by the integral conditions.

Lemma 1.2 (The Second Korn's Inequality)**.** *There exists a constant dependent only on the geometry of the domain* Ω *such that*

$$\|D\boldsymbol{u}\| \le C\|\nabla^s \boldsymbol{u}\|, \quad \forall \boldsymbol{u} \in \widehat{\boldsymbol{H}}^1(\Omega). \tag{7}$$

Proof. Given $\boldsymbol{u} \in \widehat{\boldsymbol{H}}^1(\Omega)$, let $\boldsymbol{q} = \nabla \times \boldsymbol{u} \in L_0^2(\Omega)$. There exists a symmetric matrix function $\Phi \in \boldsymbol{H}_0^1(\Omega; \mathbb{S})$ satisfying

$$\operatorname{div} \Phi = \boldsymbol{q} = \nabla \times \boldsymbol{u}, \quad \|\Phi\|_1 \le C\|\boldsymbol{q}\|.$$

This result, established in the context of Stokes' equation, extends to symmetric tensors.

We claim

$$(D\boldsymbol{u}, D\boldsymbol{u}) = (D\boldsymbol{u}, D\boldsymbol{u} - \operatorname{curl} \Phi) = (\nabla^s \boldsymbol{u}, D\boldsymbol{u} - \operatorname{curl} \Phi). \tag{8}$$

Applying the Cauchy-Schwarz inequality,

$$\|D\boldsymbol{u}\|^2 = (\nabla^s \boldsymbol{u}, D\boldsymbol{u} - \operatorname{curl} \Phi) \lesssim \|\nabla^s \boldsymbol{u}\|(\|D\boldsymbol{u}\| + \|\Phi\|_1) \lesssim \|\nabla^s \boldsymbol{u}\|\|D\boldsymbol{u}\|.$$

Eliminating one $\|D\boldsymbol{u}\|$ leads to the desired Korn inequality.

Now we justify steps in (8). The orthogonality $(D\boldsymbol{u}, \operatorname{curl} \Phi) = (\operatorname{curl} D\boldsymbol{u}, \Phi) = 0$ is verified via integration by parts, with all differential operators applied row-wise and without boundary terms since $\Phi \in \boldsymbol{H}_0^1(\Omega; \mathbb{S})$.

Recognizing $2\operatorname{skw}(D\boldsymbol{u}) = D\boldsymbol{u} - (D\boldsymbol{u})^\intercal$ and $(\operatorname{curl} \Phi, D\boldsymbol{u}) = 0$, the operations yield

$$\begin{aligned}(\operatorname{curl} \Phi, 2\operatorname{skw}(D\boldsymbol{u})) &= -(\Phi \times \nabla, \nabla \boldsymbol{u}^\intercal) = (\Phi, \nabla \boldsymbol{u}^\intercal \times \nabla) \\ &= -(\nabla \cdot \Phi, \boldsymbol{u}^\intercal \times \nabla) = (\operatorname{div} \Phi, \nabla \times \boldsymbol{u}) \\ &= (\nabla \times \boldsymbol{u}, \nabla \times \boldsymbol{u}) = (D\boldsymbol{u}, [\nabla \times \boldsymbol{u}]_\times).\end{aligned}$$

That is $(D\boldsymbol{u} - \operatorname{curl} \Phi, \operatorname{skw}(D\boldsymbol{u})) = 0$ and only symmetric gradient $\nabla^s \boldsymbol{u}$ is left.

The relationship among $\boldsymbol{u}$, $D\boldsymbol{u}$, $\operatorname{curl} \Phi$, and Φ follows the diagram:

$$\boldsymbol{u} \xrightarrow{D} D\boldsymbol{u} \oplus^\perp \operatorname{curl} \Phi \xleftarrow{\operatorname{curl}} \Phi \xrightarrow{\operatorname{div}} \nabla \times \boldsymbol{u} \to [\nabla \times \boldsymbol{u}]_\times + \nabla^s \boldsymbol{u} = D\boldsymbol{u}.$$

□

This lemma elucidates the complexities involved in tensor calculations but simplifies understanding through the introduced notation system in Tensor Calculus. The lemma adapts the 2D proof strategy in [2, Ch 11] to the three-dimensional setting.

