PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

PRIMAL AND MIXED FINITE ELEMENT METHODS FOR IMAGE REGISTRATION

NICOLÁS ALEJANDRO BARNAFI WITTWER

Thesis submitted to the Office of Research and Graduate Studies
in partial fulfillment of the requirements for the degree of
Master of Science in Engineering

Advisor:
DANIEL HURTADO
GABRIEL GATICA

Santiago de Chile, October 2017

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

PRIMAL AND MIXED FINITE ELEMENT METHODS FOR IMAGE REGISTRATION

NICOLÁS ALEJANDRO BARNAFI WITTWER

Members of the Committee:
DANIEL HURTADO
GABRIEL GATICA
CARLOS SING-LONG
MATÍAS COURDURIER
PEDRO GAZMURI

Thesis submitted to the Office of Research and Graduate Studies
in partial fulfillment of the requirements for the degree of
Master of Science in Engineering

Santiago de Chile, October 2017

Gratefully to my parents, brother
and girlfriend for their
unconditional support

ACKNOWLEDGEMENTS

I would like to thank first of all my two advisors, Daniel Hurtado and Gabriel Gatica for their guidance, wisdom, hospitality and patience. At times it was hard to cope with the requirements of both applications and theory, but in the end it was tremendously beneficial. I thank also Carlos Sing-Long for very illuminating debates regarding registration as well as optimization perspectives, Elwin van t'Wout for his help in formal aspects of applications, which of course I underestimated at first and Ricardo Ruiz-Baier for all his support in establishing my prospective studies.

Thanks to the people who were always there for small talk in my periods of frustration, in particular Natalia Ahumada, Mariana Cepeda, Ise De Smet and Constanza Toro for always receiving me with a smile. This also includes the Biomedical and Mathematical Engineering postgraduate students, who helped me see my problems from different perspectives, as well as other problems to solve when mine showed no future.

I am deeply grateful of the financial support I received during my work from VRI Puente P1617 from Pontificia Universidad Católica de Chile, CONICYT-Chile through BASAL project CMM from Universidad de Chile, and by Centro de Investigación en Ingeniería Matemática (CI^{2}MA) from Universidad de Concepción.

Thanks also to my parents, Esteban Barnafi K. and Anita Wittwer, and brother Esteban Barnafi W., for their patience with my uncertainty and exaggerated motivation regarding Mathematics. Finally, I would like to specially thank my girlfriend Daniela Cofré, for her understanding, empathy and support. I would never have had the energy to get so far without her overwhelming common sense and love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Deformable Image Registration (DIR) is a powerful computational method for image analysis, with promising applications in the diagnosis of human disease. Despite being widely used in the medical imaging community, the mathematical and numerical analysis of DIR methods still has many open questions. Further, recent applications of DIR include the quantification of mechanical quantities in addition to the aligning transformation, which justifies the development of novel DIR formulations for which the accuracy and convergence of fields other than the aligning transformation can be studied. In this work we propose and analyze a primal, mixed and augmented formulations for the DIR problem, together with their finite-element discretization schemes for their numerical solution. The DIR variational problem is equivalent to the linear elasticity problem with a source term that has a nonlinear dependence on the unknown field. Fixed point arguments and small data assumptions are employed to derive the well-posedness of both the continuous and discrete schemes for the usual primal and mixed variational formulations, as well as for an augmented version of the later. In particular, continuous piecewise linear elements for the displacement in the case of the primal method, and Brezzi-Douglas-Marini of order 1 (resp. Raviart-Thomas of order 0) for the stress together with piecewise constants (resp. continuous piecewise linear) for the displacement when using the mixed approach (resp. its augmented version), constitute feasible choices that guarantee the stability of the associated Galerkin systems. A priori error estimates derived by using Strang-type lemmas, and their associated rates of convergence depending on the corresponding approximation properties are also provided. Numerical convergence tests and DIR examples are included to demonstrate the applicability of the method.

Keywords: Image registration, mixed formulation, finite element methods.

RESUMEN

La registración deformable de imágenes (DIR) representa un poderoso método computacional para analizar imágenes, con prometedoras aplicaciones en el diagnóstico en enfermedades humanas. A pesar de ser ampliamente usado en la comunidad de imágenes médicas, el análisis matemático y numérico de métodos de DIR sigue abierto. Además, aplicaciones recientes de DIR incluyen la cuantificación de indicadores mecánicos además de la transformación de alineamiento, lo que justifica el desarrollo de formulaciones novedosas de DIR donde la precisión y convergencia de campos que no sean el alineamiento pueden ser estudiadas. En este trabajo proponemos y analizamos formulaciones primal, mixta y aumentada para DIR, junto con su esquema de discretización de elementos finitos para su resolución numérica. La formulación variacional de DIR equivale a un problema de elasticidad lineal con un término fuente que depende no-linealmente de la alineación. Argumentos de punto fijo y supuestos de datos pequeños son utilizados para obtener la correcta formulación tanto del esquema continuo como del discreto para todas las formulaciones. En particular, elementos lineales por tramo para el desplazamiento en el caso primal, elementos Brezzi-Douglas-Marini de orden 1 (resp. Raviart-Thomas de orden 0) para la tensión junto con constantes por tramo (resp. lineal por tramo) para el desplazamiento para el esquema mixto (resp. aumentado), constituyen elecciones factibles que garantizan la estabilidad de los sistemas de Galerkin asociados. Estimaciones de error a-priori se obtuvieron usando lemas de tipo Strang, y sus asociadas tasas de convergencia de acuerdo a propiedades de aproximación se muestran también. Ensayos de convergencia numérica y ejemplos de DIR se incluyen para mostrar la aplicabilidad del método.

Palabras Claves: Registración de imágenes, formulación mixta, elementos finitos.

1. INTRODUCTION

The collaborative work performed at the intersection between Biology and Mathematics has seen a tremendous increase in the last 20 years, with a very symbiotic relationship. From the perspective of Biology this happens because many biological questions can be answered through the creation of models that describe complex phenomena, now numerically solvable thanks to the increased computational capacity; from the perspective of Mathematics, it gives a plethora of new open problems to analyze, both in theory and practice. The problem we worked on, known as image registration, has acquired a great amount of attention within the academic community, and consists in finding an optimal alignment between two images. It is important in many applications, and it has been used to: determine how to match separate images of one big picture, study tumor growth, analyze the growth of bacteria populations, find cloud dynamics and generate lung dynamics from CT images, to name a few. In particular, we focus in deformable image registration (DIR), which allows for local deformation at the expense of higher computational cost.

In the classical literature, the main ingredients of image registration are: i) the transformation model, that is, a family of mappings that warp a target image into the reference image, ii) the similarity measure, that is, a functional that weighs the differences between the reference image and the resampled target image, and iii) the regularizer, which renders the problem well-conditioned by adding regularity to the DIR solution. DIR has been approached from mainly three perspectives:

- Minimization of a similarity measure, where tools from the calculus of variations have been used to establish existence of solutions to the continuous problem (Aubert & Kornprobst, 2006; Vese & Le Guyader, 2016; Horn & Schunck, 1980; Dupuis, Grenander, & Miller, 1998) and from that point the difficulty lies in the implementation. This approach leaves the discrete setting aside, and has focused mainly in the development of new techniques to deal with the nonlinearity of the problem.

- Optimal mass transport (Haker, Zhu, Tannenbaum, & Angenent, 2004; Museyko, Stiglmayr, Klamroth, & Leugering, 2009; Burger, Modersitzki, & Ruthotto, 2013; Borzi, Ito, & Kunisch, 2002). This approach uses very advanced mathematics, as it strongly relies in notions of measure theory. This has kept this approach far from the applications community.
- A level set segmentation-registration combined problem (Unal & Slabaugh, 2005; Droske & Ring, 2006). This approach has shown promising results, and depends mainly in shape optimization techniques. These techniques are very expensive computationally, and thus make this approach unattractive for the medical community.

Despite them being widely used in the medical imaging community, there are still many questions to be answered from the numerical analysis point of view. A noteworthy approach is the work of Pöschl *et al.* (Pöschl, Modersitzki, & Scherzer, 2010), where both the continuous and discretized problems are analyzed, and a solution is found using a primal finite-element approximation that is shown to be convergent. However, the analysis is restricted to polyconvex energy densities (both for the similarity measure and regularizer) and volume-preserving transformations, and does not account for the convergence of the transformation gradients and stresses. A more traditional Galerkin approach has been introduced in (Lee & Gunzburger, 2011) for optimal-control-based registration, but requires a considerable degree of regularity ($H^{2+\delta}$) of the target and reference image functions, not required by other traditional formulations. While most approaches to DIR problems are based on primal formulations, a mixed formulation of the similarity minimization problem has been proposed in the setting of fluid registration schemes (Chen & Lorenz, 2011; Ruhnau & Schnörr, 2007), where a sequence of incompressible Stokes problems are solved to find the optimal displacement and pressure fields. While directly solving for the pressure field, which is desirable to understand the mechanical behavior of the images being registered, limited analysis has been provided to understand the well-posedness of the continuous problem and convergence of numerical discretizations of mixed formulations of

DIR problems.

One important and recent application of DIR is the study of local lung tissue deformation from computed-tomography (CT) images of the thorax (Christensen, Song, Lu, El Naqa, & Low, 2007). From a continuum mechanics perspective, the transformation $\boldsymbol{u}$ obtained from the solution of the DIR problem can be considered as a displacement field, from which a strain tensor field can be computed from its gradient $\nabla \boldsymbol{u}$. Local deformation studies have revealed that deformation in the lung tissue can be highly heterogenous and anisotropic (Amelon et al., 2011; Hurtado et al., 2017), thus providing new deformation-based markers to understand lung physiology (Choi et al., 2013), showcasing the potential of DIR strain analysis as a tool in the detection and diagnosis of pulmonary disease. These advances notwithstanding, it has been recently shown that strain analysis techniques based on numerical differentiation of DIR solutions can be highly inaccurate as they are very sensitive to noise, discretization level and embedded anatomical boundaries (Hurtado, Villarroel, Retamal, Bugedo, & Bruhn, 2016). While these deficiencies are diminished with L^2 projection smoothing techniques (Hurtado et al., 2016), this approach lacks a rigorous mathematical framework that can support the stability and accuracy of the resulting strain fields. This last observation motivates the development of DIR methods that can accurately predict not only the image transformation but also its gradient by proving convergence of the associated numerical scheme. In the field of elasticity, the convergence of displacement fields and quantities associated to its gradient has been successfully approached using mixed formulations (Gatica, Márquez, & Meddahi, 2008).

In this work we present three approaches to DIR. One is a primal scheme, which has already been analyzed in more particular cases, but this approach is better suited to prove convergence of discrete schemes. It also establishes convergence for all conforming schemes, thus showing that all solutions obtained using the finite elements method (FEM) to the problem are convergent. The second one is a mixed finite elements scheme (MFEM), which incorporates stress as an unknown. This makes the problem more computationally expensive, because it adds new degrees of freedom, but it also grants best approximation properties to all unknown fields, thus giving an accurate stress estimation. This scheme is also convergent, but only to a limited number of MFEM schemes. Finally, we use an augmented formulation, which is a MFEM formulation that incorporates redundant terms. This regularizes the problem and allows for the use of any discrete scheme, and guarantees convergence. Galerkin schemes are a great tool to automatically generate transformation models which grant convergence numerical convergence to the continuous solution, which gives new arguments to better understand problems in models currently used in the community. In lung registration, evidence shows correlation between stress and lung injury, which makes precise estimations of stress critical (Hurtado et al., 2017). It was also shown in (Hurtado et al., 2017) that common techniques used for DIR present highly oscillatory stress estimations, leaving the problem of its approximation largely open.

This thesis is ordered as follows: the rest of the introduction presents the tools used for the primal, mixed and augmented schemes. Chapter 2 includes the article submitted, chapter 3 shows conclusions that can be obtained from the work and chapter 4 proposes new research directions for future work.

1.1. Image registration

The alignment between images is formulated as follows: Given a Lipschitz domain Ω, reference image $R : \Omega \to \mathbb{R}$ and a target image $T : \Omega \to \mathbb{R}$, find a function $u : \Omega \to \mathbb{R}^n$ such that

$$R(x) = T(x + u(x)) \qquad \forall x \in \Omega.$$

This problem is ill-posed because of the following reasons:

- Differences in resolution can make the problem infeasible, for instance, meaning that there could be image intensities that never match.

- The solution is not unique. A simple case is two concentric circles, one stretched with respect to the other. In this case, all rotations maintain the alignment.

It is also important to notice that the solution can be very unnatural, as this model does not enforce any physical constraints, which suggests using smaller solution spaces. For this, we consider a similarity measure, which is minimized when the images are aligned, and a regularizer, which adds smoothness to the solution (Modersitzki, 2004). Under these considerations, we restate image registration as a regularized variational problem as follows: Given a solution space $\mathcal{V}$, a similarity measure $\mathcal{D} : \mathcal{V} \to \mathbb{R}$, a regularizer $\mathcal{S} : \mathcal{V} \to \mathbb{R}$ and a regularization parameter $\alpha > 0$, we seek for minimizers to the following problem

$$\inf_{u \in \mathcal{V}} \alpha \mathcal{D}[u] + \mathcal{S}[u]. \tag{1.1}$$

The choice of the similarity measure and the regularizer is highly context-dependent, where, for example, for the alignment of pictures it might suffice to take $\mathcal{S} = 0$ and reduce the solution space to linear functions or rigid motions. In DIR, solutions spaces are usually Sobolev spaces, similarity measures are usually the norm of the difference between the images (in L^2 or H^1), the correlation between the images or their mutual information, and the regularizer is an energy term. Throughout our work we consider the following instance, known as elastic registration because of the form regularizer, which is related to the elastic energy of a body:

$$\mathcal{D}[u] = \int (T(x + u(x)) - R(x))^2 \, \mathrm{d}x,$$

$$\mathcal{S}[u] = \int \mathbb{C}\varepsilon(u) : \varepsilon(u) \, \mathrm{d}x,$$

where $\mathbb{C}$ is Hooke's tensor for an isotropic linear elastic body, defined in terms of Lamé's parameters μ, λ:

$$\mathbb{C}\tau = 2\mu \, \tau + \lambda \, \text{trace}\, (\tau) I$$

and ε is the symmetric part of the gradient:

$$\varepsilon = \frac{1}{2}\left(\nabla u + [\nabla u]^{\mathrm{t}}\right)$$

Several techniques can be extended to other kinds of registration, which is stated in remarks throughout the article.

1.2. Preliminaries

Now we introduce some preliminary results required to understand what follows; for more details, see any functional analysis textbook (Brezis, 2011). A (real) Hilbert space $\mathcal{V}$ is a complete vector space endowed with an inner product $\langle\cdot,\cdot\rangle : \mathcal{V}\times\mathcal{V}\to\mathbb{R}$ which is:

- symmetric: $\langle x, y\rangle = \langle y, x\rangle$ for every x, y in $\mathcal{V}$.
- linear in the first argument: $\langle ax_1 + bx_2, y\rangle = a\langle x_1, y\rangle + b\langle x_2, y\rangle$ for a, b scalars and x, y in $\mathcal{V}$.
- positive definite: $\langle x, x\rangle \geq 0$ for x in $\mathcal{V}$, and the equality holds only if $x = 0$.

These spaces have an induced norm given by

$$\|x\|_{\mathcal{V}} = \sqrt{\langle x, x\rangle}$$

and the Cauchy-Schwartz inequality holds on them:

$$\langle x, y\rangle \leq \|x\|_{\mathcal{V}}\|y\|_{\mathcal{V}}.$$

We let $\mathcal{V}'$ be the topological dual space of bounded linear functionals from $\mathcal{V}$ to $\mathbb{R}$. There exists an explicit isometry on Hilbert spaces onto their dual in virtue of Riesz representation theorem, and we let $\mathcal{R} : \mathcal{V}'\to\mathcal{V}$ be such an isometry. This states that functionals on Hilbert spaces can be written as integrals, so that for a given T in $\mathcal{V}'$, we can write

$$T(u) = \langle \mathcal{R}(T), u \rangle.$$

We shall adopt the convention of the duality pairing. This means that given T in $\mathcal{V}'$, we write $\langle T, u \rangle$ instead of $T(u)$. Note that no information is given through this notation if $\mathcal{V}$ is a Hilbert space, in virtue of Riesz representation theorem, and it is still widely in the mathematics community used if it is not a Hilbert space. Finally, given a distribution, i.e. a functional d in $(C_0^\infty)'$, we can define its distributional gradient as the functional given by

$$\langle \nabla d, \phi \rangle := -\langle d, \text{div } \phi \rangle \qquad \forall \phi \in C_0^\infty.$$

This allows us to write more comfortably abstract gradients to obtain first order conditions, and as this derivative matches the classical Gâteaux derivative from variational analysis in our context plus extra boundary terms. Assuming Gâteaux differentiability, we get the following strong form of the first order conditions for a critical solution u^*:

$$\nabla \mathcal{S}[u^*] = -\alpha \nabla \mathcal{D}[u^*],$$

or in weak form

$$\langle \nabla \mathcal{S}[u^*], v \rangle = -\alpha \langle \nabla \mathcal{D}[u^*], v \rangle \qquad \forall v \in \mathcal{V}.$$

1.3. Primal formulation

The primal formulation of this problem is given for the Euler-Lagrange equations of problem (1.1). Defining $a : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ and $F_\cdot : \mathcal{V} \to \mathcal{V}'$ as

$$\begin{aligned} a(u, v) &:= \langle \nabla S[u], v \rangle \qquad \forall u, v \in \mathcal{V}, \\ F_u(v) &:= -\langle \nabla \mathcal{D}[u], v \rangle \quad \forall u, v \in \mathcal{V}, \end{aligned}$$

we obtain the abstract form of problem (1.1): Find u in $\mathcal{V}$ such that

$$a(u, v) = \alpha F_u(v) \qquad \forall v \in \mathcal{V}. \tag{1.2}$$

For elastic registration, this results in the following forms:

$$a(u,v) = \int \mathbb{C}\varepsilon(u) : \varepsilon(v)\, dx,$$

$$F_u(v) = -\int (T(x+u(x)) - R(x))\nabla T(x+u(x)) \cdot v\, dx,$$

This setting is now appropriate for Lax-Milgram lemma, which we recall now.

Lemma 1. *Let $\mathcal{V}$ be a Hilbert space. $a : \mathcal{V}\times\mathcal{V} \to \mathbb{R}$ be a bilinear form and $f : \mathcal{V} \to \mathbb{R}$ a bounded linear functional. Assume also that a is elliptic, meaning that there exists $c_a > 0$ such that*

$$a(v,v) \geq c_a \|v\|_{\mathcal{V}}^2 \qquad \forall v \in \mathcal{V},$$

and that it is bounded, meaning that there exists $c_b > 0$ such that

$$a(u,v) \leq c_b \|u\|_{\mathcal{V}} \|v\|_{\mathcal{V}} \qquad \forall u, v \in \mathcal{V}.$$

Under these assumptions, the problem of finding $u \in \mathcal{V}$ such that

$$a(u,v) = f(v) \qquad \forall v \in \mathcal{V}$$

has a unique solution, and there exists a constant $C > 0$ such that

$$\|u\|_{\mathcal{V}} \leq C \|f\|_{\mathcal{V}'}.$$

PROOF. See (Brenner & Scott, 2008, Thm 2.7.7). □

The last inequality is often referred to as stability, and shows that the problem is continuous with respect to the data. Lax-Milgram lemma states existence and uniqueness for a linear problem, whereas in problem (1.2) F is nonlinear in u. This issue is dealt with using a fixed-point argument, which means that we can define an operator $T : \mathcal{V} \to \mathcal{V}$ which, given z in $\mathcal{V}$, returns the solution to the partial problem stated as follows: Find u in $\mathcal{V}$ such that

$$a(u,v) = F_z(v) \qquad \forall v \in \mathcal{V}. \tag{1.3}$$

As the partial problem is well-posed, the operator T is well defined. From this point, the existence of solutions to (1.2) can be reduced to the existence of fixed points for the operator T. For this, we use two conditions on the data for some $L_F, M_F > 0$:

$$\|F_u - F_v\|\mathcal{V}' \leq L_F\|_u - v\|_{\mathcal{W}} \qquad \forall u, v \in \mathcal{V}, \tag{1.4}$$

$$\|F_u\| \leq M_F \qquad \forall u \in \mathcal{V}, \tag{1.5}$$

where $\mathcal{W}$ is such that $\mathcal{V}$ is compactly embedded in it. These conditions are verifiable for almost every similarity measure. In practice, images are interpolated using cubic B-splines, which implies in particular that ∇T is Lipschitz in Ω. As the domain is compact, the function is also bounded, thus satisfying both conditions (1.4) and (1.5).

Now we can formulate the discrete problem. For this, we will use $\mathcal{V} = [H^1(\Omega)]^2$ and let $\{\mathcal{T}_h\}_{h>0}$ be a regular non-degenerate family of triangulations of Ω such that for every $h > 0$ we have

$$\bigcup_{\tau \in T_h} \tau = \overline{\Omega}$$

and also

$$\mathring{\tau}_i \bigcap \mathring{\tau}_j = \emptyset$$

for every τ_i, τ_j in $\mathcal{T}_h$ such that $i \neq j$. Now we define the discrete space for a given order of approximation $k \in \mathbb{N}$:

$$\mathcal{V}_h := \{v_h \in \mathcal{V} : v_h|_\tau \in \mathbb{P}_k(\tau)\},$$

where $\mathbb{P}_k$ is the space of polynomials of degree up to k. We note that $\mathcal{V}_h \subset \mathcal{V}$ holds, which makes this a conforming scheme. The discrete primal problem can now be formulated as: Find u_h in $\mathcal{V}_h$ such that

$$a(u_h, v_h) = F_{u_h}(v_h) \qquad \forall v_h \in \mathcal{V}_h. \tag{1.6}$$

We show in the article, in chapter 2, that well-posedness for this problem follows from the regularity of a and thus is a direct consequence of the continuous analysis. In particular, the ellipticity of a holds in $\mathcal{V}$, and so in particular it holds in $\mathcal{V}_h$, rendering the problem well-posed. Again, a fixed-point argument establishes the existence of solutions to the nonlinear discrete problem.

1.4. Mixed finite elements method

In this section we will space the space

$$H(\mathrm{div}\,, \Omega) := \{\sigma : \sigma \in [L^2(\Omega)]^{n\times n}, \mathrm{div}\, \sigma \in [L^2(\Omega)]^n\},$$

where the divergence operator acts row-wise. To see the connection of primal registration works with the mixed formulation, we start with a seemingly unrelated example. Suppose we want to solve a Poisson problem with Neumann boundary conditions in a Lipschitz domain Ω and a source term f in $L^2(\Omega)$, written as

$$\begin{aligned} \Delta u &= f \quad \text{in } \Omega \ a.e., \\ \nabla u \cdot \boldsymbol{n} &= 0 \quad \text{in } \partial\Omega \ a.e. \end{aligned}$$

Defining the unknown $\sigma = \nabla u$ we can restate the problem as

$$\begin{aligned} \mathrm{div}\, \sigma &= f && \text{in } \Omega \ a.e., \\ \sigma &= \nabla u && \text{in } \Omega \ a.e., \\ \sigma \cdot \boldsymbol{n} &= 0 && \text{in } \partial\Omega \ a.e. \end{aligned} \tag{1.7}$$

If we define the spaces $H := H(\mathrm{div}\,, \Omega) \cap \{\tau \cdot \boldsymbol{n} = 0 \,\mathrm{in}\, \partial\Omega\}, Q := L^2(\Omega)$ and forms $a : H \times H \to \mathbb{R}, b : H \times Q \to \mathbb{R}, F : Q \to \mathbb{R}$ given by

$$\begin{aligned} a(\sigma, \tau) &:= \int \sigma \cdot \tau \, dx && \forall \sigma, \tau \in H, \\ b(\tau, v) &:= \int v \,\mathrm{div}\, \tau \, dx && \forall \tau \in H, v \in Q, \\ F(v) &:= \int f v \, dx && \forall v \in Q, \end{aligned}$$

we can rewrite problem (1.7) as what is called a mixed formulation: Find (σ, u) in $H \times Q$ such that

$$\begin{aligned} a(\sigma,\tau) + b(u,\tau) &= 0 \qquad \forall \tau \in H, \\ b(\sigma, v) &= F(v) \quad \forall v \in Q. \end{aligned} \tag{1.8}$$

This is the classical structure of a mixed problem, and they can be written more generally in matrix notation as

$$\begin{bmatrix} A & B^{\mathrm{t}} \\ B & 0 \end{bmatrix} \begin{bmatrix} \sigma \\ u \end{bmatrix} = \begin{bmatrix} G \\ F \end{bmatrix},$$

where A and B are linear operators induced by a and b respectively. These kinds of problems are rendered well-posed using Babuška-Brezzi theory, which we recall now.

Theorem 1. *Under the context of* (1.8)*, let*

$$V = \{\tau \in H : b(\tau, v) = 0 \quad \forall v \in Q\}.$$

If a is $V - elliptic$, meaning that there exists $c_a > 0$ such that

$$a(\tau,\tau) \geq c_a \|\tau\|_H^2,$$

and that b satisfies the inf-sup condition, meaning that there exists c_b such that

$$\sup_{\tau \in H} \frac{b(\tau, v)}{\|\tau\|_H} \geq c_b \|v\|_Q \qquad \forall v \in Q,$$

then problem (1.8) *has a unique solution.*

PROOF. See (Gatica, 2014, Thm 2.3). □

For this to apply in image registration, we must note how we obtained the mixed formulation for the Poisson problem. In this case, we started from a problem that can be written as

$$a(u, v) = f(v) \qquad \forall v \in \mathcal{V}$$

and then used substitutions in the left hand side to get a new problem. Under a similar procedure, we show in the article that the elastic registration problem can be written as:

Find $((\sigma,\rho),(u,\gamma))$ in $H\times Q$ such that

$$
\begin{aligned}
A((\sigma,\rho),(\tau,\xi)) + B((\tau,\xi),(u,\gamma)) &= 0 && \forall(\tau,\xi)\in H,\\
B((\sigma,\rho),(v,\eta)) &= \bar{F}_u((v,\eta)) && \forall(v,\eta)\in Q,
\end{aligned}
$$

where:

$$
\begin{aligned}
H &:= [H(\mathrm{div}\,;\varOmega)\cap\{\tau\cdot\boldsymbol{n}=0\,\mathrm{in}\,\partial\varOmega\}]\times\mathbb{RM},\\
\mathbb{RM} &:= \{v\in[H^1(\varOmega)]^2 : \mathrm{sym}\,\nabla v = 0\},\\
Q &:= [L^2(\varOmega)]^2\times\left([L^2(\varOmega)]^{2\times 2}\cap\{\tau=-\tau^{\mathrm{t}}\}\right),\\
A((\sigma,\rho),(\tau,\xi)) &= \int \mathbb{C}^{-1}\sigma\cdot\tau\,\mathrm{d}x,\\
B((\tau,\xi),(v,\eta)) &= \int \tau:\eta\,\mathrm{d}x + \int v\cdot\,\mathrm{div}\,\tau\,\mathrm{d}x + \int \xi\cdot v\,\mathrm{d}x,\\
\bar{F}_u((v,\eta)) &= F_u(v)
\end{aligned}
$$

As in the primal case, the approach to solving this problem is to consider a partial problem, a solution operator and the Banach fixed point theorem to ensure the existence of fixed points. This imposes a restriction on α, because it uses $\alpha C L_F < 1$, where L_F is the Lipschitz constant from (1.4) and $C>0$ is a constant inherent to the problem. For the discrete case, we take the same family of triangulations we used in the primal case $\{\mathcal{T}_h\}_{h>0}$ and consider the following finite element spaces:

$$
\begin{aligned}
H_h^\sigma &= \{\tau_h\in H(\mathrm{div}\,;\varOmega):\tau_h\in[\mathbb{BDM}_1(T)]^2\quad\forall T\in\mathcal{T}_h\},\\
H_h^u &= \{u\in[L^2(\varOmega)]^2 : u\in[\mathbb{P}_0(T)]^2\quad\forall T\in\mathcal{T}_h\},\\
H_h^\gamma &= \left\{\begin{bmatrix} 0 & \psi\\ -\psi & 0\end{bmatrix} : \psi\in\mathbb{P}_0(T)\quad\forall T\in\mathcal{T}_h\right\},
\end{aligned}
$$

where $\mathbb{BDM}_k = [\mathbb{P}_k]^{n\times n}$. With this, setting $H_h = [H_h^\sigma\cap\{\tau=0\text{ in }\partial\varOmega\}]\times\mathbb{RM}$ and $Q_h = H_h^u\times H_h^\gamma$ we obtain the discrete registration problem: Find $((\sigma_h,\rho_h),(u_h,\gamma_h))\in H_h\times Q_h$

such that

$$\begin{aligned} A((\sigma_h,\rho_h),(\tau_h,\xi_h)) + B((\tau_h,\xi_h),(u_h,\gamma_h)) &= 0 && \forall(\tau_h,\xi_h)\in H_h, \\ B((\sigma_h,\rho_h),(v_h,\eta_h)) &= \bar{F}_{u_h}((v_h,\eta_h)) && \forall(v_h,\eta_h)\in Q_h, \end{aligned}.$$

Again, this problem is tackled using a fixed point strategy. The difficulty of proving inf-sup stability has already been proved for the partial problem (Gatica et al., 2008), and so we had to prove the extension to the nonlinear problem and the discrete convergence to the continuous solution. Everything is performed under data assumptions (1.4) and (1.5), and under these assumptions we get a convergence of $O(h)$, which was numerically verified.

1.5. Augmented MFEM

The augmented formulation is obtained from the mixed formulation, by adding redundant terms from the strong formulation. For that, we consider the following forms for constants $\kappa_1, \kappa_2, \kappa_3$:

$$\begin{aligned} \kappa_1 \int (\varepsilon(u) - \mathbb{C}^{-1}\sigma) : (\varepsilon(v) + \mathbb{C}^{-1}\tau)\, dx &= 0, \\ \kappa_2 \int \operatorname{div}\sigma \cdot \operatorname{div}\tau\, dx &= \kappa_2 \alpha \int f_u(x)\cdot \operatorname{div}\tau\, dx, \\ \kappa_3 \int (\gamma - \operatorname{skew}\nabla u) : (\eta + \operatorname{skew}\nabla v)\, dx &= 0, \end{aligned}$$

where

$$f_u(x) := (T(x+u(x)) - R(x))\nabla T(x+u(x)).$$

For simplicity, we let L_1, L_2, L_3 be the left hand sides of the added terms and R_2 the non-zero right hand side. Then, the new formulation is obtained as follows: we redefine out solution spaces as $H = [H(\operatorname{div}; \Omega) \cap \{\tau = 0 \operatorname{in} \partial\Omega\}] \times [H^1(\Omega)]^2 \times [L^2(\Omega)]^{2\times}$ and

$Q = \mathbb{RM}$:

$$
\begin{aligned}
\bar{A}((\sigma, u, \gamma), (\tau, v, \eta)) &:= A(\sigma, \tau) + B(\tau, (u, \gamma) - B(\sigma, (v, \eta)) \\
&\quad + L_1 + L_2 + L_3, \\
\bar{B}((\tau, v, \gamma), \xi) &= \int v \cdot \xi, \\
\bar{F}_u((\tau, v, \eta)) &= -\int f_u \cdot v \, \mathrm{d}x + \frac{1}{\alpha} R_2.
\end{aligned}
$$

Here, note that we have imposed stronger regularity on u, as it is now constrained to belong to $[H^1(\Omega)]^2$. This allows us to give a new interpretation to the original mixed formulation. As u is in $[H^1(\Omega)]^2$, it is actually a non-conforming scheme considering the true regularity of the displacement field. This is of course not true from a rigorous point of view, but it gives a way of approximating a continuous field with piecewise constants. We also note that the inf-sup condition will hold for both the continuous and discrete cases, because we can make in both cases the following calculation taking $(\tau, v, \eta) = (0, \xi, 0)$:

$$
\sup_{(\tau,v,\eta)\in H} \frac{B((\tau, v, \eta), \xi)}{\|(\tau, v, \xi)\|_H} \geq \frac{\|\xi\|_0}{\|\xi\|_1^2} \geq C\|\xi\|_0^2,
$$

for a positive constant C. Here, as $\mathbb{RM}$ is a finite dimensional space, its norms are equivalent. We now state the augmented problem: Find $((\sigma, u, \gamma), \rho)$ in $H \times Q$ such that

$$
\begin{aligned}
\bar{A}((\sigma, u, \gamma), (\tau, v, \eta)) + \bar{B}((\tau, v, \xi), \rho) &= \alpha \bar{F}_u((\tau, v, \eta)) \quad && \forall (\tau, v, \eta) \in H, \\
B((\sigma, u, \gamma), \xi) &= 0 && \forall \xi \in Q.
\end{aligned} \tag{1.9}
$$

The existence of solutions to this problem can be stated independently of α. This can be done because there exists a compact immersion from $H^1(\Omega)$ to $L^2(\Omega)$, which allows using the Schauder fixed point theorem. This does not state uniqueness, which again can be obtained from the Banach fixed point theorem and $\alpha C L_F < 1$.

As the inf-sup condition holds for every conforming discrete space, we can use whichever space is adequate for the context. In this case, we simply use the least expensive ones in terms of the amount of degrees of freedom they generate, given by:

$$
\begin{aligned}
H_h^\sigma &:= \{\tau \in H(\text{div}\,;\Omega) : \tau \in [\mathbb{RT}_0(T)]^2 \quad \forall T \in \mathcal{T}_h\},\\
H_h^u &:= \{u \in [H^1(\Omega)] : u \in [\mathbb{P}(T)_1]^2 \quad \forall T \in \mathcal{T}_h\},\\
H_h^\gamma &= \left\{ \begin{bmatrix} 0 & \psi \\ -\psi & 0 \end{bmatrix} : \psi \in \mathbb{P}_0(T) \quad \forall T \in \mathcal{T}_h \right\}.
\end{aligned}
$$

With that, and setting the discrete spaces $H_h = [H_h^\sigma \cap \{\tau = 0 \,\text{in}\, \partial\Omega\}] \times H_h^u \times H_h^\gamma$ and $Q_h = Q$, we arrive at the discrete problem: Find $((\sigma_h, u_h, \gamma_h), \rho_h)$ in $H_h \times Q_h$ such that

$$
\begin{aligned}
\bar{A}((\sigma_h, u_h, \gamma_h), (\tau_h, v_h, \eta_h)) + \bar{B}((\tau_h, v_h, \eta_h),) &= \bar{F}_{u_h}((\tau_h, v_h, \eta_h)) \quad &\forall (\tau_h, v_h, \eta_h) \in H_h,\\
\bar{B}((\sigma_h, u_h, \gamma_h), \xi_h) &= 0 \quad &\forall \xi_h \in Q_h
\end{aligned}
\tag{1.10}
$$

This problem is analyzed with a similar fixed point strategy, but the existence of solutions is obtained with the simpler Brouwer fixed point theorem. This grants also convergence of the discrete solution to the continuous one, and guarantees a convergence of $O(h)$. It is important to mention that convergence of u in L^2 is in this case $O(h^2)$.

2. PRIMAL AND MIXED FINITE ELEMENT METHODS FOR THE DEFORMABLE IMAGE REGISTRATION PROBLEM

2.1. Introduction

Deformable image registration (DIR) concerns the problem of aligning two or more images by means of a transformation that allows for distortion (warping) of the images under analysis. Such problem arises in a number of important applications, particularly in the field of medical imaging (Sotiras, Davatzikos, & Paragios, 2013). DIR is commonly formulated as a variational problem, where the main ingredients of the DIR are i) the transformation model, a family of mappings that warp a target image into the reference image, ii) the similarity measure, a functional that weighs the differences between the reference image and the resampled target image, and iii) the regularizer, which renders the problem well-conditioned by adding regularity to the DIR solution.

Despite DIR is widely used in the medical imaging community, the mathematical and numerical analysis of DIR remains understudied. The DIR continuous problem has been formulated using mainly three approaches: minimization of similarity measures (with or without constraints), as an optimal mass transport problem (Haker et al., 2004; Museyko et al., 2009; Burger et al., 2013), or as a level set segmentation-registration combined problem (Unal & Slabaugh, 2005; Droske & Ring, 2006). The problem of minimizing similarity measures has been studied in (Aubert & Kornprobst, 2006; Vese & Le Guyader, 2016), where the direct method of calculus of variations has been used to establish existence of solutions. The optical flow formulation, an associated problem which can be seen as a sequence of registration problems in time, was proposed by Horn & Schunk in 1980 (Horn & Schunck, 1980), and has been the subject of analysis from an optimal-control problem point of view (Borzi et al., 2002; Lee & Gunzburger, 2010). Well-posedness of optical flow schemes has been established for Dirichlet boundary conditions under reasonable assumptions (Dupuis et al., 1998; Trouvé, 1998). Besides providing existence and uniqueness of the solution, by assuming only uniform boundedness on the images, these studies show that the solution is a step-wise diffeomorphism, which is a desirable regularity property when it comes to warping images. The analysis of the numerical schemes proposed to solve similarity-minimization formulations has received less attention. A noteworthy approach is the work of Pöschl *et al.* (Pöschl et al., 2010), where both the continuous and discretized problems are analyzed, and a solution is found using a primal finite-element approximation that is shown to be convergent. However, the analysis is restricted to polyconvex energy densities (both for the similarity measure and regularizer) and volume-preserving transformations, and does not account for the convergence of the transformation gradients and stresses. A more traditional Galerkin approach has been introduced in (Lee & Gunzburger, 2011) for optimal-control-based registration, but requires a considerable degree of regularity ($H^{2+\delta}$) of the target and reference image functions, not required by other traditional formulations. While most approaches to DIR problems are based on primal formulations, a mixed formulation of the similarity minimization problem has been proposed in the setting of fluid registration schemes (Chen & Lorenz, 2011; Ruhnau & Schnörr, 2007), where a sequence of incompressible Stokes problems are solved to find the optimal displacement and pressure fields. While directly solving for the pressure field, which is desirable to understand the mechanical behavior of the images being registered, limited analysis has been provided to understand the well-posedness of the continuous problem and convergence of numerical discretizations of mixed formulations of DIR problems that use elastic regularizers.