For the general case on Lipschitz domains, we rely on the norm equivalence:

$$\|\boldsymbol{v}\|^2 \approx \|\nabla \boldsymbol{v}\|_{-1}^2 + \|\boldsymbol{v}\|_{-1}^2 \quad \text{for all } \boldsymbol{v} \in \boldsymbol{L}^2(\Omega). \tag{9}$$

Lemma 1.3 (Lion's Lemma)**.** *For Lipschitz domains, the space* $X(\Omega) = L^2(\Omega)$*, where* $X(\Omega) = \{v \mid v \in H^{-1}(\Omega), \operatorname{grad} v \in (H^{-1}(\Omega))^n\}$ *with norm* $\|v\|_X^2 = \|v\|_{-1}^2 + \|\operatorname{grad} v\|_{-1}^2$.

Proof. A proof $\|v\|_X \lesssim \|v\|$, consequently $L^2(\Omega) \subseteq X(\Omega)$, is trivial (using the definition of the dual norm). The non-trivial part is to prove the inequality

$$\|v\|^2 \lesssim \|v\|_{-1}^2 + \|\operatorname{grad} v\|_{-1}^2 = \|v\|_{-1}^2 + \sum_{i=1}^{d} \|\partial_i v\|_{-1}^2. \tag{10}$$

The difficulty is associated to the non-computable dual norm $\|\cdot\|_{-1}$. We only present a special case $\Omega = \mathbb{R}^n$ by the characterization of H^{-1} norm using Fourier transform. Let $\hat{u}(\xi) = \mathscr{F}(u)$ be the Fourier transform of u. Then

$$\|u\|_{\mathbb{R}^n}^2 = \|\hat{u}\|_{\mathbb{R}^n}^2 = \left\|1/(\sqrt{1+|\xi|^2})\hat{u}\right\|_{\mathbb{R}^n}^2 + \sum_{i=1}^{d} \left\|\xi_i/(\sqrt{1+|\xi|^2})\hat{u}\right\|_{\mathbb{R}^n}^2 = \|u\|_X^2.$$

In general cases, careful extension from $H^{-1}(\Omega)$ to $H^{-1}(\mathbb{R}^d)$ is needed; see, e.g. [1]. □

The following identity for C^2 function can be easily verified by definition of symmetric graident

$$\partial_{ij}^2 u_k = \partial_j \varepsilon_{ki}(\boldsymbol{u}) + \partial_i \varepsilon_{jk}(\boldsymbol{u}) - \partial_k \varepsilon_{ij}(\boldsymbol{u}). \tag{11}$$

We now use Lemma 1.3 and identity (11) to prove the following Korn's inequality.

Theorem 1.4 (Korn's inequality with L^2-norm)**.** *There exists a constant depending only on the geometry of domain* Ω *s. t.*

$$\|D\boldsymbol{u}\| \leq C\left(\|\nabla^s \boldsymbol{u}\| + \|\boldsymbol{u}\|\right), \quad \forall \boldsymbol{u} \in \boldsymbol{H}^1(\Omega). \tag{12}$$

Proof. By the norm equivalence and using the identity to switch derivatives (11)

$$\|\partial_i \boldsymbol{u}\| \lesssim \|\partial_i \boldsymbol{u}\|_{-1} + \|\nabla \partial_i \boldsymbol{u}\|_{-1} \lesssim \|\boldsymbol{u}\| + \sum_j \|\partial_j \nabla^s \boldsymbol{u}\|_{-1} \lesssim \|\boldsymbol{u}\| + \|\nabla^s \boldsymbol{u}\|.$$

□

We are ready to prove the coercivity of the displacement formulation (2).

Lemma 1.5 (Korn's inequality for non-trivial zero trace)**.** *Assume* $|\Gamma_D| \neq 0$*. There exists a constant depending only on the geometry of domain* Ω *s. t.*

$$\|D\boldsymbol{u}\| \leq C\|\nabla^s \boldsymbol{u}\|, \quad \forall \boldsymbol{u} \in \boldsymbol{H}_{0,D}^1(\Omega). \tag{13}$$

Proof. Assuming the contrary, that no such constant exists, leads to finding a sequence $\{\boldsymbol{u}_k\} \subset \boldsymbol{H}^1(\Omega)$ such that

$$\|D\boldsymbol{u}_k\| = 1, \quad \text{and } \|\nabla^s \boldsymbol{u}_k\| \to 0 \text{ as } k \to \infty.$$

Since $\boldsymbol{H}^1(\Omega)$ compactly embeds into $\boldsymbol{L}^2(\Omega)$, a convergent subsequence exists, $\boldsymbol{u}_k \to \boldsymbol{u}$ in $\boldsymbol{L}^2(\Omega)$. Korn's inequality implies $\{\boldsymbol{u}_k\}$ is also a Cauchy sequence in $\boldsymbol{H}^1(\Omega)$, hence $\boldsymbol{u}_k \to \boldsymbol{u}$ in $\boldsymbol{H}^1(\Omega)$. Consequently, $\|\nabla^s \boldsymbol{u}\| = 0$.