One important and recent application of DIR is the study of regional deformation of lung tissue from computed-tomography (CT) images of the thorax (Christensen et al., 2007). From a continuum-mechanics perspective, the optimal transformation $\boldsymbol{u}$ obtained from the solution of the DIR problem can be considered as a displacement field, from which a strain tensor field can be computed using the gradient of the displacement field $\nabla\boldsymbol{u}$. The study of not only motion but also deformation at a regional level in the lung has revealed that deformation in the lung tissue can be highly heterogenous and anisotropic (Amelon et al., 2011; Hurtado et al., 2017), thus providing new deformation-based markers to understand lung physiology (Choi et al., 2013), showcasing the potential of DIR strain analysis as a tool in the detection and diagnosis of pulmonary disease. These advances notwithstanding, it has been recently shown that state-of-the-art strain analysis techniques based on direct differentiation of DIR solutions can be highly inaccurate as they are very sensitive to noise, discretization level and embedded anatomical boundaries (Hurtado et al., 2016). While these deficiencies are ameliorated when using L^2 projection smoothing techniques (Hurtado et al., 2016), such approach remains largely heuristic as it lacks of a rigorous mathematical framework that can support the stability and accuracy of the resulting strain fields. This last observation motivates the development of DIR methods that can accurately predict not only the image transformation but also its gradient by proving convergence of the associated numerical scheme.

Motivated by discussion above, the main goal of this work is to present a rigorous formulation and analysis of the registration problem, both for the continuous and discrete settings. Formulations of DIR using elastic regularizers can be interpreted as an elasticity problem with non-linear sources, where well-posedness is handled by requiring Lipschitz continuity and boundedness of the distributional gradient of the similarity-measure density. Taking this approach, the primal formulation of the DIR problem can be conveniently analyzed using Schauder's and Brower's fixed-point theorems for the continuous and discrete formulations, respectively. To directly account for the solution of not only the displacement field, but also of the associated stresses (and displacement gradients thereof), we propose and analyze mixed and augmented formulations of the DIR problem. Finite-element schemes are developed for all the formulations considered in this work, using a $\mathbb{P}_1$ formulation for the primal problem, a $BDM_1 - \mathbb{P}_0$ scheme for the mixed problem, and a $RT_0 - \mathbb{P}_1$ scheme for the augmented problem, all of which can be shown to be stable under the small data assumption.

In what follows, given a scalar expression A, we let $\mathbf{A}$ and $\boldsymbol{A}$ be its vectorial and tensor versions, respectively. In general, the regular font will be used for scalars, bold for vectors and bold slanted for tensors. In the same fashion, we define the usual Sobolev framework with

$$L^2(\Omega) := \{v : \Omega \to \mathbb{R} : \int v^2 < \infty\}, \quad \boldsymbol{L}^2(\Omega) := [L^2(\Omega)]^n, \quad \mathbb{L}^2(\Omega) := [L^2(\Omega)]^{n\times n},$$

and given $m \in \mathbb{N} \cup \{0\}$ use accordingly the spaces $H^m(\Omega), \boldsymbol{H}^m(\Omega)$ and $\mathbb{H}^m(\Omega)$, where each v in these spaces has at least m distributional derivatives in $L^2(\Omega), \boldsymbol{L}^2(\Omega)$ and $\mathbb{L}^2(\Omega)$ respectively. These use the inner product

$$\langle u, v\rangle_{m,\Omega} = \sum_{i=0}^{m} \langle D^i u, D^i v\rangle_{0,\Omega},$$

where $m = 0$ is just the $L^2(\Omega)$, meaningly $H^0 = L^2$. Typically, mixed-FEM schemes employ the space

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

which is a Hilbert space with inner product

$$\langle \boldsymbol{\tau}, \boldsymbol{\sigma}\rangle_{\mathrm{div};\Omega} = \int_\Omega \boldsymbol{\sigma} : \boldsymbol{\tau} + \int_\Omega \mathrm{div}\boldsymbol{\tau} \cdot \mathrm{div}\boldsymbol{\sigma}.$$

Hereafter, div stands for the usual divergence operator acting along the rows of a tensor. Finally, we define the trace operator $\gamma_D(\boldsymbol{u}) = \boldsymbol{u}|_\Gamma$, and the space

$$\boldsymbol{H}^{1/2}(\Gamma) := \left\{\gamma_D(\boldsymbol{u}) : \boldsymbol{u} \in \boldsymbol{H}^1(\Omega)\right\},$$

where its dual is written $\boldsymbol{H}^{-1/2}(\Gamma) = (\boldsymbol{H}^{1/2}(\Gamma))'$ with the usual operator norm. In turn, we let $\gamma_{\boldsymbol{\nu}} : H(\mathrm{div};\Omega) \to H^{-1/2}(\partial\Omega)$ be the normal trace operator on the boundary, which is defined distributionally (see (Gatica, 2014) for more details).

2.2. Elastic deformable image registration (DIR) problem

Let $n \in \{2, 3\}$ be the dimension of the images we are interested in analyzing and $\Omega \subset \mathbb{R}^n$ be a compact domain with Lipschitz boundary $\Gamma := \partial\Omega$. Let $R : \Omega \to \mathbb{R}$ be the reference image and $T : \tilde{\Omega} \to \mathbb{R}$ be the target image with $\Omega \subseteq \tilde{\Omega}$, both in H^1. The requirement that the domain of T is larger than that of R is necessary because in the definition of the registration problem we will need to evaluate T possibly outside Ω. In practice, images used in DIR problems have the same domain, and therefore we consider the underlying image $T_0 : \Omega \to \mathbb{R}$ and construct T by extending it (typically by zero) to $\tilde{\Omega}$. The DIR problem consists in finding a transformation $\boldsymbol{u} : \Omega \to \mathbb{R}^n$, also known as the displacement field, that best aligns the images R and T, which is expressed as a variational problem (Modersitzki, 2004) that reads

$$\inf_{\boldsymbol{u}\in\mathcal{V}} \alpha\,\mathcal{D}[\boldsymbol{u}; R, T] + \mathcal{S}[\boldsymbol{u}], \tag{2.1}$$

where $\mathcal{V}$ is tipically $\boldsymbol{H}^1(\Omega)$, $\mathcal{D} : \mathcal{V} \to \mathbb{R}$ is the similarity measure between the images R and T, $\alpha > 0$ is a weighting constant and $\mathcal{S} : \mathcal{V} \to \mathbb{R}$ is the regularization term, required to render the problem well-posed. A common choice for the similarity measure is the sum of squares difference, i.e, an L^2 error that takes the form

$$\mathcal{D}[\boldsymbol{u}; R, T] := \frac{1}{2}\int_{\Omega} (T(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - R(\boldsymbol{x}))^2. \tag{2.2}$$

For the case of elastic DIR, the regularizing term is commonly taken to be the elastic deformation energy, defined by

$$\mathcal{S}[\boldsymbol{u}] := \frac{1}{2}\int_{\Omega} \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{u}), \tag{2.3}$$

where

$$\boldsymbol{\varepsilon}(\boldsymbol{u}) := \frac{1}{2}(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^{\mathrm{t}}) \tag{2.4}$$

is the infinitesimal strain tensor, i.e., the symmetric component of the displacement field gradient, and $\mathbb{C}$ is the elasticity tensor, which for isotropic solids is defined by the expression

$$\mathbb{C}\boldsymbol{\tau} := \lambda\,\mathrm{trace}(\boldsymbol{\tau})\boldsymbol{I} + n\mu\boldsymbol{\tau}.$$

Assuming that (2.1) has at least one solution with sufficient regularity, the associated Euler-Lagrange equations deliver the following strong problem: Find $\boldsymbol{u} \in \boldsymbol{C}^2(\Omega) \cap \boldsymbol{C}^1(\bar{\Omega})$ such that

$$\begin{aligned} div(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})) &= \alpha\boldsymbol{f}_{\boldsymbol{u}} \quad \text{in} \quad \Omega\,, \\ \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\boldsymbol{n} &= 0 \quad \text{on} \quad \partial\Omega\,, \end{aligned}$$

where

$$\boldsymbol{f}_{\boldsymbol{u}}(\boldsymbol{x}) := \{T(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x})) - R(\boldsymbol{x})\}\nabla T(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x})), \qquad \forall \boldsymbol{x} \in \Omega \ a.e. \tag{2.5}$$

2.3. Primal DIR formulation

Let $\mathcal{V} := \boldsymbol{H}^1(\Omega)$ and define the following forms by considering the registration problem defined in (2.1)

$$\begin{aligned} a(\boldsymbol{u},\boldsymbol{v}) &:= \int \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{v}) \quad && \forall \boldsymbol{u},\boldsymbol{v} \in \mathcal{V}, \\ F_{\boldsymbol{u}}(\boldsymbol{v}) &:= -\int \boldsymbol{f}_{\boldsymbol{u}} \cdot \boldsymbol{v} && \forall \boldsymbol{u},\boldsymbol{v} \in \mathcal{V}, \end{aligned}$$

where $a : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is bilinear and $F_{\boldsymbol{u}} : \mathcal{V} \to R$ is linear for every $\boldsymbol{u}$. We can then write the DIR problem as

$$\min_{\boldsymbol{v}\in\mathcal{V}} \Big\{ \alpha\, \mathcal{D}[\boldsymbol{v}] + a(\boldsymbol{v},\boldsymbol{v}) \Big\}, \tag{2.6}$$

and its Euler-Lagrange equations in weak form give the following problem: Find $\boldsymbol{u} \in \mathcal{V}$ such that

$$a(\boldsymbol{u},\boldsymbol{v}) = \alpha F_{\boldsymbol{u}}(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in \mathcal{V}. \tag{2.7}$$

We will make the following assumptions on the data:

$$\|F_{\boldsymbol{u}} - F_{\boldsymbol{v}}\|_{\mathcal{V}'} \le L_F \|\boldsymbol{u}-\boldsymbol{v}\|_{0,\Omega} \qquad \forall \boldsymbol{u},\boldsymbol{v} \in \mathcal{V}, \tag{2.8}$$

$$\|F_{\boldsymbol{u}}\|_{\mathcal{V}'} \le M_F \qquad \forall \boldsymbol{u} \in \mathcal{V}. \tag{2.9}$$

We remark that assumptions (2.8) and (2.9) are achieved by imposing the following conditions on the nonlinear load term $\boldsymbol{f}_{\boldsymbol{u}}$:

$$\begin{aligned} |\boldsymbol{f}_{\boldsymbol{u}}(\boldsymbol{x}) - \boldsymbol{f}_{\boldsymbol{v}}(\boldsymbol{x})| &\le L_f |\boldsymbol{u}(\boldsymbol{x}) - \boldsymbol{v}(\boldsymbol{x})| \quad && \forall \boldsymbol{x} \in \Omega\, a.e, \\ |\boldsymbol{f}_{\boldsymbol{u}}(\boldsymbol{x})| &\le M_f && \forall \boldsymbol{x} \in \Omega\, a.e. \end{aligned}$$

In addition, we notice that in engineering practice, R, T are interpolations of an array of data (image) where values are defined at the nodes of a Cartesian grid. The most popular interpolation schemes used to construct R and T are cubic B-splines, which implies that $R, T \in C^2$. This in turn implies that the load term is C^1, and therefore locally Lipschitz. This can be extended to the entire domain because Ω is compact.

In the following, we will consider the compact embedding $\boldsymbol{i}_c : \boldsymbol{H}^1(\Omega) \to \boldsymbol{L}^2(\Omega)$ given by Rellich-Kondrachov's theorem. We further recall Schauder's fixed point theorem (see (Ciarlet, 2013) for a proof):

Theorem 2 (Schauder's Fixed Point Theorem)**.** *Let W be a closed and convex subset of a Banach space $\mathcal{V}$ and let $T : W \to W$ be a continuous mapping with the property that $\overline{T(W)}$ is compact. Then T has at least one fixed point in W.*

We define the following partial problem: Given $\boldsymbol{z} \in \mathcal{V}$, find $\boldsymbol{u} \in \mathcal{V}$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) = \alpha F_{\boldsymbol{z}}(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in \mathcal{V}. \tag{2.10}$$

This problem does not have unisolvence, so we will modify the problem by imposing weak orthogonality to the rigid motions space, denoted by $\mathbb{RM}(\Omega)$ and defined as (see (Brenner & Scott, 2008, Eq 11.1.7))

$$\mathbb{RM}(\Omega) := \left\{ \boldsymbol{v} \in H^1(\Omega) : \boldsymbol{\varepsilon}(\boldsymbol{v}) = 0 \right\},$$

which guarantees unisolvence of problem (2.10) since $\mathbb{RM}(\Omega)$ is precisely its null space. Defining $H = \mathbb{RM}(\Omega)^{\perp}$, we define the restricted problem as: Given $\boldsymbol{z} \in H$, find $\boldsymbol{u} \in H$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) = \alpha F_{\boldsymbol{z}}(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in H. \tag{2.11}$$

An application of the Lax-Milgram lemma and Korn's inequality yields the following result:

Theorem 3. *Given $\boldsymbol{z} \in H$, problem* (2.11) *has a unique solution $\boldsymbol{u} \in H$ which satisfies the following a priori estimate for a constant $C > 0$:*

$$\|\boldsymbol{u}\|_{1,\Omega} \leq \alpha\, C\, \|F_{\boldsymbol{z}}\|_{\mathcal{V}'}. \tag{2.12}$$

PROOF. See (Brenner & Scott, 2008, Thm 11.2.30) □

We now define the operator $T : H \to H$ given by

$$T(\boldsymbol{z}) = \boldsymbol{u}, \tag{2.13}$$

where $\boldsymbol{u}$ is the solution to problem (2.11) and thus rewrite problem (2.7) as: Find $\boldsymbol{u}$ such that

$$T(\boldsymbol{u}) = \boldsymbol{u}, \tag{2.14}$$

which shows that the existence of solutions to the original nonlinear problem reduces to the existence of fixed points for the operator T. The following lemmas prove that the conditions required by Schauder's fixed point theorem hold.

Lemma 2. *Let T be an operator defined by* (2.13)*. Then, under data assumption* (2.8)*, there exists $L_T > 0$ such that*

$$\|T(\boldsymbol{z}_1) - T(\boldsymbol{z}_2)\|_{1,\Omega} \leq \alpha C L_F \|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_{0,\Omega} \qquad \forall \boldsymbol{z}_1, \boldsymbol{z}_2 \in H.$$

PROOF. Given $\boldsymbol{z}_1, \boldsymbol{z}_2 \in H$, we let $\boldsymbol{u}_1 := T(\boldsymbol{z}_1)$ and $\boldsymbol{u}_2 := T(\boldsymbol{z}_2)$, that is $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are the unique solutions to the following problems:

$$a(\boldsymbol{u}_1, \boldsymbol{v}) = \alpha F_{\boldsymbol{z}_1}(\boldsymbol{v}) \quad \forall \boldsymbol{v} \in H \;\text{ and }\; a(\boldsymbol{u}_2, \boldsymbol{v}) = \alpha F_{\boldsymbol{z}_2}(\boldsymbol{v}) \quad \forall \boldsymbol{v} \in H.$$

Their difference gives a new problem

$$a(\boldsymbol{u}_1 - \boldsymbol{u}_2, \boldsymbol{v}) = \alpha\,(F_{\boldsymbol{z}_1} - F_{\boldsymbol{z}_2})(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in H,$$

which satisfies the a priori estimate

$$\|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{1,\Omega} \;\leq\; \alpha\, C\, \|F_{\boldsymbol{z}_1} - F_{\boldsymbol{z}_2}\|_{\mathcal{V}'}\,.$$

Using the condition over the data (2.8), we arrive to the desired result:

$$\|T(\boldsymbol{z}_1) - T(\boldsymbol{z}_2)\|_{1,\Omega} = \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{1,\Omega} \;\leq\; \alpha\, C\, \|F_{\boldsymbol{z}_1} - F_{\boldsymbol{z}_2}\|_{\mathcal{V}'} \;\leq\; \alpha\, C\, L_F\, \|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_{0,\Omega},$$

with Lipschitz constant $L_T = \alpha C L_F$. $\square$

Lemma 3. *Let T be an operator defined by* (2.13). *Then, under data assumption* (2.9)*, there exists $r_0 > 0$ such that*

$$T(\overline{B}(0, r_0)) \subseteq \overline{B}(0, r_0) := \Big\{ \boldsymbol{z} \in H : \|\boldsymbol{z}\|_{1,\Omega} \leq r_0 \Big\}.$$

PROOF. We conclude from the a priori estimate. Given $\boldsymbol{z} \in H$, we have from (2.12) and (2.13) that

$$\|T(\boldsymbol{z})\|_{1,\Omega} = \|\boldsymbol{u}\|_{1,\Omega} \leq \alpha C \|F_{\boldsymbol{z}}\|_{\mathcal{V}'} \leq \alpha C M_F,$$

which shows that $T(\boldsymbol{z}) \in \overline{B}(0, r_0)$, with $r_0 = \alpha C M_F$, and hence in particular $T(\overline{B}(0, r_0)) \subseteq \overline{B}(0, r_0)$. $\square$

We are now in a position to prove the existence of solution to the fixed-point problem (2.14).

Theorem 4. *Let T be the operator defined by* (2.13). *Then, under data assumptions* (2.8) *and* (2.9)*, T has at least one fixed point. Moreover, if $\alpha C L_F < 1$, the fixed point is unique.*

PROOF. Let $\{z_j\}_{j=1}^{\infty}$ be a sequence in $B(0, r_0)$ with $r_0 = \alpha C M_F$ as shown in Lemma 3. It follows that there exists a subsequence $\{z_j^{(1)}\}_{j \in \mathbb{N}} \subseteq \{z_j\}_{j \in \mathbb{N}}$ weakly convergent to some $\boldsymbol{z}$ in H. Using the compact embedding $\boldsymbol{i}_c$ we have that $z_j^{(1)} \xrightarrow{j} \boldsymbol{z}$ in $\boldsymbol{L}^2(\Omega)$. In this way, using Lemma 2 we see that $T(z_j^{(1)}) \xrightarrow{j} T(z)$ in H, which means that $\overline{T(\overline{B}(0, r_0)}$ is compact and thus by Schauder's fixed point theorem we conclude the existence of a fixed point. Finally, if $\alpha C L_F < 1$, T is contraction, and so the conclusion is a consequence of Banach's fixed point theorem. $\square$

At this point we observe that, in the context of image registration, the foregoing result shows the existence of solutions to classic schemes such as diffusion and elastic registration together with SSD, cross-correlation or mutual information similarities. Also, data conditions (2.8) and (2.9) give a rule for how much regularity to ask from the images. For example, similarities involving gradients require a Lipschitz gradient, so at least H^2 is required.

2.3.1. Analysis of the discrete problem

In the following we formulate a Galerkin scheme to the primal DIR formulation. To this end, let $(\mathcal{V}_h)_{h>0}$ be a conforming family of discrete spaces indexed by a mesh size $h > 0$. We define $H_h := \mathbb{RM}^\perp \cap \mathcal{V}_h$, and formulate the nonlinear discrete problem as follows: Find $\boldsymbol{u}_h \in H_h$ such that

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) = \alpha F_{\boldsymbol{u}_h}(\boldsymbol{v}_h) \qquad \forall \boldsymbol{v}_h \in H_h. \tag{2.15}$$

Analogously to the continuous case, we consider the (discrete) partial problem: Given $\boldsymbol{z}_h \in H_h$, find $\boldsymbol{u}_h \in H_h$ such that

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) = \alpha F_{\boldsymbol{z}_h}(\boldsymbol{v}_h) \qquad \forall \boldsymbol{v}_h \in H_h, \tag{2.16}$$

and also let $T_h : H_h \to H_h$ be the discrete operator given by

$$T_h(\boldsymbol{z}_h) = \boldsymbol{u}_h,$$

where $\boldsymbol{u}_h$ is the unique solution of problem (2.16) given $\boldsymbol{z}_h$. To prove the existence of fixed points of T_h, we rely on Brouwer's fixed point theorem, which we include next for reference (Ciarlet, 2013):

Theorem 5 (Brouwer's Fixed Point Theorem)**.** *Let K be a compact and convex subset of a finite-dimensional normed vector space and let $T_h : K \to K$ be a continuous mapping. Then T_h has at least one fixed point.*

Considering the same data assumptions as in the continuous case, as well as the continuity and bound obtained before, we arrive at the following result:

Theorem 6. *Let T_h be the discrete operator and assume data conditions* (2.8) *and* (2.9) *hold. Then, T_h has at least one fixed point. Further, if $\alpha C L_F < 1$, the fixed point it is unique.*

PROOF. From the discrete analysis of Lemmas 2 and 3 we know that T_h is continuous and that there exists r_0 such that $T_h : \overline{B}(0, r_0) \to \overline{B}(0, r_0)$. Since H_h is finite-dimensional, $\overline{B}(0, r_0)$ is compact, and so Brouwer's conditions hold, from where we can conclude the existence of a fixed point. Uniqueness is established as in the continuous case with Banach's fixed point theorem. $\square$

Having proved the existence of solutions for the discrete problem, we can show convergence under uniqueness regime, and so we leave Cea's estimate for reference.

Theorem 7. *Let $\boldsymbol{u} \in H$ and $\boldsymbol{u}_h \in H_h$ be the solutions to the continuous and discrete problems* (2.7)*,* (2.15)*. Then, there exist α, $C > 0$ such that*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega} \leq C \inf_{\boldsymbol{v}_h \in H_h} \|\boldsymbol{u} - \boldsymbol{v}_h\|_H. \tag{2.17}$$

PROOF. Strang's first Lemma (see (Steinbach, 2007, Theorem 8.2)) implies that there exists a constant $\tilde{C} > 0$ such that

$$\|\boldsymbol{u} - \boldsymbol{u}_h\| \;\leq\; \tilde{C} \left\{ \inf_{\boldsymbol{v}_h \in H_h} \|\boldsymbol{u} - \boldsymbol{v}_h\|_H \;+\; \alpha \, \|F_{\boldsymbol{u}} - F_{\boldsymbol{u}_h}\|_{\mathcal{V}'} \right\} .$$

Using this inequality, the continuity of the compact embedding $\boldsymbol{i}_c$ and data condition (2.8), we obtain

$$\|\boldsymbol{u} - \boldsymbol{u}_h\| \leq \tilde{C} \inf_{\boldsymbol{v}_h \in H_h} \|\boldsymbol{u} - \boldsymbol{v}_h\|_H + \alpha \tilde{C} L_F \|\boldsymbol{u} - \boldsymbol{u}_h\|_H.$$

Imposing $\alpha \tilde{C} L_F < 1$ gives the desired result for $C = \frac{\tilde{C}}{1 - \alpha \tilde{C} L_F}$. $\square$

Sufficiently small α allows the bound to be independent of it as shown in the following corollary.

Corollary 1. *Under the same assumptions of the previous Theorem, sufficiently small* α *allows* C *to depend only on* $\tilde{C}$.

PROOF. Repeat the argument in Theorem 2.17 assuming $\alpha\,\tilde{C}\,L_F \;\le\; \frac{1}{2}$, then $C \;\ge\; 2\,\tilde{C}$. □

2.4. Mixed DIR formulation

In the following, we introduce a mixed variational formulation for (2.20). Following (Barrientos, Gatica, & Stephan, 2002) and particularizing the method for the 2D case, we note that the constitutive relation can be inverted, to obtain that

$$\mathbb{C}^{-1}\boldsymbol{\sigma} := \frac{1}{2\mu}\boldsymbol{\sigma} - \frac{\lambda}{2\mu(2\mu+n\lambda)}\,\mathrm{trace}(\boldsymbol{\sigma})\boldsymbol{I}. \tag{2.18}$$

We further define an auxiliary field given by the skew symmetric component of the displacement field gradient as

$$\boldsymbol{\rho} := \frac{1}{2}(\nabla\boldsymbol{u} - \nabla\boldsymbol{u}^t). \tag{2.19}$$

We note that from a continuum mechanics perspective, $\boldsymbol{\rho}$ corresponds to the rotation tensor, which accounts for displacement gradients that do not induce deformation energy. Then, the strong form of the mixed registration BVP problem reads: Find $\boldsymbol{u} \in \boldsymbol{C}^1(\Omega)$, $\boldsymbol{\sigma} \in \boldsymbol{C}^1(\Omega) \cap \boldsymbol{C}(\bar{\Omega})$ and $\boldsymbol{\rho} \in \boldsymbol{C}^0_{\mathrm{skew}}(\Omega)$ such that

$$\begin{gathered}\mathbb{C}^{-1}\boldsymbol{\sigma} \;=\; \nabla\boldsymbol{u} - \boldsymbol{\rho} \quad \text{in} \quad \Omega\,, \quad \mathrm{div}\,\boldsymbol{\sigma} \;=\; \alpha\boldsymbol{f}_{\boldsymbol{u}} \quad \text{in} \quad \Omega\,,\\ \boldsymbol{\sigma} \;=\; \boldsymbol{\sigma}^{\mathrm{t}} \quad \text{in} \quad \Omega\,, \quad \boldsymbol{\sigma}\boldsymbol{n} \;=\; \boldsymbol{0} \quad \text{in} \quad \partial\Omega\,,\end{gathered} \tag{2.20}$$

2.4.1. Analysis of the continuous problem

Following the standard integration by parts procedure, the weak variational formulation of the mixed registration problem (2.20) reads: Find $(\boldsymbol{\sigma}, (\boldsymbol{u}, \boldsymbol{\rho})) \in \mathcal{H} \times Q$ such that

$$\begin{aligned} a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, (\boldsymbol{u}, \boldsymbol{\rho})) = 0 \qquad\qquad & \forall \boldsymbol{\tau} \in \mathcal{H}, \\ b(\boldsymbol{\sigma}, (\boldsymbol{v}, \boldsymbol{\eta})) = \alpha F_{\boldsymbol{u}}(\boldsymbol{v}, \boldsymbol{\eta}) \quad & \forall (\boldsymbol{v}, \boldsymbol{\eta}) \in Q. \end{aligned} \tag{2.21}$$

where

$$\begin{aligned} &\mathcal{H} := \boldsymbol{H}_0(\mathrm{div}; \Omega) = \{\boldsymbol{\tau} \in \boldsymbol{H}(\mathrm{div}; \Omega) : \gamma_\nu \boldsymbol{\tau} = 0\}, \\ &Q := \boldsymbol{L}^2(\Omega) \times \boldsymbol{L}^2_{\text{skew}}(\Omega), \end{aligned}$$

and the bilinear forms $a : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ and $b : \mathcal{H} \times Q \to \mathbb{R}$ are defined as follows:

$$\begin{aligned} a(\boldsymbol{\sigma}, \boldsymbol{\tau}) &:= \int_\Omega \mathbb{C}^{-1} \boldsymbol{\sigma} : \boldsymbol{\tau} && \forall\, \boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathcal{H}, \\ b(\boldsymbol{\tau}, (\boldsymbol{v}, \boldsymbol{\eta})) &:= \int_\Omega \boldsymbol{v} \cdot \mathrm{div} \boldsymbol{\tau} + \int_\Omega \boldsymbol{\eta} : \boldsymbol{\tau} && \forall \boldsymbol{\tau} \in \mathcal{H}, (\boldsymbol{v}, \boldsymbol{\eta}) \in Q, \end{aligned}$$

whereas given $\boldsymbol{u} \in \boldsymbol{L}^2(\Omega)$, $F_{\boldsymbol{u}} : Q \to \mathbb{R}$ is the linear functional given by

$$F_{\boldsymbol{u}}(\boldsymbol{v}, \boldsymbol{\eta}) := \alpha \int_\Omega \boldsymbol{f_u} \cdot \boldsymbol{v} \quad \forall (\boldsymbol{v}, \boldsymbol{\eta}) \in Q.$$

To study the solvability of (2.21), we define the partial problem: Given $\boldsymbol{z} \in \boldsymbol{L}^2(\Omega)$, find $(\boldsymbol{\sigma}, (\boldsymbol{u}, \boldsymbol{\rho})) \in \mathcal{H} \times Q$ such that

$$\begin{aligned} a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, (\boldsymbol{u}, \boldsymbol{\rho})) = 0 \qquad\qquad & \forall \boldsymbol{\tau} \in \mathcal{H}, \\ b(\boldsymbol{\sigma}, (\boldsymbol{v}, \boldsymbol{\eta})) = \alpha F_{\boldsymbol{z}}(\boldsymbol{v}, \boldsymbol{\eta}) \quad & \forall (\boldsymbol{v}, \boldsymbol{\eta}) \in Q, \end{aligned} \tag{2.22}$$

which is a linear elasticity problem with Neumann boundary conditions. This problem does not have unisolvence, so we impose orthogonality to the rigid motions space $\mathbb{RM}$ weakly. With this consideration, we define $H := \mathcal{H} \times \mathbb{RM}(\Omega)$, as well as the bilinear forms

$$A((\boldsymbol{\sigma},\boldsymbol{\rho}),(\boldsymbol{\tau},\boldsymbol{\xi})) := a(\boldsymbol{\sigma},\boldsymbol{\tau}) \qquad \forall(\boldsymbol{\sigma},\boldsymbol{\rho}),(\boldsymbol{\tau},\boldsymbol{\xi}) \in H$$

$$B((\boldsymbol{\tau},\boldsymbol{\xi}),(\boldsymbol{v},\boldsymbol{\eta})) := b(\boldsymbol{\tau},(\boldsymbol{v},\boldsymbol{\eta})) + \int_\Omega \boldsymbol{\xi}\cdot\boldsymbol{v} \quad \forall((\boldsymbol{\tau},\boldsymbol{\xi}),(\boldsymbol{v},\boldsymbol{\eta})) \in H \times \mathcal{Q}.$$

We then consider the following equivalent mixed variational formulation: Given $\boldsymbol{z} \in \boldsymbol{L}^2(\Omega)$, find $((\boldsymbol{\sigma},\boldsymbol{\rho}),(\boldsymbol{u},\boldsymbol{\gamma})) \in H \times Q$ such that

$$\begin{aligned} A((\boldsymbol{\sigma},\boldsymbol{\rho}),(\boldsymbol{\tau},\boldsymbol{\xi})) + B((\boldsymbol{\tau},\boldsymbol{\xi}),(\boldsymbol{u},\boldsymbol{\gamma})) &= 0 && \forall(\boldsymbol{\tau},\boldsymbol{\xi}) \in H, \\ B((\boldsymbol{\sigma},\boldsymbol{\rho}),(\boldsymbol{v},\boldsymbol{\eta})) &= \alpha F_{\boldsymbol{z}}((\boldsymbol{v},\boldsymbol{\eta})) && \forall(\boldsymbol{v},\boldsymbol{\eta}) \in Q. \end{aligned} \tag{2.23}$$

This formulation delivers a well-posed problem, as proved in (Gatica et al., 2008, Theorem 3.1), and thus it allows for the definition of a fixed-point operator. Let $\boldsymbol{T} : \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$ given by

$$\boldsymbol{T}(\boldsymbol{z}) := \boldsymbol{u} \qquad \forall \boldsymbol{z} \in \boldsymbol{L}^2(\Omega), \tag{2.24}$$

where $\boldsymbol{u}$ is the displacement component of the unique solution of problem (2.23), and so the mixed formulation can be restated as: Find $\boldsymbol{u} \in \boldsymbol{L}^2(\Omega)$ such that

$$\boldsymbol{T}(\boldsymbol{u}) = \boldsymbol{u}. \tag{2.25}$$

Note here that if $((\boldsymbol{\tau},\boldsymbol{\xi}),(\boldsymbol{v},\boldsymbol{\eta})) = ((\mathbf{o},\boldsymbol{\xi}),(\boldsymbol{\xi},\nabla\boldsymbol{\xi}))$, we obtain

$$\int \boldsymbol{\rho}\cdot\boldsymbol{\xi} = \int \boldsymbol{f}_{\boldsymbol{z}}\cdot\boldsymbol{\xi} \quad \forall \boldsymbol{\xi} \in \mathbb{RM},$$

so the Lagrange multiplier $\boldsymbol{\rho}$ removes the compatibility requirement from the data.

To prove the existence of a solution to problem (2.25) we use Banach's fixed point theorem for a sufficiently small α and data conditions (2.9) and (2.8). The continuity of the operator $\boldsymbol{T}$ is established in the following lemma.

Lemma 4. *Let the operator $\boldsymbol{T}$ defined by* (2.24)*, assume the data condition* (2.8)*, and let $C > 0$ be the constant of continuous dependence on data of* (2.23)*. Then there holds:*

$$\|\boldsymbol{T}(\boldsymbol{z}_1) - \boldsymbol{T}(\boldsymbol{z}_2)\|_{0,\Omega} \le \alpha C L_F \|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_{0,\Omega} \qquad \forall \boldsymbol{z}_1, \boldsymbol{z}_2 \in \boldsymbol{L}^2(\Omega). \tag{2.26}$$

PROOF. Set $\boldsymbol{u}_i = \boldsymbol{T}(\boldsymbol{z}_i), i \in \{1,2\}$. Substracting terms we arrive at the following system of equations:

$$A((\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \boldsymbol{\rho}_1 - \boldsymbol{\rho}_2), (\boldsymbol{\tau}, \boldsymbol{\xi})) + B((\boldsymbol{\tau}, \boldsymbol{\xi}), (\boldsymbol{u}_1 - \boldsymbol{u}_2, \boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2)) = 0 \qquad \forall (\boldsymbol{\tau}, \boldsymbol{\xi}) \in H,$$
$$B((\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \boldsymbol{\rho}_1 - \boldsymbol{\rho}_2), (\boldsymbol{v}, \boldsymbol{\eta})) = \alpha(F_{\boldsymbol{z}_1} - F_{\boldsymbol{z}_2})(\boldsymbol{v}, \boldsymbol{\eta}) \quad \forall (\boldsymbol{v}, \boldsymbol{\eta}) \in Q,$$

from which, using the continuous dependence of (2.23) and the Lipschitz continuity of F, we get

$$\|\boldsymbol{T}(\boldsymbol{z}_1) - \boldsymbol{T}(\boldsymbol{z}_2)\|_{0,\Omega} = \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{0,\Omega} \leq \alpha C \, \|F_{\boldsymbol{z}_1} - F_{\boldsymbol{z}_2}\|_{Q'} \leq \alpha C \, L_F \, \|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_{0,\Omega},$$

thus completing the proof. □

Theorem 8. *Under data conditions* (2.8), (2.9) *and assuming* $\alpha C L_F < 1$*, there is a unique fixed point for problem* (2.25)*. With this, the mixed formulation* (2.21) *has a unique solution and the a priori estimation*

$$\|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma}))\|_H \, \leq \, \alpha C M_F$$

holds.

PROOF. From Lemma 4 we have that

$$\|\boldsymbol{T}(\boldsymbol{z}_1) - \boldsymbol{T}(\boldsymbol{z}_2)\|_{0,\Omega} \leq \alpha C L_F \|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_{0,\Omega} \qquad \forall \boldsymbol{z}_1, \boldsymbol{z}_2 \in \boldsymbol{L}^2(\Omega),$$

which thanks to the assumption $\alpha \, C \, L_F < 1$ makes $\boldsymbol{T}$ a contraction, and thus prove the existence of the unique fixed point by virtue of Banach's fixed point theorem. For the a priori bound, we use the same estimate of the partial problem setting $\boldsymbol{z} = \boldsymbol{u}$ and condition (2.9), which gives

$$\|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma}))\|_H \leq \alpha C \|F_{\boldsymbol{u}}\|_{Q'} \leq \alpha C M_F,$$

which concludes the proof. □

2.4.2. Analysis of the discrete problem

In this section we analyze a Galerkin scheme for problem (2.21). Let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangularizations of $\overline{\Omega}$ of characteristic size h, and the following set of inf-sup stable discrete spaces:

$$
\begin{aligned}
H_h^{\boldsymbol{\sigma}} &:= \left\{ \boldsymbol{\tau}_h \in \boldsymbol{H}(\mathrm{div};\Omega) : \boldsymbol{\tau}_{h,i}|_T \in BDM_1(T) \quad \forall T \in \mathcal{T}_h \right\}, \\
Q_h^{\boldsymbol{u}} &:= \left\{ \boldsymbol{v}_h \in \boldsymbol{L}^2(\Omega) : \boldsymbol{v}_h \in [\mathbb{P}_0(T)]^n \quad \forall T \in \mathcal{T}_h \right\}, \\
Q_h^{\boldsymbol{\gamma}} &:= \left\{ \begin{bmatrix} 0 & \psi_h \\ -\psi_h & 0 \end{bmatrix} : \psi_h \in \mathbb{P}_1(T) \quad \forall T \in \mathcal{T}_h \right\},
\end{aligned}
$$

where $BDM_k = [\mathbb{P}_k]^n$ are the Brezzi-Douglas-Marini elements (Brezzi, Douglas, & Marini, 1984) and $\boldsymbol{\tau}_{h,i}$ is the ith row of $\boldsymbol{\tau}_h$. Then, we introduce

$$
H_{0,h}^{\boldsymbol{\sigma}} := H_h^{\boldsymbol{\sigma}} \cap \boldsymbol{H}_0(\mathrm{div};\Omega), \qquad H_h = H_{0,h}^{\boldsymbol{\sigma}} \times \mathbb{RM}(\Omega),
$$

and

$$
Q_h = Q_h^{\boldsymbol{u}} \times Q_h^{\boldsymbol{\gamma}}.
$$

Thus, we define the discrete version of (2.23) as follows: Given $\boldsymbol{z}_h \in Q_h^{\boldsymbol{u}}$, find $((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h)) \in H_h \times Q_h$ such that

$$
\begin{aligned}
A((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h)) + B((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h)) &= 0 \qquad \forall (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in H_h \\
B((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{v}_h, \boldsymbol{\eta}_h)) &= \alpha F_{\boldsymbol{z}_h}((\boldsymbol{v}_h, \boldsymbol{\eta}_h)) \quad \forall (\boldsymbol{v}_h, \boldsymbol{\eta}_h) \in Q_h .
\end{aligned}
\tag{2.27}
$$

The unique solvability and stability of (2.27), being the Galerkin scheme of a linear elasticity problem with Neumann boundary conditions has already been established in (Gatica et al., 2008, Theorem 4.1). This allows us to define the discrete operator $\boldsymbol{T}_h : Q_h^{\boldsymbol{u}} \to Q_h^{\boldsymbol{u}}$, given by

$$
\boldsymbol{T}_h(\boldsymbol{z}_h) = \boldsymbol{u}_h,
$$

and then rewrite the discretized nonlinear problem as: Find $\boldsymbol{u}_h \in Q_h^{\boldsymbol{u}}$ such that

$$\boldsymbol{T}_h(\boldsymbol{u}_h) = \boldsymbol{u}_h. \tag{2.28}$$

Concerning the other components $\boldsymbol{\sigma}_h, \boldsymbol{\gamma}_h$ and $\boldsymbol{\rho}_h$ of this problem, we will refer to $((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h))$ as the mixed solution for a given $\boldsymbol{z}_h$.