So $\boldsymbol{u} = \boldsymbol{\omega} \times \boldsymbol{x} + \boldsymbol{c}$. The condition $\boldsymbol{u}|_{\Gamma_D} = 0$ implies $\boldsymbol{u} = 0$. Contradicts with the condition $\|D\boldsymbol{u}\| = 1$. □

A set of functional L, consisting of components $l_i(\cdot)$ for $i = 1, 2, \cdots, 6$, is introduced, requiring $\ker(L) \cap \ker(\nabla^s) = \{\boldsymbol{0}\}$. This means if $\boldsymbol{u} \in \ker(\nabla^s)$ and $l_i(\boldsymbol{u}) = 0$ for all i, then $\boldsymbol{u} = 0$. Modifying the proof of Lemma 1.5 leads to the following version of Korn's inequality [3].

Lemma 1.6 (Korn's inequality with functionals)**.** *Let* $L = (l_1, \dots, l_6)$*. Assuming* $\ker(L) \cap \ker(\nabla^s) = \{\mathbf{0}\}$*, there exists a constant dependent only on the domain's geometry* Ω*, such that*

$$\|D\boldsymbol{u}\| \leq C\|\nabla^s \boldsymbol{u}\| + \sum_{i=1}^{6} |l_i(\boldsymbol{u})|, \quad \forall \boldsymbol{u} \in \boldsymbol{H}^1(\Omega). \tag{14}$$

Proof. Following the proof for non-trivial zero trace, we conclude there exists $\boldsymbol{u} \in \ker(\nabla^s)$ with $l_i(\boldsymbol{u}) = 0$ for $i = 1, 2, \dots, 6$, and $\|D\boldsymbol{u}\| = 1$. This implies $\boldsymbol{u} = 0$, contradicting the fact $\|D\boldsymbol{u}\| = 1$. □

One specific example of the functionals are $\int_\Omega \boldsymbol{v}\, \mathrm{d}\boldsymbol{x}, \int_\Omega \nabla \times \boldsymbol{v}\, \mathrm{d}\boldsymbol{x}$. In contra to the constructive proof in Lemma 1.2, here it is using the abstract compactness result.

1.3. **Pure Traction Boundary Condition.** For cases similar to the pure Neumann boundary condition in the Poisson equation, where $\Gamma_D = \varnothing$ and $\Gamma_N = \partial\Omega$, the differential operator exhibits a non-trivial kernel $\ker(\nabla^s)$. By selecting $\boldsymbol{v} \in \ker(\nabla^s) \cap \boldsymbol{H}^1$ in (2), we derive the compatibility condition for the force $\boldsymbol{f}$ and the boundary force $\boldsymbol{t}_N$:

$$\int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\, \mathrm{d}\boldsymbol{x} + \int_{\partial\Omega} \boldsymbol{t}_N \cdot \boldsymbol{v}\, \mathrm{d}S = 0 \quad \forall \boldsymbol{v} \in \ker(\nabla^s). \tag{15}$$

The kernel $\ker(\nabla^s)$ can be identified by considering the dual space of $\ker(\nabla^s)$. Specifically, when the compatibility condition (15) is met, a unique solution can be found in the space $\widehat{\boldsymbol{H}}^1(\Omega)$ with constraints, as defined in (6).

1.4. **Robustness.** In scenarios where $\lambda \gg 1$ and $\mu \sim 1$, the operator $-\lambda\,\mathrm{grad}\,\mathrm{div}$ will predominate in the displacement formulation, leading to a singularly perturbed operator

$$-\mathrm{grad}\ \mathrm{div} - \epsilon\, \mathrm{div}\, \nabla^s, \quad \epsilon = 2\mu/\lambda \ll 1,$$

with the degenerate case of $\epsilon = 0$, as $\mathrm{img}(\mathrm{curl}\,) \subseteq \ker(\mathrm{div})$. This condition poses challenges for both finite element discretizations and multigrid solvers.

When λ is significantly larger, indicating the material is nearly incompressible, the quantity $\mathrm{div}\,\boldsymbol{u}$ measures the material's incompressibility. For nearly incompressible materials ($\lambda \gg 1$), $\mathrm{div}\,\boldsymbol{u}$ should also be small. In smooth or convex domains, uniform regularity results hold:

$$\|\boldsymbol{u}\|_2 + \lambda\|\,\mathrm{div}\,\boldsymbol{u}\|_1 \leq C\|\boldsymbol{f}\|. \tag{16}$$

Robust numerical methods can be designed for the displacement-pressure formulation by introducing an artificial pressure $p = \lambda\,\mathrm{div}\,\boldsymbol{u}$. See Finite Element Methods for Linear Elasticity.

2. Stress-displacement Formulation

Consider the Sobolev space

$$\boldsymbol{H}(\mathrm{div}, \Omega; \mathbb{S}) := \{\boldsymbol{\tau} \in L^2(\Omega; \mathbb{S}) : \mathrm{div}\,\boldsymbol{\tau} \in L^2(\Omega)\},$$

endowed with the norm

$$\|\boldsymbol{\tau}\|_{\mathrm{div}} = \left(\|\boldsymbol{\tau}\|^2 + \|\,\mathrm{div}\,\boldsymbol{\tau}\|^2\right)^{1/2}.$$

Here, $\mathbb{S}$ denotes the space of symmetric tensors, which is relevant for stress tensors that are inherently symmetric due to physical laws of equilibrium.

2.1. **Hellinger-Reissner formulation.** A variational form of linear elasticity, known as the Hellinger-Reissner formulation, that focuses on computing the stresses directly. Define $\boldsymbol{H}_{t,N}(\mathrm{div},\Omega,\mathbb{S})$ as the set of symmetric tensor fields that agree with the prescribed traction on the Neumann boundary, Γ_N:

$$\boldsymbol{H}_{t,N}(\mathrm{div},\Omega,\mathbb{S}) = \{\boldsymbol{\tau}\in\boldsymbol{H}(\mathrm{div},\Omega,\mathbb{S}), \boldsymbol{\tau n} = \boldsymbol{t}_N \text{ on } \Gamma_N\}.$$

Consider the optimization problem:

$$\inf_{\boldsymbol{\sigma}\in\boldsymbol{H}_{t,N}(\mathrm{div},\Omega,\mathbb{S})} \frac{1}{2}(\mathcal{A}\boldsymbol{\sigma},\boldsymbol{\sigma}),$$

subject to the constraint:

$$-\operatorname{div}\boldsymbol{\sigma} = \boldsymbol{f} \text{ in } \Omega, \quad \boldsymbol{\sigma n} = \boldsymbol{t}_N \text{ on } \Gamma_N.$$

Here $\mathcal{A}\boldsymbol{\sigma} = \frac{1}{2\mu}\left(\boldsymbol{\sigma} - \frac{\lambda}{2\mu+d\lambda}\mathrm{tr}(\boldsymbol{\sigma})I\right)$. The condition $-\operatorname{div}\boldsymbol{\sigma} = \boldsymbol{f}$ ensures equilibrium with external forces $\boldsymbol{f}$, and $\boldsymbol{\sigma n} = \boldsymbol{t}_N$ represents traction forces on Γ_N.

In this formulation, the displacement field, although not directly solved for in the optimization problem, can be interpreted mathematically as a Lagrange multiplier to impose the equilibrium constraint: adding $(\operatorname{div}\boldsymbol{\sigma} - \boldsymbol{f}, u)$ into the Lagrangian and consider the saddle point problem:

$$\inf_{\boldsymbol{\sigma}\in\boldsymbol{H}_{t,N}(\mathrm{div},\Omega,\mathbb{S})} \sup_{u\in\boldsymbol{L}^2(\Omega)} \frac{1}{2}(\mathcal{A}\boldsymbol{\sigma},\boldsymbol{\sigma}) + (\operatorname{div}\boldsymbol{\sigma} - \boldsymbol{f}, u),$$