Lemma 5. *Let $\boldsymbol{T}_h$ be the discrete operator given by* (2.28) *and C be the constant of continuous dependence on data of* (2.27)*. Then, given $\boldsymbol{z}_1, \boldsymbol{z}_2$ in $Q_h^{\boldsymbol{u}}$, there holds*

$$\|\boldsymbol{T}_h(\boldsymbol{z}_1) - \boldsymbol{T}_h(\boldsymbol{z}_2)\|_{0,\Omega} \leq \alpha C L_F \|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_{0,\Omega}.$$

PROOF. Repeat argument in Lemma 4 to $\boldsymbol{T}_h$. □

Now we are in position to establish the well-posedness of problem (2.28), as well as to prove the Cea's best approximation property.

Theorem 9. *Under data assumptions* (2.8), (2.9) *and assuming $\alpha\, C\, L_F < 1$, the problem* (2.28) *has a unique solution $((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h)) \in H_h \times Q_h$ such that*

$$\|((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h))\|_{H\times Q} \leq \alpha C M_F.$$

In addition, there exists $\tilde{C} > 0$ such that Cea's estimate holds for the unique solution $((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma}))$ to problem (2.25)*, i.e,*

$$\|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h))\|_{H\times Q} \leq \tilde{C} \inf_{\substack{(\boldsymbol{\tau}_h, \mathbf{0}) \in H_h, \\ (\boldsymbol{v}_h, \boldsymbol{\eta}_h) \in Q_h}} \|((\boldsymbol{\sigma}, \mathbf{0}), (\boldsymbol{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\tau}_h, \mathbf{0}), (\boldsymbol{v}_h, \boldsymbol{\eta}_h))\|_{H\times Q}.$$

PROOF. The first part is analogous to the continuous case, so we are only left with Cea's estimate. Let $((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma})) \in H \times Q$ and $((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h)) \in H_h \times Q_h$ be the solutions arising from the continuous and discrete problems (2.25) and (2.28), respectively. Equivalently, $((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma}))$ (resp. $((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h))$) solves (2.23) with $\boldsymbol{z} = \boldsymbol{u}$ (resp. (2.27) with $\boldsymbol{z}_h = \boldsymbol{u}_h$). In addition, let also $((\boldsymbol{\zeta}_h, \boldsymbol{\varphi}_h), (\boldsymbol{w}_h, \boldsymbol{s}_h)) \in H_h \times Q_h$ be the solution to the discrete version of (2.23) with $\boldsymbol{z} = \boldsymbol{u}$, but without changing $\boldsymbol{u}$ by $\boldsymbol{u}_h$, that is

$$\begin{aligned} A((\boldsymbol{\zeta}_h, \boldsymbol{\varphi}_h), (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h)) + B((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\boldsymbol{w}_h, \boldsymbol{s}_h)) &= 0 & \forall (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in H_h \\ B((\boldsymbol{\zeta}_h, \boldsymbol{\varphi}_h), (\boldsymbol{v}_h, \boldsymbol{\eta}_h)) &= \alpha F_{\boldsymbol{u}}((\boldsymbol{v}_h, \boldsymbol{\eta}_h)) & \forall (\boldsymbol{v}_h, \boldsymbol{\eta}_h) \in Q_h \,. \end{aligned} \tag{2.29}$$

By virtue of (Gatica, 2014, Theorem 2.6), the following estimate holds for a constant $\bar{C} > 0$, which depends only on the bilinear forms A, B:

$$\begin{aligned} &\|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\zeta}_h, \boldsymbol{\varphi}_h), (\boldsymbol{w}_h, \boldsymbol{s}_h))\|_{H\times Q} \\ &\qquad \leq \; \bar{C} \inf_{\substack{(\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in H_h, \\ (\boldsymbol{v}_h, \boldsymbol{\eta}_h) \in Q_h}} \|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\boldsymbol{v}_h, \boldsymbol{\eta}_h))\|_{H\times Q} \,. \end{aligned} \tag{2.30}$$

Substracting systems (2.27) with $\boldsymbol{z}_h = \boldsymbol{u}_h$ and (2.29) it holds that

$$\begin{aligned} A((\boldsymbol{\sigma}_h - \boldsymbol{\zeta}_h, \boldsymbol{\rho}_h - \boldsymbol{\varphi}_h), (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h)) + B((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\boldsymbol{u}_h - \boldsymbol{w}_h, \boldsymbol{\gamma}_h - \boldsymbol{s}_h)) &= 0 \\ B((\boldsymbol{\sigma}_h - \boldsymbol{\zeta}_h, \boldsymbol{\rho}_h - \boldsymbol{\varphi}_h), (\boldsymbol{v}_h, \boldsymbol{\eta}_h)) &= \alpha (F_{\boldsymbol{u}_h} - F_{\boldsymbol{u}})((\boldsymbol{v}_h, \boldsymbol{\eta}_h)), \end{aligned} \tag{2.31}$$

for all $(\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in H_h, (\boldsymbol{v}_h, \boldsymbol{\eta}_h) \in Q_h$, and once again the Babuška-Brezzi theory gives the estimate

$$\|((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h)) - ((\boldsymbol{\zeta}_h, \boldsymbol{\varphi}_h), (\boldsymbol{w}_h, \boldsymbol{s}_h))\|_{H\times Q} \;\leq\; \alpha\, C\, \|F_{\boldsymbol{u}} - F_{\boldsymbol{u}_h}\|_{Q'} \,. \tag{2.32}$$

Finally, using (2.30) and (2.32) we obtain the following Strang-type estimate

$$\begin{aligned} &\|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h))\|_{H\times Q} \\ &\quad \leq \|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\zeta}_h, \boldsymbol{\varphi}_h), (\boldsymbol{w}_h, \boldsymbol{s}_h))\|_{H\times Q} + \|((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h)) - ((\boldsymbol{\zeta}_h, \boldsymbol{\varphi}_h), (\boldsymbol{w}_h, \boldsymbol{s}_h))\|_{H\times Q} \\ &\quad \leq \bar{C} \inf_{\substack{(\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in H_h, \\ (\boldsymbol{v}_h, \boldsymbol{\eta}_h) \in Q_h}} \|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\boldsymbol{v}_h, \boldsymbol{\eta}_h))\|_{H\times Q} + \alpha C \|F_{\boldsymbol{u}} - F_{\boldsymbol{u}_h}\|_{Q'} \,. \end{aligned}$$

In this way, using the data condition (2.8), and assuming that $\alpha\, C\, L_F < 1$, we find that

$$\begin{aligned}
&\|((\boldsymbol{\sigma},\boldsymbol{\rho}),(\boldsymbol{u},\boldsymbol{\gamma})) - ((\boldsymbol{\sigma}_h,\boldsymbol{\rho}_h),(\boldsymbol{u}_h,\boldsymbol{\gamma}_h))\|_{H\times Q} \\
&\qquad \le\ \tilde{C} \inf_{\substack{(\boldsymbol{\tau}_h,\boldsymbol{\xi}_h)\in H_h,\\ (\boldsymbol{v}_h,\boldsymbol{\eta}_h)\in Q_h}} \|((\boldsymbol{\sigma},\boldsymbol{\rho}),(\boldsymbol{u},\boldsymbol{\gamma})) - ((\boldsymbol{\tau}_h,\boldsymbol{\xi}_h),(\boldsymbol{v}_h,\boldsymbol{\eta}_h))\|_{H\times Q}\,,
\end{aligned}$$

where $\tilde{C} = \frac{\bar{C}}{1-\alpha C L_F}$. The absence of the rigid motion variable is due to the space not being discretized, and thus we have that $d(\boldsymbol{\rho}, \mathbb{RM}(\Omega)) = 0$. □

Notice that the proof used to establish this Strang-type estimate is analogous to the one employed in the primal case in (Steinbach, 2007, Theorem 8.2).

Corollary 2. *The above Theorem holds for $\tilde{C} = 2\bar{C}$ provided α is sufficiently small.*

PROOF. In the preceding proof, take $\alpha\, C\, L_F\ \le\ \frac{1}{2}$. □

We end this section by providing the rate of convergence of the solution to (2.28). We first recall the classic approximation results from (Brezzi & Fortin, 1991) and proceed as in (Gatica et al., 2008):

$(AP_h^{\boldsymbol{\sigma}})$ For each $\boldsymbol{\tau} \in \boldsymbol{H}^1(\Omega) \cap \boldsymbol{H}_0(\mathrm{div};\Omega)$ with $\mathrm{div}\boldsymbol{\tau} \in \boldsymbol{H}^1(\Omega)$ there exists $\boldsymbol{\tau}_h \in H_{0,h}^{\boldsymbol{\sigma}}$ such that

$$\|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{\mathrm{div};\Omega} \le Ch\{\|\boldsymbol{\tau}\|_{1,\Omega} + \|\mathrm{div}\boldsymbol{\tau}\|_{1,\Omega}\}.$$

$(AP_h^{\boldsymbol{u}})$ For each $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ there there exists $\boldsymbol{v}_h \in H_h^{\boldsymbol{u}}$ such that

$$\|\boldsymbol{v} - \boldsymbol{v}_h\|_{0,\Omega} \le Ch\|\boldsymbol{v}\|_{1,\Omega}.$$

$(AP_h^{\boldsymbol{\gamma}})$ For each $\boldsymbol{\eta} \in \boldsymbol{H}^1(\Omega) \cap \boldsymbol{L}^2_{asym}(\Omega)$ there exists $\boldsymbol{\eta}_h \in H_h^{\boldsymbol{\gamma}}$ such that

$$\|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{0,\Omega} \le Ch\|\boldsymbol{\eta}\|_{1,\Omega}.$$

These allow us to establish the following theorem:

Theorem 10. *Under data assumptions* (2.8), (2.9) *and assuming* $\alpha C L_F \leq \frac{1}{2}$, *let* $((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h)) \in H_h \times Q_h$ *be the mixed solution of* (2.28) *and* $((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma})) \in H \times Q$ *the solution of the continuous mixed problem* (2.25). *Then, there exists a constant* $C > 0$, *independent of* h, *such that whenever* $\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega), \operatorname{div}\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega), \boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ *and* $\boldsymbol{\gamma} \in \boldsymbol{H}^1(\Omega)$, *there holds*

$$\|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h))\|_{H \times Q} \leq Ch \left\{ \|\boldsymbol{\tau}\|_{1,\Omega} + \|\operatorname{div}\boldsymbol{\tau}\|_{1,\Omega} + \|\boldsymbol{v}\|_{1,\Omega} + \|\boldsymbol{\eta}\|_{1,\Omega} \right\}.$$

PROOF. It follows from applying Cea's estimate and the approximation properties $(AP_h^{\boldsymbol{\sigma}}), (AP_h^{\boldsymbol{u}})$ and $(AP_h^{\boldsymbol{\gamma}})$. $\square$

2.5. Augmented DIR formulation

This section proposes an augmented mixed variational formulation for the BVP (2.20). This scheme gives additional regularity to the displacement field, allows the use of a weaker fixed point theorem guaranteeing the existence of solutions for any positive α, and permits more flexibility for the choice of the finite element subspaces. Again, sufficiently small α grants uniqueness. It has been studied in (Gatica, 2006) that the augmented Dirichlet and mixed-boundary condition linear elasticity problems have a unique solution, and here we prove that the same holds for the null traction problem.

2.5.1. Analysis of the continuous problem

Let $H := \boldsymbol{H}(\mathrm{div};\Omega)\times\boldsymbol{H}^1(\Omega)\times\boldsymbol{L}^2_{\mathrm{skew}}$, $Q := \mathbb{RM}(\Omega)$. We further define the operators $A : H\times H\to\mathbb{R}$, $B : H\times Q\to\mathbb{R}$ and $F_{\boldsymbol{z}}\in H'$ as follows

$$
\begin{aligned}
A((\boldsymbol{\sigma},\boldsymbol{u},\boldsymbol{\gamma}),(\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta})) &:= a(\boldsymbol{\sigma},\boldsymbol{\tau}) + b(\boldsymbol{\tau},(\boldsymbol{u},\boldsymbol{\gamma})) - b(\boldsymbol{\sigma},(\boldsymbol{v},\boldsymbol{\eta}))\\
&\quad + \kappa_1\int_\Omega (\boldsymbol{\varepsilon}(\boldsymbol{u})) - \mathbb{C}^{-1}\boldsymbol{\sigma}) : (\boldsymbol{\varepsilon}(\boldsymbol{v}) + \mathbb{C}^{-1}\boldsymbol{\tau})\\
&\quad + \kappa_2\int_\Omega \mathrm{div}\boldsymbol{\sigma}\cdot\mathrm{div}\boldsymbol{\tau}\\
&\quad + \kappa_3\int_\Omega (\boldsymbol{\gamma} - \mathrm{skew}\,\nabla\boldsymbol{u}) : (\boldsymbol{\eta} + \mathrm{skew}\,\nabla\boldsymbol{v}),\\
B((\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta}),\boldsymbol{\xi}) &:= \int_\Omega \boldsymbol{v}\cdot\boldsymbol{\xi},\\
F_{\boldsymbol{z}}((\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta})) &:= \int_\Omega \boldsymbol{f}_{\boldsymbol{z}}\cdot(-\boldsymbol{v} + \kappa_2\mathrm{div}\boldsymbol{\tau})\,,
\end{aligned}
\tag{2.33}
$$

where κ_1, κ_2, and κ_3 are positive parameters to be suitably chosen later on. With the definitions above, the augmented formulation reads: Given $\boldsymbol{z}\in\boldsymbol{H}^1(\Omega)$, find $((\boldsymbol{\sigma},\boldsymbol{u},\boldsymbol{\gamma}),\boldsymbol{\rho})\in H\times Q$ such that:

$$
\begin{aligned}
A((\boldsymbol{\sigma},\boldsymbol{u},\boldsymbol{\gamma}),(\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta})) + B((\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta}),\boldsymbol{\rho}) &= \alpha F_{\boldsymbol{z}}((\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta})) \quad && \forall(\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta})\in H,\\
B((\boldsymbol{\sigma},\boldsymbol{u},\boldsymbol{\gamma}),\boldsymbol{\xi}) &= 0 && \forall\boldsymbol{\xi}\in Q.
\end{aligned}
$$

Now we define the augmented fixed point operator $\boldsymbol{T} : \boldsymbol{H}^1(\Omega)\to\boldsymbol{H}^1(\Omega)$ given by $\boldsymbol{T}(\boldsymbol{z}) = \boldsymbol{u}$, where $\boldsymbol{u}$ is the displacement component of the solution to problem (2.34), and so we write the nonlinear augmented problem as: Find $\boldsymbol{u}\in\boldsymbol{H}^1(\Omega)$ such that

$$
\boldsymbol{T}(\boldsymbol{u}) = \boldsymbol{u}. \tag{2.34}
$$

We apply the Babuška-Brezzi conditions to see that the proposed problem has a unique solution and depends continuously on the data. We will use the following definitions:

$$
\boldsymbol{\tau}^d = \boldsymbol{\tau} - \frac{1}{n}\,\mathrm{trace}(\boldsymbol{\tau})\boldsymbol{I} \quad\text{and}\quad \boldsymbol{\tau}_0 = \boldsymbol{\tau} - \frac{1}{|\Omega|}\int \mathrm{trace}(\boldsymbol{\tau})\boldsymbol{I},
$$

where the first one is known as the deviatoric tensor, and is such that it has null trace. The second one is the decomposition that arises from $\boldsymbol{H}(\mathrm{div};\Omega) = \tilde{\boldsymbol{H}}(\mathrm{div};\Omega) \oplus \mathbb{R}\boldsymbol{I}$, where $\boldsymbol{I}$ is the identity tensor and

$$\tilde{\boldsymbol{H}}(\mathrm{div};\Omega) := \left\{\boldsymbol{\tau} \in \boldsymbol{H}(\mathrm{div};\Omega) : \int \mathrm{trace}\,\boldsymbol{\tau} = 0\right\}.$$

We will use the following lemmas:

Lemma 6. *There exists $C_1 > 0$ such that*

$$\|\boldsymbol{\tau}^d\|_{0,\Omega}^2 + \|\mathrm{div}\boldsymbol{\tau}\|_{0,\Omega}^2 \geq C_1\|\boldsymbol{\tau}\|_{\mathrm{div};\Omega}^2 \qquad \forall \boldsymbol{\tau} \in \boldsymbol{H}_0(\mathrm{div};\Omega). \tag{2.35}$$

PROOF. We start from (Gatica, 2006, Lemma 2.1), which guarantees the existence of $c_1 > 0$ such that

$$\|\mathrm{div}\boldsymbol{\tau}\|_{0,\Omega}^2 + \|\boldsymbol{\tau}^d\|_{0,\Omega}^2 \geq c_1\|\boldsymbol{\tau}_0\|_{0,\Omega}^2 \qquad \forall \boldsymbol{\tau} \in H(\mathrm{div};\Omega)\,.$$

Then, noting that $\boldsymbol{\tau}_0 = \boldsymbol{\tau}_0^d$ and $\mathrm{div}\boldsymbol{\tau} = \mathrm{div}\boldsymbol{\tau}_0$, we readily find that

$$\|\boldsymbol{\tau}^d\|_{0,\Omega}^2 + \|\mathrm{div}\boldsymbol{\tau}\|_{0,\Omega}^2 \geq \frac{c_1}{2}\|\boldsymbol{\tau}_0\|_{0,\Omega}^2 + \frac{1}{2}\|\mathrm{div}\boldsymbol{\tau}\|_{0,\Omega}^2 \geq \frac{1}{2}\min\{c_1, 1\}\,\|\boldsymbol{\tau}_0\|_{\mathrm{div};\Omega}^2. \tag{2.36}$$

Now, writing explicitly $\boldsymbol{\tau} = \boldsymbol{\tau}_0 + d\boldsymbol{I}$ we impose the null normal trace condition from $0 = \gamma_\nu(\boldsymbol{\tau}) = \gamma_\nu(\boldsymbol{\tau}_0 + d\boldsymbol{I})$ and obtain

$$|d|\|\nu\|_{-1/2,\Gamma} = \|\gamma_\nu\boldsymbol{\tau}_0\|_{-1/2,\Omega} \leq \|\boldsymbol{\tau}_0\|_{\mathrm{div};\Omega}\,,$$

or equivalently,

$$|d| \leq \frac{1}{\|\nu\|_{-1/2,\Gamma}}\|\boldsymbol{\tau}_0\|_{\mathrm{div};\Omega},$$

where the continuity of the normal trace was used. Then, aplying this to the norm of $\boldsymbol{\tau}$ we get

$$\|\boldsymbol{\tau}\|_{\mathrm{div};\Omega} = \|\boldsymbol{\tau}_0\|_{\mathrm{div};\Omega} + n|d||\Omega| \leq \left(1 + \frac{n|\Omega|}{\|\nu\|_{-1/2,\Gamma}}\right)\|\boldsymbol{\tau}_0\|_{\mathrm{div};\Omega},$$

and everything together in equation (2.36) gives the desired result, with $C_1 = 2\dfrac{\left(1+\frac{n|\Omega|}{\|\nu\|_{-1/2,\Gamma}}\right)^2}{\min\{1,c_1\}}$.