The strong formulation of linear elasticity problem, integrating both equilibrium and material constitutive relations, is given as follows:

$$\begin{aligned} \mathcal{A}\boldsymbol{\sigma} &= \nabla^s\boldsymbol{u} \quad \text{in } \Omega, & \boldsymbol{u} &= 0 \quad \text{on } \Gamma_D,\\ -\operatorname{div}\boldsymbol{\sigma} &= \boldsymbol{f} \quad \text{in } \Omega, & \boldsymbol{\sigma}\cdot\boldsymbol{n} &= \boldsymbol{t}_N \quad \text{on } \Gamma_N. \end{aligned}$$

The weak or variational formulation seeks $\boldsymbol{\sigma}\in\boldsymbol{H}_{t,N}(\mathrm{div},\Omega,\mathbb{S})$ and $\boldsymbol{u}\in L^2(\Omega)$ such that:

$$(\mathcal{A}\boldsymbol{\sigma},\boldsymbol{\tau}) + (\boldsymbol{u},\operatorname{div}\boldsymbol{\tau}) = 0 \quad \forall\boldsymbol{\tau}\in\boldsymbol{H}_{0,N}(\mathrm{div},\Omega,\mathbb{S}), \tag{17}$$

$$(\operatorname{div}\boldsymbol{\sigma},\boldsymbol{v}) = -(\boldsymbol{f},\boldsymbol{v}) \quad \forall\boldsymbol{v}\in L^2(\Omega). \tag{18}$$

In scenarios where $\Gamma_N = \varnothing$, meaning the whole boundary is subject to Dirichlet conditions ($\boldsymbol{u} = 0$ on $\partial\Omega$), we encounter a mixed formulation akin to a pure Neumann problem where solutions are not uniquely determined. To address this, we consider the quotient space:

$$\widehat{\boldsymbol{H}}(\mathrm{div},\Omega,\mathbb{S}) = \{\boldsymbol{\tau}\in\boldsymbol{H}(\mathrm{div},\Omega,\mathbb{S}) : \int_\Omega \mathrm{tr}(\boldsymbol{\tau})\,\mathrm{d}\boldsymbol{x} = 0\}.$$

The constraint comes from by choosing $\boldsymbol{\tau} = \boldsymbol{I}$ in (17).

2.2. **Inf-sup Conditions.** In the context of mixed finite element methods for linear elasticity, the inf-sup condition (also known as the Ladyzhenskaya-Babuška-Brezzi (LBB) condition) is crucial for establishing the well-posedness of the problem and ensuring the stability of the solution. Let us explore the framework for the stress space $\boldsymbol{\Sigma}$ and the displacement space $\boldsymbol{V}$, each equipped with their respective norms.

Introducing the linear operator $\mathcal{L}:\boldsymbol{\Sigma}\times\boldsymbol{V}\to(\boldsymbol{\Sigma}\times\boldsymbol{V})^*$ as

$$\langle\mathcal{L}(\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})\rangle := (\mathcal{A}\boldsymbol{\sigma},\boldsymbol{\tau}) - (\nabla^s\boldsymbol{u},\boldsymbol{\tau}) + (\operatorname{div}\boldsymbol{\sigma},\boldsymbol{v}),$$

we define the bilinear forms as follows:

$$a(\boldsymbol{\sigma},\boldsymbol{\tau}) := (\mathcal{A}\boldsymbol{\sigma},\boldsymbol{\tau}),$$
$$b(\boldsymbol{\tau},\boldsymbol{v}) := -(\operatorname{div}\boldsymbol{\tau},\boldsymbol{v}).$$

The isomorphism of $\mathcal{L}$ from $\boldsymbol{\Sigma}\times\boldsymbol{V}$ onto $(\boldsymbol{\Sigma}\times\boldsymbol{V})^*$ hinges on satisfying the following Brezzi conditions:

Continuity of Bilinear Forms: There exist constants $c_a, c_b > 0$ such that for all $\boldsymbol{\sigma},\boldsymbol{\tau}\in\boldsymbol{\Sigma}$ and $\boldsymbol{v}\in\boldsymbol{V}$,

$$a(\boldsymbol{\sigma},\boldsymbol{\tau}) \le c_a\|\boldsymbol{\sigma}\|_{\boldsymbol{\Sigma}}\|\tau\|_{\boldsymbol{\Sigma}},\quad b(\boldsymbol{\tau},\boldsymbol{v}) \le c_b\|\boldsymbol{\tau}\|_{\boldsymbol{\Sigma}}\|\boldsymbol{v}\|_{\boldsymbol{V}}.$$

Coercivity of $a(\cdot,\cdot)$ in the Kernel Space: There exists a constant $\alpha>0$ such that

$$a(\boldsymbol{\sigma},\boldsymbol{\sigma}) \ge \alpha\|\boldsymbol{\sigma}\|_{\boldsymbol{\Sigma}}^2 \quad \text{for all } \boldsymbol{\sigma}\in\ker(B),$$

where $\ker(B) = \{\boldsymbol{\tau}\in\boldsymbol{\Sigma} : b(\boldsymbol{\tau},\boldsymbol{v}) = 0 \text{ for all } \boldsymbol{v}\in\boldsymbol{V}\}$.

inf-sup Condition of $b(\cdot,\cdot)$: There exists a constant $\beta>0$ such that

$$\inf_{\boldsymbol{v}\in\boldsymbol{V},\boldsymbol{v}\neq 0}\ \sup_{\boldsymbol{\tau}\in\boldsymbol{\Sigma},\boldsymbol{\tau}\neq 0}\frac{b(\boldsymbol{\tau},\boldsymbol{v})}{\|\boldsymbol{\tau}\|_{\boldsymbol{\Sigma}}\|\boldsymbol{v}\|_{\boldsymbol{V}}} \ge \beta.$$

The continuity condition typically follows from the choice of appropriate norms. The interplay between the coercivity of $a(\cdot,\cdot)$ and the inf-sup condition for $b(\cdot,\cdot)$ highlights the intricate balance necessary for the mixed formulation to be well-posed.

The inf-sup condition for $b(\cdot,\cdot)$ can be straightforwardly confirmed. Given $\boldsymbol{v}\in\boldsymbol{L}^2(\Omega)$, we approach this through a simplified displacement scenario: seeking a $\phi\in\boldsymbol{H}_0^1$ such that

$$(\nabla^s\boldsymbol{\phi},\nabla^s\boldsymbol{\psi}) = (\boldsymbol{v},\boldsymbol{\phi}),\quad \forall\boldsymbol{\psi}\in\boldsymbol{H}_0^1.$$

Thanks to the first Korn inequality, we ascertain that $\boldsymbol{\phi}$ is both existent and unique. Setting $\boldsymbol{\tau} = \nabla^s\boldsymbol{\phi}$, we discover $-\operatorname{div}\boldsymbol{\tau} = \boldsymbol{v}$ and $\|\boldsymbol{\tau}\|_{\mathrm{div}} \lesssim \|\boldsymbol{v}\|$. With this specific $\boldsymbol{\tau}$, we successfully verify the inf-sup condition for $b(\cdot,\cdot)$.

2.3. **A non-robust coercivity.** The coercivity of $a(\cdot,\cdot)$ in the L^2-norm requires careful consideration. The L^2-inner product for two tensor functions integrates two inner product structures: the Frobenius inner product $(\cdot,\cdot)_F$ among matrices, and the L^2-inner product of functions, specifically $\int_\Omega fg\,\mathrm{d}\boldsymbol{x}$.

Let $\mathbb{M}$ be the linear space of $d\times d$ matrices. The subspace of all traceless matrices is denoted by $\mathbb{T}$. The Frobenius inner product in $\mathbb{M}$ as

$$(\boldsymbol{A},\boldsymbol{B})_F = \boldsymbol{A}:\boldsymbol{B} := \sum_{ij} a_{ij}b_{ij}.$$

We begin by exploring an orthogonal decomposition in $(\cdot,\cdot)_F$: $\mathbb{M} = \mathbb{T}\oplus^{\perp_F}\mathbb{R}\boldsymbol{I}$ and

$$\|\boldsymbol{\sigma}\|_F^2 = \|\boldsymbol{\sigma}^D\|_F^2 + \|P_{\mathbb{R}}\boldsymbol{\sigma}\|_F^2,$$

where $P_{\mathbb{R}}\boldsymbol{\sigma} = \operatorname{tr}(\boldsymbol{\sigma})\boldsymbol{I}/d$ is the orthogonal projector in $(\cdot,\cdot)_F$ inner product, and $\boldsymbol{\sigma}^D = (I - P_{\mathbb{R}})\boldsymbol{\sigma}$. Recall that

$$\mathcal{A}\boldsymbol{\sigma} = \frac{1}{2\mu}\left(\boldsymbol{\sigma} - \frac{\lambda}{2\mu + d\lambda}\operatorname{tr}(\boldsymbol{\sigma})I\right).$$

Considering the orthogonal decomposition $\mathbb{M} = \mathbb{T}\oplus^{\perp}\mathbb{R}\boldsymbol{I}$, we observe the compliance tensor $\mathcal{A}$, represented as $\operatorname{diag}(1/(2\mu), 1/(d\lambda + 2\mu))$, in this coordinate.