$\square$

Lemma 7. *There exists $\widetilde{\kappa} > 0$ such that :*

$$\|\boldsymbol{\varepsilon}(\boldsymbol{v})\|_{0,\Omega}^2 \,\geq\, \widetilde{\kappa}\,\|\boldsymbol{v}\|_{1,\Omega}^2 \qquad \forall \boldsymbol{v} \in \mathbb{RM}^{\perp}. \tag{2.37}$$

PROOF. The proof is a combination of (Brenner & Scott, 2008, Theorem 9.2.12) (Korn's second inequality) and the fact that $H^1 = \widehat{H}^1 \oplus \mathbb{RM}$, where

$$\widehat{H}^1(\Omega) = \left\{ v \in H^1(\Omega) : \int_\Omega v = 0, \int \operatorname{rot} v = 0 \right\}.$$

$\square$

Theorem 11. *Let $V := N(\mathbb{B})$, where $\mathbb{B}$ is the operator induced by the bilinear form B defined in* (2.34)*. Then, there exist κ_1, κ_2 and κ_3 for* (2.33) *such that*

(i) *the bilinear form A is V-elliptic,*

(ii) *the bilinear form B satisfies the inf-sup condition,*

so that there exists a unique solution $((\boldsymbol{\sigma},\boldsymbol{u},\boldsymbol{\gamma}),\boldsymbol{\xi}) \in H \times Q$ to the problem (2.34) *for a given $\boldsymbol{z} \in \boldsymbol{H}^1(\Omega)$. In addition, there exists a constant $C > 0$ such that*

$$\|((\boldsymbol{\sigma},\boldsymbol{u},\boldsymbol{\gamma}),\boldsymbol{\xi})\|_{H\times Q} \,\leq\, \alpha\, C\,\|F_{\boldsymbol{z}}\|_{H'}\,.$$

PROOF. It is important to mention that we are considering rigid motions endowed with the $L^2(\Omega)$ inner product, and that the main idea is to find values for $\kappa_1, \kappa_2, \kappa_3$ (cf. (2.33)) such that ellipticity holds. First we prove ellipticity in V, where the norm used will be the following:

$$\|(\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta})\|_V^2 := \|\boldsymbol{\tau}\|_{\mathrm{div};\Omega}^2 + \|\boldsymbol{v}\|_{1,\Omega}^2 + \|\boldsymbol{\eta}\|_{0,\Omega}^2.$$

In fact, since $V := \{(\boldsymbol{\sigma},\boldsymbol{u},\boldsymbol{\gamma}) \in H : B((\boldsymbol{\sigma},\boldsymbol{u},\boldsymbol{\gamma}),\boldsymbol{\xi}) = 0 \quad \forall \boldsymbol{\xi} \in Q\}$, we find that

$$V \,=\, \{(\boldsymbol{\sigma},\boldsymbol{u},\boldsymbol{\gamma}) \in H : \int \boldsymbol{u}\cdot\boldsymbol{\xi} = 0 \quad \forall \boldsymbol{\xi} \in Q\} \,=\, \{(\boldsymbol{\sigma},\boldsymbol{u},\boldsymbol{\gamma}) \in H : \boldsymbol{u} \in \mathbb{RM}^{\perp}\}\,.$$

In turn, for the ellipticity we are seeking, we use from (Gatica, 2006, Theorem 3.1) that

$$
\begin{aligned}
A((\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta}),(\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta})) = \int_{\Omega} \mathbb{C}^{-1}\boldsymbol{\tau} : \boldsymbol{\tau} - \kappa_1 \|\mathbb{C}^{-1}\boldsymbol{\tau}\|_{0,\Omega}^2 + (\kappa_1+\kappa_3)\|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega}^2 \\
+\kappa_2\|\mathrm{div}\boldsymbol{\tau}\|_{0,\Omega}^2 + \kappa_3\|\boldsymbol{\eta}\|_{0,\Omega}^2 - \kappa_3|\boldsymbol{v}|_{1,\Omega}^2 \,,
\end{aligned} \tag{2.38}
$$

and that

$$
\int_{\Omega} \mathbb{C}^{-1}\boldsymbol{\tau} : \boldsymbol{\tau} - \kappa_1 \|\mathbb{C}^{-1}\boldsymbol{\tau}\|_{0,\Omega}^2 \geq \frac{1}{2\mu}\left(1-\frac{\kappa_1}{2\mu}\right)\|\boldsymbol{\tau}^d\|_{0,\Omega}^2 \,, \tag{2.39}
$$

which require $0 < \kappa_1 < 2\mu$. Then, proceeding as in (Gatica, 2006, Section 3.2), that is applying Lemma 6 and Korn's second inequality (cf. Lemma 7), it follows from (2.38) and (2.39) that

$$
A((\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta}),(\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta})) \;\geq\; \alpha_1\,\|\boldsymbol{\tau}\|_{\mathrm{div};\Omega}^2 + \left\{\kappa_1\,\widetilde{\kappa} - \kappa_3\,(1-\widetilde{\kappa})\right\}\|\boldsymbol{v}\|_{1,\Omega}^2 + \kappa_3\|\boldsymbol{\eta}\|_{0,\Omega}^2 \,, \tag{2.40}
$$

where $\alpha_1 \;:=\; C_1\,\min\left\{\frac{1}{2\mu}\left(1-\frac{\kappa 1}{2\mu}\right),\kappa_2\right\}$. In this way, the bilinear form becomes V-elliptic for the ranges of the parameters given by $0<\kappa_1<2\mu$, $0<\kappa_2$, and $\kappa_3>0$ if $\widetilde{\kappa}\geq 1$, or $0<\kappa_3<\frac{\kappa_1\widetilde{\kappa}}{1-\widetilde{\kappa}}$ if $\widetilde{\kappa}<1$. Next, the inf-sup condition for B is established directly, because taking $(\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta}) = (0,\boldsymbol{\xi},0)$ gives

$$
\sup_{(\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta})\in H} \frac{B((\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta}),\boldsymbol{\xi})}{\|(\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{\eta})\|_H} \geq \frac{\|\boldsymbol{\xi}\|_{0,\Omega}^2}{\|\boldsymbol{\xi}\|_{1,\Omega}} \geq c^2\|\boldsymbol{\xi}\|_{1,\Omega},
$$

where the last step comes from the fact that $\mathbb{RM}$ is a finite dimensional space, thus the Open Mapping theorem gives the existence of a constant $c>0$ such that $\|\boldsymbol{\xi}\|_{0,\Omega}\geq c\|\boldsymbol{\xi}\|_{1,\Omega}$. To end the proof, we note that both bilinear forms have only L^2 inner products, so that their boundedness is directly established and thus Babuška-Brezzi conditions hold, which imply continuous dependence on data. $\square$

The existence of a solution for problem (2.34) is given by Schauder's Fixed Point Theorem.

Lemma 8. *Let $\boldsymbol{T}$ be the augmented fixed point operator given by* (2.34) *and assume that data assumptions* (2.8) *and* (2.9) *hold. Then*

$$
\boldsymbol{T}(\overline{B}(0,r_0)) \;\subset\; \overline{B}(0,r_0) \;:=\; \left\{\boldsymbol{v}\in\boldsymbol{H}^1(\Omega) : \|\boldsymbol{v}\|_{1,\Omega}\leq r_0\right\},
$$

where $r \leq r_0 := \alpha\, C\, M_F$ with C the constant arising from the stability estimate of the augmented problem. In addition, there holds

$$\|\boldsymbol{T}(z_1)-\boldsymbol{T}(z_2)\|_{1,\Omega} \leq \alpha\, \max(1,|\kappa_2|)\, C\, L_F\, \|z_1-z_2\|_{0,\Omega} \qquad \forall\, z_1, z_2 \in \boldsymbol{H}^1(\Omega)\,. \quad (2.41)$$

PROOF. Let $z \in \boldsymbol{H}^1(\Omega)$. Then, from the continuous dependence on data and (2.9) it follows:

$$\|\boldsymbol{T}(z)\|_{1,\Omega} = \|\boldsymbol{u}\|_{1,\Omega} \leq \alpha C \|F_z\|_{H'} \leq \alpha C M_F =: r_0.$$

Now, given $z_1, z_2 \in \boldsymbol{H}^1(\Omega)$, we subtract both augmented systems and get a new solution $u_1 - u_2$ with $F := F_{z_1} - F_{z_2}$. Applying continuity on data and (2.8) we get the desired result:

$$\begin{aligned}\|\boldsymbol{T}(z_1) - \boldsymbol{T}(z_2)\|_{1,\Omega} &\leq \alpha\, C\, \|F\|_{H'} \leq \alpha\, \max(1,|\kappa_2|)\, C\, \|F_{z_1} - F_{z_2}\|_{H'} \\ &\leq \alpha\, \max(1,|\kappa_2|)\, C\, L_F\, \|z_1 - z_2\|_{0,\Omega}\,.\end{aligned}$$

□

These results are enough to guarantee the existence of at least one solution of problem (2.34).

Theorem 12. *Assume that the data satisfy* (2.8) *and let $r_0 := \alpha C M_F$. Then the augmented problem* (2.34) *has at least one solution in $\overline{B}(0, r_0)$, all of which have the following a priori estimate:*

$$\|((\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\gamma}), \boldsymbol{\xi})\|_{H\times Q} \leq \alpha\, C\, M_F\,.$$

Moreover, if $\alpha\, C\, L_F\, \max(1,|\kappa_2|) < 1$, the solution is unique.

PROOF. Let $\{z_k\}_{k\in\mathbb{N}} \subset \overline{B}(0, r_0)$ be a bounded sequence given by iterated solutions of problem (2.34), then it has a subsequence $\{z_k^{(1)}\}_{k\in\mathbb{N}} \subseteq \{z_k\}_{k\in\mathbb{N}}$ weakly convergent to some z in $\boldsymbol{H}^1(\Omega)$. Rellich-Kondrachov's compactness theorem states that there exists a compact embedding $\boldsymbol{i}_c : \boldsymbol{H}^1(\Omega) \to \boldsymbol{L}^2(\Omega)$, so that $z_{k_i} \to z$ in $\boldsymbol{L}^2(\Omega)$. Using estimate

(2.41) from Lemma 8 we see that also $\boldsymbol{T}(z_{k_i}) \to \boldsymbol{T}(z)$ in $\boldsymbol{H}^1(\Omega)$ and so there exists a fixed point due to Schauder's fixed point Theorem. The bound on α is established if $\boldsymbol{T}$ is forced to be a contraction. $\square$

2.5.2. Analysis of the discrete problem

For a Galerkin scheme of the augmented formulation, let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangularizations of $\overline{\Omega}$ of size h. Now, it is crucial for the analysis to notice that the bilinear form B does not change when a discretization is made because the rigid motions space is already discrete. With this in mind, the inf-sup condition is proved trivially and so the ellipticity of A in the discrete kernel is just a consequence of the continuous case. In virtue of this, any conforming set of discrete spaces will suffice, and as such we take the following:

$$
\begin{aligned}
H_h^{\boldsymbol{\sigma}} &:= \big\{\boldsymbol{\tau}_h \in \boldsymbol{H}(\mathrm{div};\Omega) : \boldsymbol{\tau}_{h,i}|_T \in \mathbb{RT}_0 \quad \forall T \in \mathcal{T}_h\big\},\\
H_h^{\boldsymbol{u}} &:= \big\{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) : \boldsymbol{v}|_T \in [\mathbb{P}_1(T)]^n \quad \forall T \in \mathcal{T}_h\big\},\\
H_h^{\boldsymbol{\gamma}} &: = \left\{ \begin{bmatrix} 0 & \psi_h \\ -\psi_h & 0 \end{bmatrix} : \psi_h \in \mathbb{P}_0(T) \quad \forall T \in \mathcal{T}_h \right\}.
\end{aligned}
$$

Then, setting $H_h = (H_h^{\boldsymbol{\sigma}} \cap \boldsymbol{H}_0(\mathrm{div};\Omega)) \times H_h^{\boldsymbol{v}} \times H_h^{\boldsymbol{\gamma}}, Q_h = \mathbb{RM}$, the Galerkin scheme of the augmented partial problem reads: Given $\boldsymbol{z}_h \in H_h^{\boldsymbol{u}}$, find $((\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\xi}_h)) \in H_h \times Q_h$ such that

$$
\begin{aligned}
A((\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\gamma}_h), (\boldsymbol{\tau}_h, \boldsymbol{v}_h, \boldsymbol{\eta}_h)) + B((\boldsymbol{\tau}_h, \boldsymbol{v}_h, \boldsymbol{\eta}_h), \boldsymbol{\xi}_h) &= \alpha F_{\boldsymbol{z}_h}((\boldsymbol{\tau}_h, \boldsymbol{v}_h, \boldsymbol{\eta}_h)) \quad \forall (\boldsymbol{\tau}_h, \boldsymbol{v}_h, \boldsymbol{\eta}_h) \in H_h,\\
B((\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\rho}_h) &= 0 \qquad\qquad\qquad\qquad \forall \boldsymbol{\rho} \in Q_h.
\end{aligned}
\tag{2.42}
$$

We finally define the discrete operator $\boldsymbol{T}_h : H_h^{\boldsymbol{u}} \to H_h^{\boldsymbol{u}}$ given by

$$\boldsymbol{T}_h(\boldsymbol{z}_h) = \boldsymbol{u}_h,$$

where $\boldsymbol{u}_h$ is the second component of the solution in H_h from the discrete partial augmented problem (2.42), so that the nonlinear discrete problem can be stated as: Find

$\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ such that

$$\boldsymbol{T}_h(\boldsymbol{u}) \;=\; \boldsymbol{u}\,. \tag{2.43}$$

As we did with the mixed formulation, we will first establish the well posedness of the discrete partial problem, and then extend this to the nonlinear problem.

Theorem 13. *Given $\boldsymbol{z} \in H_h^{\boldsymbol{u}}$, problem* (2.42) *has a unique solution $((\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) \in H_h \times Q_h$ and there exists a constant C such that*

$$\|((\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\xi}_h)\|_{H\times Q} \leq \alpha C \|F_{\boldsymbol{z}_h}\|_{H_h'}.$$

PROOF. First we note that $N(\mathbb{B}_h) \subseteq N(\mathbb{B})$. From there, the inf-sup condition and ellipticity are a consequence of the continuous case and thus Babuška-Brezzi conditions hold. □

Now we establish the regularity of the discrete operator. The following lemma is analogous to the continuous case, so we leave it without proof.

Lemma 9. *Let $\boldsymbol{T}_h$ be the discrete fixed point operator associated to the discrete augmented problem* (2.43) *and assume that data conditions* (2.8) *and* (2.9) *hold. Then*

$$\boldsymbol{T}_h(\overline{B}_h(0, r_0)) \subset \overline{B}_h(0, r_0) := \left\{\boldsymbol{v} \in H_h^{\boldsymbol{u}} : \|\boldsymbol{u}\|_{1,\Omega} \leq r_0\right\},$$

where $r_0 := \alpha C M_F$ with C the constant that arises from the a priori estimate of the partial problem. In addition, there holds

$$\|\boldsymbol{T}_h(z_1) - \boldsymbol{T}_h(z_2)\|_{1,\Omega} \leq \alpha \max(1, |\kappa_2|) C L_F \|z_1 - z_2\|_{1,\Omega} \qquad \forall z_1, z_2 \in H_h^{\boldsymbol{u}}.$$

This is enough to prove the well-posedness of the original nonlinear problem, which we show in the following theorem.

Theorem 14. *Assuming data conditions* (2.8) *and* (2.9)*, problem* (2.43) *has at least one solution. All solutions verify the a priori estimate*

$$\|((\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\xi}_h)\|_{H\times Q} \leq \alpha C M_F. \tag{2.44}$$

Moreover, if $\alpha\, C\, L_F\, \max(1, |\kappa_2|) < 1$*, the solution is unique. If under the same hypothesis* $((\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) \in H \times Q$ *is the unique solution to the continuous problem, Cea's estimate holds for a certain constant* $C_4 > 0$*:*

$$\|((\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) - ((\boldsymbol{\sigma}, \boldsymbol{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\xi}_h)\|_{H\times Q} \leq C_4 \inf_{(\boldsymbol{\tau}_h, \boldsymbol{v}_h, \boldsymbol{\eta}_h)\in H_h} \|(\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\gamma}) - (\boldsymbol{\tau}_h, \boldsymbol{v}_h, \boldsymbol{\eta}_h)\|_H. \tag{2.45}$$

PROOF. First for the a priori estimate, we note from Lemma 9 that Brower's fixed point theorem holds, so that there exists a fixed point with

$$\boldsymbol{T}_h(\boldsymbol{u}_h) = \boldsymbol{u}_h,$$

and then the rest of the proof follows exactly as in the mixed case. We obtain the a priori estimate from the partial a priori estimate

$$\|((\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\gamma}), \boldsymbol{\xi})\|_{H\times Q} \;\leq\; \alpha\, C\, \|F_{\boldsymbol{u}_h}\|_{H_h'} \;\leq\; \alpha\, C\, M_F\,.$$

In the case of the mixed formulation, proving Cea's estimate was done by taking partial problems and then subtracting. We only need for the same technique to work to have an approximation property for a given $\boldsymbol{z}$. Indeed, plugging $\boldsymbol{z}$ in (2.42), then the solution $((\boldsymbol{\zeta}_h, \boldsymbol{w}_h, \boldsymbol{s}_h), \boldsymbol{\varphi}_h) \in H_h \times Q_h$ to such problem would have the desired property for a constant $C_3 > 0$ independent of $\boldsymbol{z}$:

$$\begin{aligned} &\|((\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) - ((\boldsymbol{\zeta}_h, \boldsymbol{w}_h, \boldsymbol{s}_h), \boldsymbol{\varphi}_h)\|_{H\times Q} \\ &\qquad \leq\; C_3 \inf_{(\boldsymbol{\tau}_h, \boldsymbol{v}_h, \boldsymbol{\eta}_h)\in H_h} \|(\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\gamma}) - (\boldsymbol{\tau}_h, \boldsymbol{v}_h, \boldsymbol{\eta}_h)\|_H\,. \end{aligned} \tag{2.46}$$

Using equation (2.46) and proceeding as in the proof of the mixed case, the result holds for the constant $C_4 := \frac{C_3}{1-\alpha C \max(1,|\kappa_2|) L_F}$. □

Corollary 3. *Under the assumptions from the previous theorem, a sufficiently small α grants C_4 independent of it.*

PROOF. The proof is analogous to that of the primal and mixed cases. $\square$

We end this section with the rates of convergence for the solutions of (2.43). We first recall classic approximation results from (Brezzi & Fortin, 1991) and proceed as in (Gatica, 2006):

$(AP_h^{\boldsymbol{\sigma}})$ For each $\boldsymbol{\tau} \in \mathbf{H}^1(\Omega) \cap \mathbf{H}_0$ with $\mathrm{div}\boldsymbol{\tau} \in \boldsymbol{H}^1(\Omega)$ there exists $\boldsymbol{\tau}_h \in H_{0,h}^{\boldsymbol{\sigma}}$ such that

$$\|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{\mathrm{div};\Omega} \leq Ch\left\{\|\boldsymbol{\tau}\|_{1,\Omega} + \|\mathrm{div}\boldsymbol{\tau}\|_{1,\Omega}\right\}.$$

$(AP_h^{\boldsymbol{u}})$ For each $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ there there exists $\boldsymbol{v}_h \in H_h^{\boldsymbol{u}}$ such that

$$\|\boldsymbol{v} - \boldsymbol{v}_h\|_{0,\Omega} \leq Ch\|\boldsymbol{v}\|_{1,\Omega}.$$

$(AP_h^{\boldsymbol{\gamma}})$ For each $\boldsymbol{\eta} \in \mathbf{H}^1(\Omega) \cap \mathbf{L}^2_{asym}(\Omega)$ there exists $\boldsymbol{\eta}_h \in H_h^{\boldsymbol{\gamma}}$ such that

$$\|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{0,\Omega} \leq Ch\|\boldsymbol{\eta}\|_{1,\Omega}.$$

Consequently, we can prove the following result:

Theorem 15. *Under data assumptions* (2.8) *and* (2.9) *and assuming* $\alpha\, C\, L_F\, \max(1, |\kappa_2|) \leq \frac{1}{2}$, *let* $((\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\gamma}), \boldsymbol{\rho}) \in H \times Q$ *and* $((\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\rho}_h) \in H_h \times Q_h$ *be the unique solutions of the continuous problem* (2.34) *and the discrete one* (2.43) *respectively. Then, there exists a constant* $C > 0$ *independendt of* h *such that whenever* $\boldsymbol{\sigma} \in \mathbf{H}^1(\Omega)$, $\mathrm{div}\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega)$, $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega)$ *and* $\boldsymbol{\gamma} \in \mathbf{H}^1(\Omega)$, *there holds*

$$\|((\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) - ((\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\xi}_h)\|_{H\times Q} \leq C\, h\Big\{\|\boldsymbol{\tau}\|_{1,\Omega} + \|\mathrm{div}\boldsymbol{\tau}\|_{1,\Omega} + \|\boldsymbol{v}\|_{1,\Omega} + \|\boldsymbol{\eta}\|_{1,\Omega}\Big\}. \tag{2.47}$$

PROOF. The proof relies in the application of Cea's estimate together with the approximation properties $(AP_h^{\boldsymbol{\sigma}})$, $(AP_h^{\boldsymbol{u}})$ and $(AP_h^{\boldsymbol{\gamma}})$. $\square$

2.6. Numerical simulations

The registration problem was implemented in `python` using the library FEniCS ((Alnæs et al., 2015), for reference see (Logg, Mardal, Wells, et al., 2012)). Numerical results were visualized using Paraview (Ahrens, Geveci, & Law, 2005). The numerical solution of the discrete problems was computed using a regularization technique of the gradient-flow type, which is extensively used in the DIR community (see (Schmidt-Richberg, 2014), also known as the proximal point method in the optimization community (Rockafellar, 1976)). To this end, we consider an artificial time variable such that the velocity of the movement points towards the greatest-descent direction of the original functional and then discretize in time using an implicit Euler method. The resulting iterative scheme (time-discretized gradient-flow problem) reads

$$\frac{1}{\varDelta t}\langle \boldsymbol{u} - \boldsymbol{u}_n, \boldsymbol{v}\rangle = -a(\boldsymbol{u}, \boldsymbol{v}) + f_{\boldsymbol{u}}(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in \mathcal{V}. \tag{2.48}$$

It is easy to show that the previous problem derives from the following variational principle

$$\min_{\boldsymbol{u}\in\mathcal{V}} \frac{1}{2} a(\boldsymbol{u}, \boldsymbol{u}) - \mathcal{D}[\boldsymbol{u}] + \frac{1}{2\varDelta t}\|\boldsymbol{u} - \boldsymbol{u}_n\|_{0,\varOmega}^2 := \varPi_n[\boldsymbol{u}], \tag{2.49}$$

where $\varDelta t$ is a regularizing parameter chosen so that the incremental functional is strictly convex (Hurtado & Henao, 2014), guaranteeing the convergence of the iterative problem. The stop criterion is based on the L^2 error of the solution to (2.49) with respect to the previous converged solution.

In the following, the DIR methods developed in this work are tested in the registration of a synthetic image. To this end, we let $\varOmega = [0, 1]^2$, and consider the reference image described by

$$R(x_0, x_1) = \sin(2\pi x_0)\sin(2\pi x_1). \tag{2.50}$$

To construct the target image, we first consider an affine displacement field of the form

$$\boldsymbol{u}(x) = \begin{bmatrix} 0.4 & 0 \\ 0 & -0.2 \end{bmatrix} \boldsymbol{x}, \tag{2.51}$$

where $\boldsymbol{x} = (x_0, x_1)^{\mathrm{t}}$. Then, the target image is found by composition, i.e., $T = R \circ (\boldsymbol{id} + \boldsymbol{u})^{-1}$. Figure 2.1 shows the resulting reference and target images.

We perform numerical convergence tests for the three formulations using the images just defined. To this end, a fine mesh with 65,536 triangular elements and characteristic length $h = 0.0055$ is considered as the baseline solution. The degrees of freedom involved in each case are denoted by N and the corresponding errors are quantified as

$$e_0(\boldsymbol{\sigma}_h) := \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega}, \qquad e(\boldsymbol{\sigma}_h) := \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathrm{div};\Omega},$$
$$e_0(\boldsymbol{u}_h) := \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega}, \qquad e(\boldsymbol{u}_h) := \|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega}\,,$$

and

$$e(\boldsymbol{\gamma}_h) := \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,\Omega},$$

where $(\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\gamma})$ is the baseline solution to the mixed or the augmented case when it corresponds, and $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\gamma}_h)$ is the associated finite element solution with coarser mesh sizes. We further define the experimental rates of numerical convergence as

$$r_0(\boldsymbol{\sigma}_h) := \frac{\log(e_0(\boldsymbol{\sigma})/e_0'(\boldsymbol{\sigma}_h))}{\log(h/h')}, \qquad r(\boldsymbol{\sigma}_h) := \frac{\log(e(\boldsymbol{\sigma})/e'(\boldsymbol{\sigma}_h))}{\log(h/h')},$$
$$r_0(\boldsymbol{u}_h) := \frac{\log(e_0(\boldsymbol{u}_h)/e_0'(\boldsymbol{u}_h)}{\log(h/h')}, \qquad r(\boldsymbol{u}_h) := \frac{\log(e(\boldsymbol{u}_h)/e'(\boldsymbol{u}_h)}{\log(h/h')},$$

and

$$r(\boldsymbol{\gamma}_h) := \frac{\log(e(\boldsymbol{\gamma}_h)/e'(\boldsymbol{\gamma}_h))}{\log(h/h')}\,,$$

where h and h' denote two consecutive mesh sizes with corresponding errors e and e'. The experiments were performed using the primal, mixed and augmented schemes with $\alpha = 0.5$, $\Delta t = 0.1$ and $\mu = \lambda = 0.5$, and the termination criterion used was a relative error of 10^{-8}. In turn, the stabilization parameters κ_i, $i \in \{1, 2, 3\}$, of the augmented scheme were chosen according to the ranges derived in the proof of Theorem 11. Figures 2.2, 2.3 and 2.4 show the results of the numerical convergence studies measured under several error norms for the primal, mixed and augmented schemes, respectively. In all cases, the error norms monotonically decrease as the number of degrees of freedom increases (equivalently, the mesh size is reduced). The convergence rates obtained for the primal, mixed and augmented schemes are documented in Tables 2.1, 2.2, and 2.3, respectively. We note that, in most of the cases, the expected (linear) convergence rate is met by the numerical tests, and sometimes exceeded. The exception is the $H(\mathrm{div})$ norm of $\boldsymbol{\sigma}$ in the mixed and augmented schemes where convergence rates are positive but well below 1, which simply says that the computation of the rates of convergence for this unknown depends strongly on how accurately the continuous solution is approximated.

Figure 2.5 shows the reference image, and the composed image $T(\boldsymbol{x} + \boldsymbol{u}_h(\boldsymbol{x}))$ for all three formulations using a structured mesh with 2048 elements, for the sake of comparison of the registration solution between methods. All three DIR methods delivered very similar solutions for the composed target image, which are qualitatively similar to the reference image. The associated stress fields are depicted in Figure (2.6) in terms of the Frobenius norm of the stress tensor. Stress fields are displayed using an L^2-projection to a $\mathbb{P}_1$ FE space constructed on the mesh, for visualization purposes. While all three solutions qualitatively agree, the solution of the primal scheme differs from the solutions provided by the mixed and augmented schemes, which are nearly identical.

2.7. Discussion

In this work we have proposed and analyzed a mixed and augmented formulations for the DIR problem, along with suitable finite-element discretization for their numerical solution. To this end, we consider the variational formulation of the Euler-Lagrange equations associated to the original DIR problem, which presents a structure similar to that of a linear elasticity problem with a nonlinear load source. In this form, we leverage a large body of results in mathematical and numerical analysis to study the DIR problem, namely the Babŭska-Brezzi theory for mixed formulations as well as fixed-point theorems. Two key assumptions needed to prove existence of solution, both in the continuous and discrete settings, are the Lipschitz continuity and boundedness of the distributional gradient of the similarity measure, which we argue can be easily verified in DIR applications where images are typically constructed using smooth interpolation schemes. This result, which we particularize for similarity measures of the sum-of-squares-difference's type, is readily extendable to other standard similarity measures, such as correlation and mutual information. For the case of the elastic regularization term, we assume the bijectivity of its Gâteaux gradient, possibly quotiented by its kernel, or simply its surjectivity and the availability of a space reduction that renders the operator injective. Under these assumptions, fixed point arguments allowed us to establish well-posedness of the continuous and discrete problems. A key result of our work, that can be exploited in the feasibility analysis of DIR problems, is the necessary condition for uniqueness of the DIR solution: $\alpha C L_F < 1$. This inequality establishes a novel relation between the similarity weighting constant α and the Lipschitz constant of the nonlinear source term L_F given by $\alpha \propto L_F^{-1}$. The Lipschitz constant, is interpreted as the largest slope between two points as seen in the definition:

$$\frac{\|F(\boldsymbol{u}) - F(\boldsymbol{v})\|}{\|\boldsymbol{u} - \boldsymbol{v}\|} \leq L_F \qquad \forall \boldsymbol{u}, \boldsymbol{v}.$$

In virtue of the preceding discussion, we show that DIR of images with higher gradients (i.e., high image contrast, or rapidly oscillating intensity fields due to noise) require the reducing the relevance of the similarity term for well-posedness to hold. This result provides a new insight for the success of DIR, for which a standard practice is to preprocess the reference and target images to reduce image contrast. One such approach is pyramidal gaussian convolution, where images are filtered and then subsampled to perform a sequence of chained registrations, with increasing level of detail, but more sophisticated strategies have been proposed to reduce noise (Paquin, Levy, & Xing, 2008; Athavale & Tadmor, 2011).

Another key contribution resulting from this work is the formulation and analysis of mixed and augmented numerical schemes for DIR, where convergence can be proven not only for the transformation (displacement) mapping but also for the stress (and in turn

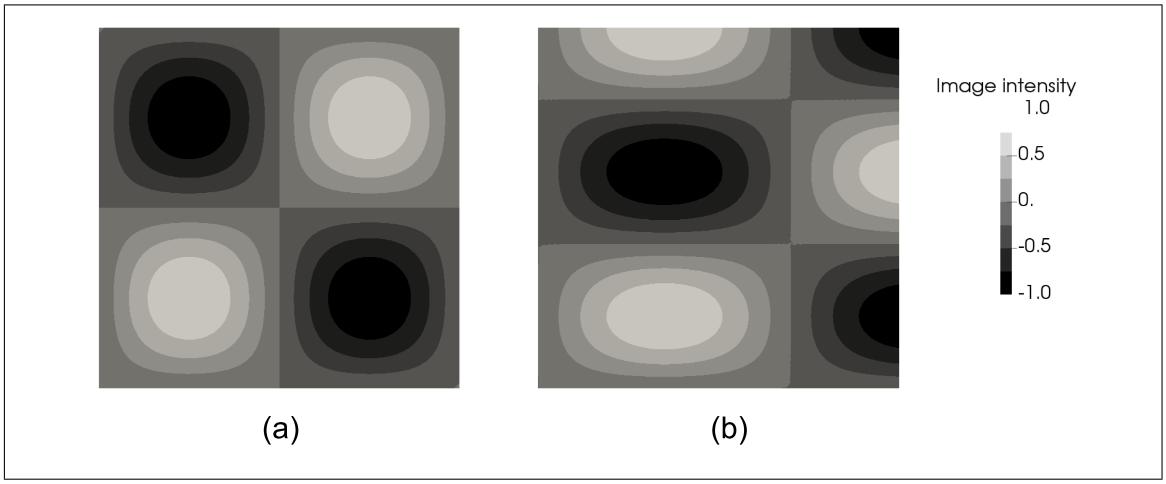

Figure 2.1. Synthetic images used in numerical tests: (a) Reference image, (b) Target image.

strain) and the rotation tensor. Our results show that, while the transformation mapping that results from solving the primal DIR problem can be very similar to those obtained from the mixed and augmented formulation, the stress fields can be different. Traditionally, the transformation mapping has been the fundamental field sought in DIR applications, particularly in medical applications where the goal is to align the reference and target images (Sotiras et al., 2013). However, recent applications of medical image quantification have highlighted the importance of guaranteeing certain accuracy and convergence when estimating stress tensor fields (Sotelo et al., 2016) and rotation tensor fields (Sotelo et al., 2017), due to their important connection to medical conditions. Further developments to improve the accuracy of the numerical solution, that are naturally developed within the finite-element framework adopted in this work, are the introduction of *a-posteriori* mesh refinement methods, where recent results in linear elasticity for mixed formulations with Neumann boundary conditions (Domínguez, Gatica, & Márquez, 2016) can be extended to the case of DIR problems addressed here.

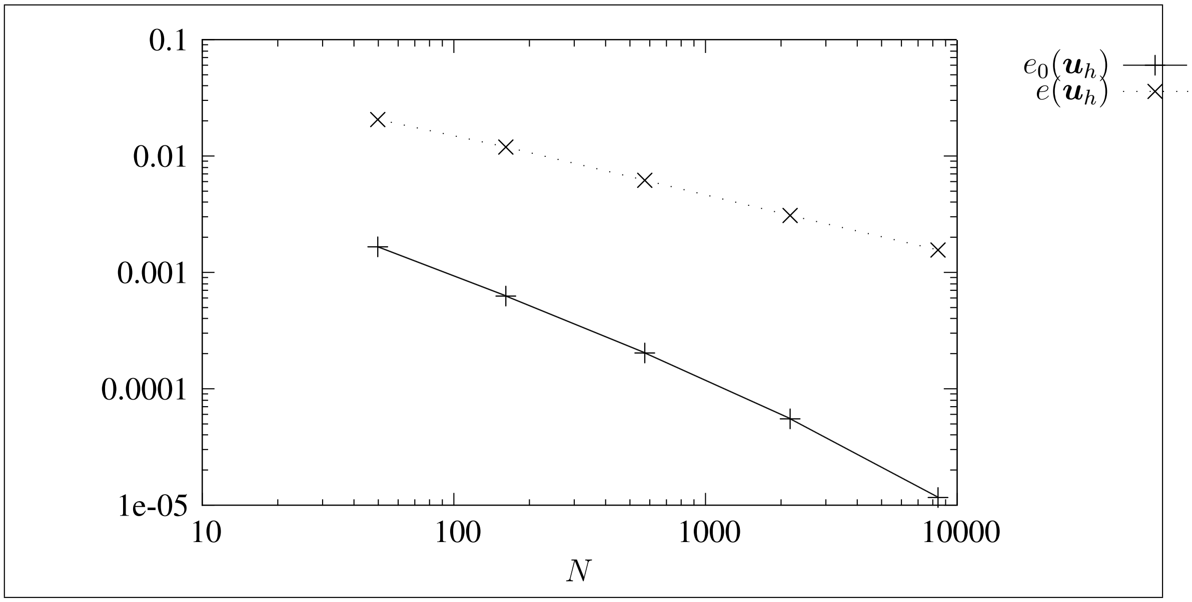

Figure 2.2. Errors vs. N for the primal formulation

N	$e_0(\boldsymbol{u}_h)$	$r_0(\boldsymbol{u}_h)$	$e(\boldsymbol{u}_h)$	$r(\boldsymbol{u}_h)$
50	1.66E-3	—	2.06E-2	—
162	6.23E-4	1.42	1.19E-2	0.80
578	2.02E-4	1.62	6.18E-3	0.94
2178	5.55E-5	1.86	3.08E-3	1.01
8450	1.16E-5	2.26	1.55E-3	0.99

Table 2.1. Errors and convergence rates for the primal scheme.

N	$e_0(\boldsymbol{\sigma}_h)$	$r_0(\boldsymbol{\sigma}_h)$	$e(\boldsymbol{\sigma}_h)$	$r(\boldsymbol{\sigma}_h)$	$e_0(\boldsymbol{u}_h)$	$r_0(\boldsymbol{u}_h)$	$e(\boldsymbol{\gamma}_h)$	$r(\boldsymbol{\gamma}_h)$
323	0.80	—	17.49	—	7.25E-2	—	0.53	—
1219	0.46	0.81	14.88	0.23	3.93E-2	0.88	0.39	0.45
4739	0.32	0.50	12.94	0.20	2.25E-2	0.81	0.27	0.52
18691	0.23	0.50	11.34	0.19	1.26E-2	0.83	0.19	0.51
74243	0.13	0.85	8.69	0.38	5.72E-3	1.14	0.11	0.75

Table 2.2. Errors and convergence rates for the mixed scheme.