Lemma 2.1. *Let* $P_{\mathbb{R}}\boldsymbol{\sigma} = d^{-1}\operatorname{tr}(\boldsymbol{\sigma})\boldsymbol{I}$ *and* $\boldsymbol{\sigma}^D = (I - P_{\mathbb{R}})\boldsymbol{\sigma}$. *Then,*

$$\mathcal{A}\boldsymbol{\sigma} : \boldsymbol{\tau} = \frac{1}{2\mu}\boldsymbol{\sigma}^D : \boldsymbol{\tau}^D + \frac{1}{d\lambda + 2\mu} P_{\mathbb{R}}(\boldsymbol{\sigma}) : P_{\mathbb{R}}(\boldsymbol{\tau}),$$

Proof. Using the formulae of $\mathcal{A}$, and let $\rho = 2\mu/(d\lambda + 2\mu) \in (0,1)$, we have

$$2\mu\mathcal{A}\boldsymbol{\sigma} = \boldsymbol{\sigma} + (1-\rho)P_{\mathbb{R}}\boldsymbol{\sigma} = \boldsymbol{\sigma}^D + \rho P_{\mathbb{R}}\boldsymbol{\sigma}.$$

Using the property of orthogonal projectors, we expand the product

$$2\mu\mathcal{A}\boldsymbol{\sigma} : \boldsymbol{\tau} = (\boldsymbol{\sigma}^D + \rho P_{\mathbb{R}}\boldsymbol{\sigma}) : (\boldsymbol{\tau}^D + P_{\mathbb{R}}\boldsymbol{\tau}) = \boldsymbol{\sigma}^D : \boldsymbol{\tau}^D + \rho P_{\mathbb{R}}(\boldsymbol{\sigma}) : P_{\mathbb{R}}(\boldsymbol{\tau}),$$

and the identity then follows. □

This decomposition directly leads to a coercivity condition:

$$a(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \geq \min\left\{\frac{1}{2\mu}, \frac{1}{d\lambda + 2\mu}\right\} \|\boldsymbol{\sigma}\|^2, \quad \forall \boldsymbol{\sigma} \in \boldsymbol{\Sigma}. \tag{19}$$

Here, the constant α is on the order of $\mathcal{O}(1/\lambda)$ as λ approaches infinity. This characteristic indicates that the coercivity is not robust with respect to λ, becoming nearly singular as $\lambda \gg 1$.

It is crucial to note that while the norm for the space $\boldsymbol{\Sigma}$ is defined as $\|\cdot\|_{\operatorname{div}}$, the coercivity requirement in the L^2-norm specifically pertains to elements within $\ker(\operatorname{div})$. This distinction necessitates additional consideration, especially when $\ker(B)$ does not coincide with $\ker(\operatorname{div})$.

2.4. **A robust coercivity.** Recall that coercivity is required only in the kernel space of the divergence operator, and the compliance tensor $\mathcal{A}$ approaches near singularity in the subspace $\mathbb{R}\boldsymbol{I}$. To tackle this, we employ a strategy that connects linear elasticity with the inf-sup stability of Stokes equations, enhancing our approach by controlling the norm of the trace through the addition of $\|\operatorname{div}\boldsymbol{\sigma}\|_{-1}$.

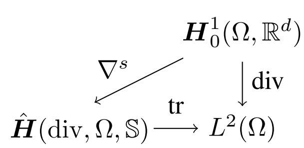

FIGURE 1. Linear elasticity and the inf-sup stability of Stokes equations.

Lemma 2.2. *There is a constant* β, *dependent only on the domain* Ω, *such that*

$$\|P_{\mathbb{R}}\boldsymbol{\sigma}\|^2 \leq \beta\left(\|\boldsymbol{\sigma}^D\|^2 + \|\operatorname{div}\boldsymbol{\sigma}\|_{-1}^2\right), \quad \textit{for all } \boldsymbol{\sigma} \in \widehat{\boldsymbol{H}}(\operatorname{div}, \Omega; \mathbb{S}). \tag{20}$$

Proof. Let $p = \operatorname{tr}\boldsymbol{\sigma} \in L_0^2(\Omega)$. Thanks to the inf-sup stability of the Stokes equation, there exists $\boldsymbol{v} \in \boldsymbol{H}_0^1$ such that $\operatorname{div}\boldsymbol{v} = p$ and $\|\boldsymbol{v}\|_1 \lesssim \|p\| = \|\operatorname{tr}\boldsymbol{\sigma}\|$.

Considering $\operatorname{div}\boldsymbol{v} = \operatorname{tr}\nabla^s\boldsymbol{v} = \operatorname{tr}\boldsymbol{\sigma}$, we have

$$\begin{aligned}
\|\operatorname{tr}(\boldsymbol{\sigma})\|^2 &= (\operatorname{tr}\nabla^s\boldsymbol{v}, \operatorname{tr}\boldsymbol{\sigma}) = d(P_{\mathbb{R}}\nabla^s\boldsymbol{v}, P_{\mathbb{R}}\boldsymbol{\sigma}) = d(\nabla^s\boldsymbol{v}, P_{\mathbb{R}}\boldsymbol{\sigma}) \\
&= -d(\nabla^s\boldsymbol{v}, \boldsymbol{\sigma}^D) + d(\nabla^s\boldsymbol{v}, \boldsymbol{\sigma}) \\
&= -d(\nabla^s\boldsymbol{v}, \boldsymbol{\sigma}^D) - d(\boldsymbol{v}, \operatorname{div}\boldsymbol{\sigma}).
\end{aligned}$$

Applying the Cauchy-Schwarz inequality and the definition of $\|\cdot\|_{-1}$, we find

$$\|\operatorname{tr}(\boldsymbol{\sigma})\|^2 \lesssim \|\boldsymbol{\sigma}^D\|\|\nabla^s \boldsymbol{v}\| + \|\operatorname{div}\boldsymbol{\sigma}\|_{-1}\|D\boldsymbol{v}\| \lesssim \left(\|\boldsymbol{\sigma}^D\|^2 + \|\operatorname{div}\boldsymbol{\sigma}\|_{-1}^2\right)\|\operatorname{tr}\boldsymbol{\sigma}\|.$$

Cancelling one $\|\operatorname{tr}\boldsymbol{\sigma}\|$ yields the desired inequality. □

We then obtain a robust coercivity of $a(\cdot,\cdot)$ restricted to the null space $\ker(\operatorname{div})$.

Theorem 2.3. *There exists a constant α depending on Ω and μ, but independent of λ such that*

$$a(\boldsymbol{\sigma},\boldsymbol{\sigma}) \geq \alpha\|\boldsymbol{\sigma}\|^2 \quad \textit{for all } \boldsymbol{\sigma} \in \boldsymbol{\Sigma} \cap \ker(\operatorname{div}). \tag{21}$$

Proof. By Lemma 2.2, we have

$$2\mu\, a(\boldsymbol{\sigma},\boldsymbol{\sigma}) = \|\boldsymbol{\sigma}^D\|^2 + \rho\|P_{\mathbb{R}}\boldsymbol{\sigma}\|^2 \geq \|\boldsymbol{\sigma}^D\|^2,$$

where we drop the term with $\rho = 2\mu/(d\lambda + 2\mu) \to 0^+$ as $\lambda \to +\infty$. On the other hand, we can control

$$\|\boldsymbol{\sigma}\|^2 = \|\boldsymbol{\sigma}^D\|^2 + \|P_{\mathbb{R}}\boldsymbol{\sigma}\|^2 \leq (1+\beta)\|\boldsymbol{\sigma}^D\|^2.$$

The desired coercivity then follows for $\alpha = (1+\beta)/(2\mu)$. □

REFERENCES

[1] J. Bramble. A proof of the inf-sup condition for the Stokes equations on Lipschitz domains. *Mathematical Models and Methods in Applied Sciences*, 13(3):361–372, 2003. 4

[2] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15. Springer-Verlag, New York, second edition, 2002. 3

[3] P. G. Ciarlet. On Korn's inequality. *Chinese Annals of Mathematics, Series B*, 31(5):607–618, Sept. 2010. 4