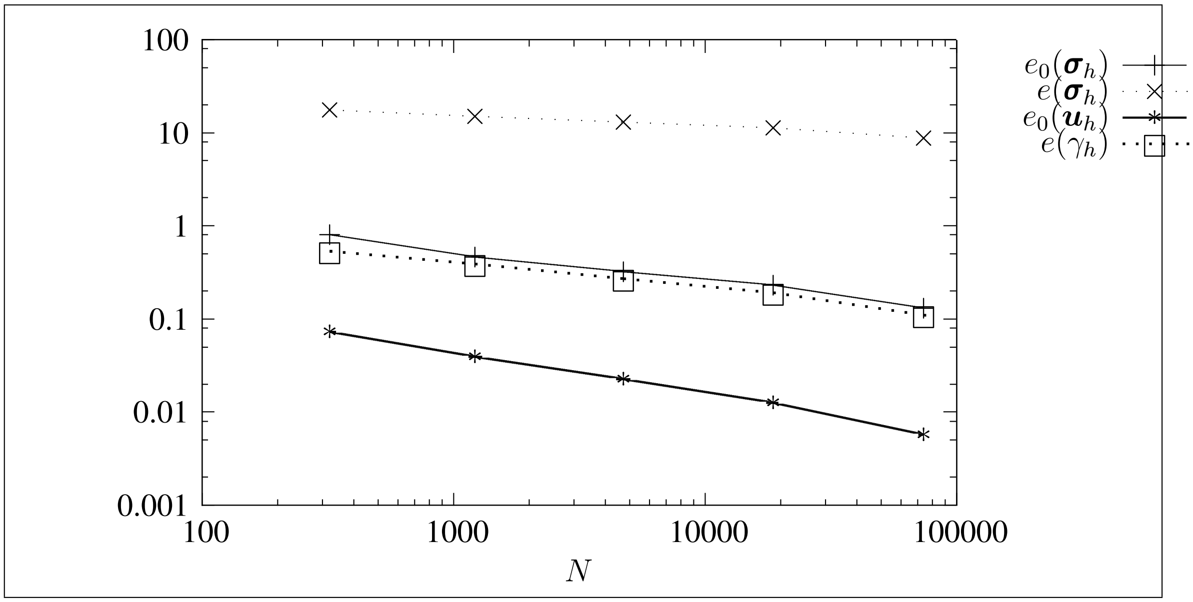

Figure 2.3. Errors vs. N for the mixed formulation

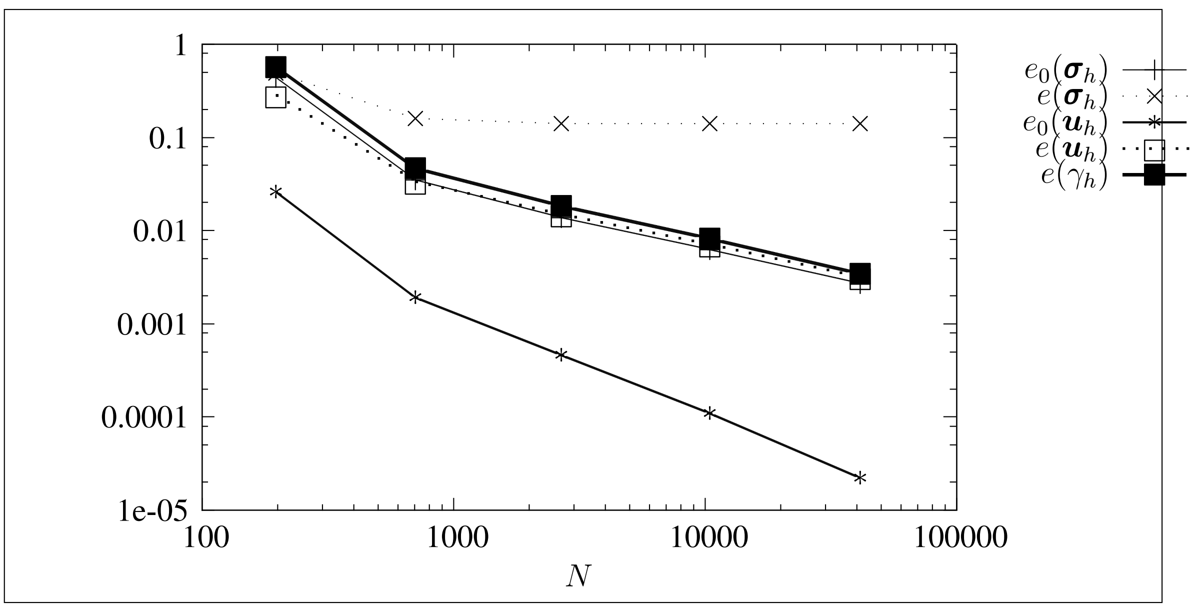

Figure 2.4. Errors vs. N for the augmented formulation

N	$e_0(\boldsymbol{\sigma}_h)$	$r_0(\boldsymbol{\sigma}_h)$	$e(\boldsymbol{\sigma}_h)$	$r(\boldsymbol{\sigma}_h)$	$e_0(\boldsymbol{u}_h)$	$r_0(\boldsymbol{u}_h)$	$e(\boldsymbol{u}_h)$	$r(\boldsymbol{u}_h)$	$e(\boldsymbol{\gamma}_h)$	$r(\boldsymbol{\gamma}_h)$
197	4.40E-1	—	0.49	—	2.58E-2	—	2.80E-1	—	5.70E-1	—
709	3.51E-2	3.65	0.16	1.570	1.91E-3	3.76	3.33E-2	3.08	4.66E-2	3.62
2693	1.38E-2	1.35	0.14	0.190	4.57E-4	2.06	1.48E-2	1.17	1.82E-2	1.36
10501	6.26E-3	1.14	0.14	0.061	1.09E-4	2.06	7.06E-3	1.07	8.14E-3	1.16
41477	2.71E-3	1.21	0.14	0.004	2.22E-5	2.30	3.14E-3	1.17	3.43E-3	1.25

Table 2.3. Errors and convergence rates for the augmented scheme.

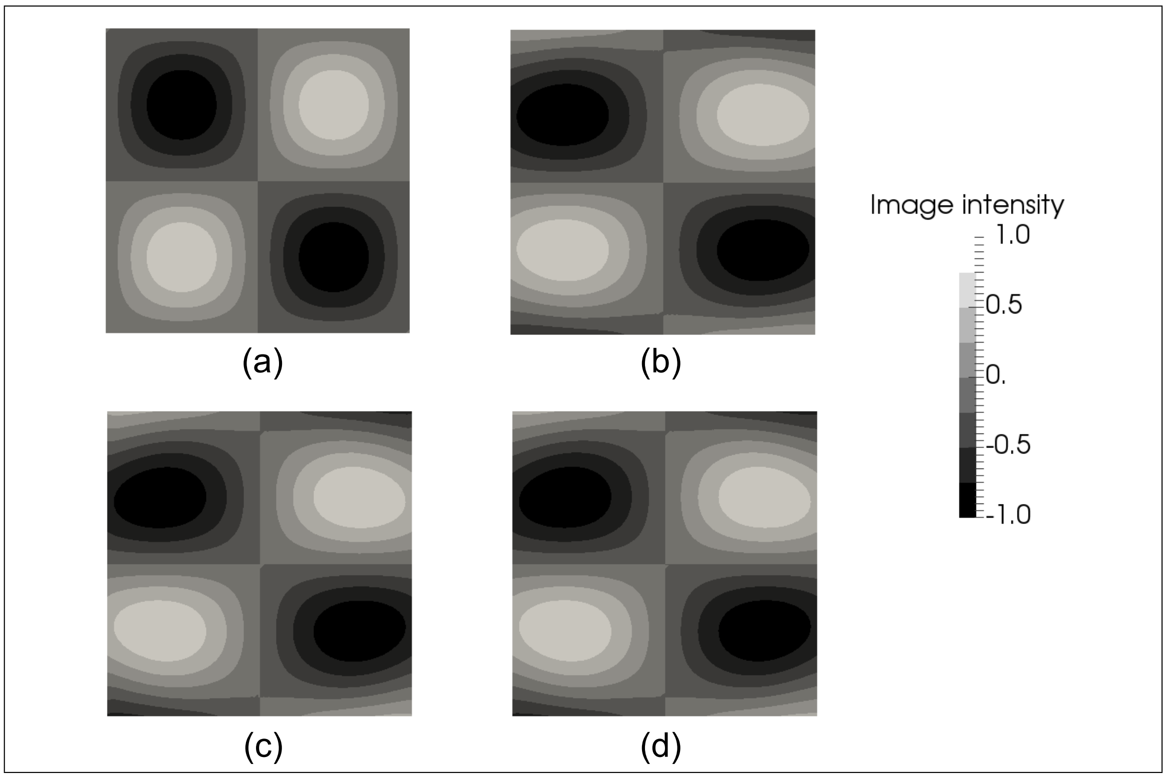

Figure 2.5. Registration results and comparison for $\alpha = \mu = \lambda = 0.5$. (a) Reference image R. Composed image $T(\boldsymbol{x} + \boldsymbol{u}_h(\boldsymbol{x}))$ using solutions of (b) the primal method, (c) the mixed method, and (d) the augmented method.

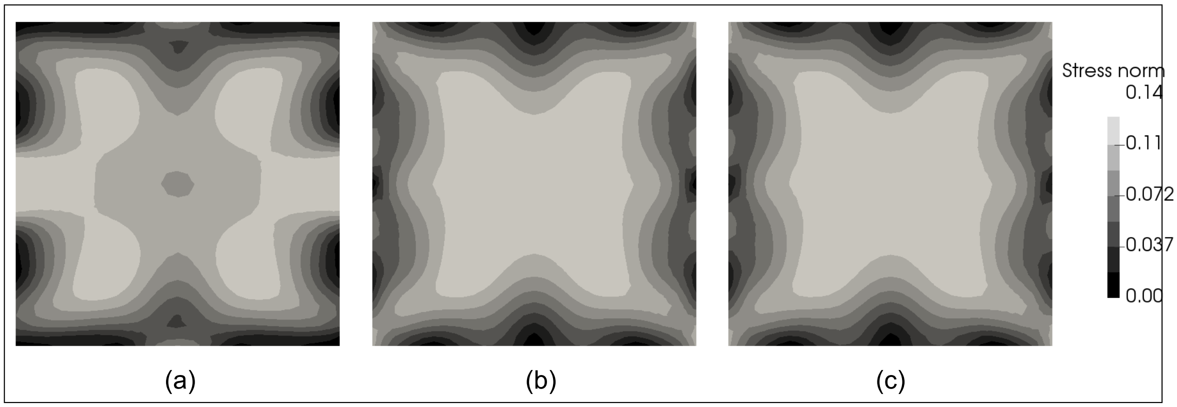

Figure 2.6. Frobenius norm of the stress tensor field for $\alpha = \mu = \lambda = 0.5$ and $\Delta t = 0.1$ with 2^5 elements per side: (a) primal scheme, (b) mixed scheme, and (c) augmented scheme. Stress fields are displayed using an L^2-projection to a $\mathbb{P}_1$ FE space constructed on the mesh.

3. CONCLUSIONS

In this work we presented a mixed formulation for elastic DIR. One of the advantages is that the problem can be seen as a linear elasticity equation with a nonlinear load term problem, for which there exists a lot of analysis. We require only Lipschitz continuity and boundedness on the load, which is easily verifiable for the most common interpolation schemes for the images and similarity terms. We also obtained a sufficient condition for uniqueness, namely $\alpha C L_F < 1$, which establishes a novel relationship between the similarity weight α and the Lipschitz constant L_F. It had been noticed that high gradients affect registration, but no formal proof of this had been given. It also gives a new way of understanding the success of image smoothing techniques, such as gaussian filtering.

Another important consequence of this approach is that it states convergence of all the unknowns used, namely the displacement field, the rigid motion Lagrange multiplier, the skew-symmetric component of the displacement's gradient and the stress. It also delivers the rates of convergence at which this fields converge, which helps estimate the required computational cost of the method. This approach gives a more straightforward way of analyzing the problem and the relationship between it and the discrete problem.

Finally, we show that the incorporation of an extra rigid motion space removes stiffness from the scheme. This happens because the left hand side of the system has a non trivial kernel, and thus restricts the solution to be orthogonal to such space. This is inexpensive in practice, as it adds only 3 degrees of freedom to the problem.

4. FUTURE WORK

From here, in the future we will prove the well-posedness of the extended problem with the added degrees of freedom in $\mathbb{RM}$ and analyze how they affect the registration of real lung images. We expect to show that the stiffness incurred severely deteriorates schemes with Neumann boundary conditions. This kind of schemes cannot be circumvented as Dirichlet conditions impose spurious stresses near the boundary.

As the images are highly nonlinear, semi-implicit time-stabilized techniques are common. In this regard, we will also prove the convergence and stability of such stabilization techniques in order to find theoretical bounds for the time step, which is still largely an open question. This will be performed to both implicit and semi-implicit schemes. In principle the implicit schemes are avoided because they are more expensive, but we hope to see comparable computation times in virtue of the the added stability of implicit schemes.

We will also study and compare mixed methods with the other existing methods to benchmark their usefulness in stress estimation with respect to state-of-the-art software. Currently, the most popular deformation model is B-splines, which have been shown to produce numerical artifacts when estimating gradients (Hurtado et al., 2017). This was one of the main motivations for the development of a mixed scheme in lung image registration, and we expect it to perform better than current approaches.

Finally, we will develop an a posteriori method to establish a way of performing mesh refinement in registration, which is largely an open question. This will help the development of future registration codes as it will give a guideline for refining. One of the biggest challenges in registration is its computational cost, because the images are highly nonlinear, so methods for cost reduction are critical. This problem has been traditionally approached with smoothing techniques, and we expect to give a new way of reducing the cost of registration.

REFERENCES

Ahrens, J., Geveci, B., & Law, C. (2005). Paraview: An end-user tool for large data visualization. *The Visualization Handbook*, *717*.

Alnæs, M. S., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., … Wells, G. N. (2015). The fenics project version 1.5. *Archive of Numerical Software*, *3*(100). doi: 10.11588/ans.2015.100.20553

Amelon, R. E., Cao, K., Ding, K., Christensen, G. E., Reinhardt, J. M., & Raghavan, M. L. (2011). Three-dimensional characterization of regional lung deformation. *Journal of Biomechanics*, *44*(13), 2489–2495. Retrieved from `http://dx.doi.org/10.1016/j.jbiomech.2011.06.009` doi: 10.1016/j.jbiomech.2011.06.009

Athavale, P., & Tadmor, E. (2011). Integro-differential equations based on (BV, L^1) image decomposition. *SIAM Journal on Imaging Sciences*, *4*(1), 300–312. Retrieved from `http://dx.doi.org/10.1137/100795504` doi: 10.1137/100795504

Aubert, G., & Kornprobst, P. (2006). *Mathematical Problems in Image Processing* (Second ed., Vol. 147). Springer, New York. (Partial differential equations and the calculus of variations, With a foreword by Olivier Faugeras)

Barrientos, M. A., Gatica, G. N., & Stephan, E. P. (2002). A mixed finite element method for nonlinear elasticity: two-fold saddle point approach and a-posteriori error estimate. *Numerische Mathematik*, *91*(2), 197–222. Retrieved from `http://dx.doi.org/10.1007/s002110100337` doi: 10.1007/s002110100337

Borzi, A., Ito, K., & Kunisch, K. (2002). Optimal control formulation for determining optical flow. *SIAM Journal on Scientific Computing*, *24*(3), 818–847. Retrieved from `http://dx.doi.org/10.1137/S1064827501386481` doi: 10.1137/S1064827501386481

Brenner, S. C., & Scott, L. R. (2008). *The Mathematical Theory of Finite Element Methods* (Third ed., Vol. 15). Springer, New York. Retrieved from `http://dx.doi.org/10.1007/978-0-387-75934-0` doi: 10.1007/978-0-387-75934-0

Brezis, H. (2011). *Functional analysis, Sobolev spaces and partial differential equations*. Springer, New York.

Brezzi, F., Douglas, J., Jr., & Marini, L. D. (1984). Two families of mixed finite elements for second order elliptic problems. *Istituto di Analisi Numerica del Consiglio Nazionale delle Ricerche [Institute of Numerical Analysis of the National Research Council]*, *435*, 32.

Brezzi, F., & Fortin, M. (1991). *Mixed and Hybrid Finite Element Methods* (Vol. 15). Springer-Verlag, New York. Retrieved from `http://dx.doi.org/10.1007/978-1-4612-3172-1` doi: 10.1007/978-1-4612-3172-1

Burger, M., Modersitzki, J., & Ruthotto, L. (2013). A hyperelastic regularization energy for image registration. *SIAM Journal on Scientific Computing*, *35*(1), B132-B148. Retrieved from `https://doi.org/10.1137/110835955` doi: 10.1137/110835955

Chen, K., & Lorenz, D. A. (2011). Image sequence interpolation using optimal control. *Journal of Mathematical Imaging and Vision*, *41*(3), 222–238. Retrieved from `http://dx.doi.org/10.1007/s10851-011-0274-2` doi: 10.1007/s10851-011-0274-2

Choi, S., Hoffman, E. A., Wenzel, S. E., Tawhai, M. H., Yin, Y., Castro, M., & Lin, C. L. (2013). Registration-based assessment of regional lung function via volumetric CT images of normal subjects vs. severe asthmatics. *Journal of Applied Physiology*, *115*(5), 730–742. doi: 10.1152/japplphysiol.00113.2013

Christensen, G. E., Song, J. H., Lu, W., El Naqa, I., & Low, D. a. (2007). Tracking lung tissue motion and expansion/compression with inverse consistent image registration and spirometry. *Medical Physics*, *34*, 2155–2163. doi: 10.1118/1.2731029

Ciarlet, P. G. (2013). *Linear and Nonlinear Functional Analysis with Applications*. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Domínguez, C., Gatica, G. N., & Márquez, A. (2016). A residual-based a posteriori error estimator for the plane linear elasticity problem with pure traction boundary conditions. *Journal of Computational and Applied Mathematics*, *292*.

Droske, M., & Ring, W. (2006). A Mumford-Shah level-set approach for geometric image registration. *SIAM Journal on Applied Mathematics*, *66*(6), 2127–2148. Retrieved from http://dx.doi.org/10.1137/050630209 doi: 10.1137/050630209

Dupuis, P., Grenander, U., & Miller, M. I. (1998). Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, *56*(3), 587–600. Retrieved from http://dx.doi.org/10.1090/qam/1632326 doi: 10.1090/qam/1632326

Gatica, G. N. (2006). Analysis of a new augmented mixed finite element method for linear elasticity allowing $\mathbb{RT}_0$-$\mathbb{P}_1$-$\mathbb{P}_0$ approximations. *M2AN Mathematical Modelling and Numerical Analysis*, *40*(1), 1–28. Retrieved from http://dx.doi.org/10.1051/m2an:2006003 doi: 10.1051/m2an:2006003

Gatica, G. N. (2014). A Simple Introduction to the Mixed Finite Element Method. Theory and Applications. , xii+132. Retrieved from http://dx.doi.org/10.1007/978-3-319-03695-3 doi: 10.1007/978-3-319-03695-3

Gatica, G. N., Márquez, A., & Meddahi, S. (2008). A new dual-mixed finite element method for the plane linear elasticity problem with pure traction boundary conditions. *Computer Methods in Applied Mechanics and Engineering*, *197*(9-12), 1115–1130. Retrieved from http://dx.doi.org/10.1016/j.cma.2007.10.003 doi: 10.1016/j.cma.2007.10.003

Haker, S., Zhu, L., Tannenbaum, A., & Angenent, S. (2004, Dec 01). Optimal mass transport for registration and warping. *International Journal of Computer Vision*, *60*(3), 225–240. Retrieved from https://doi.org/10.1023/B:VISI.0000036836.66311.97 doi: 10.1023/B:VISI.0000036836.66311.97

Horn, B. K., & Schunck, B. G. (1980). *Determining Optical Flow* (Tech. Rep.). Cambridge, MA, USA.

Hurtado, D. E., & Henao, D. (2014). Gradient flows and variational principles for cardiac electrophysiology: toward efficient and robust numerical simulations of the electrical activity of the heart. *Computer Methods in Applied Mechanics and Engineering*, *273*, 238–254. Retrieved from http://dx.doi.org/10.1016/j.cma.2014.02.002 doi: 10.1016/j.cma.2014.02.002

Hurtado, D. E., Villarroel, N., Andrade, C., Retamal, J., Bugedo, G., & Bruhn, A. R. (2017). Spatial patterns and frequency distributions of regional deformation in the healthy human lung. *Biomechanics and Modeling in Mechanobiology*, *16*(4), 1413–1423. doi: 10.1007/s10237-017-0895-5

Hurtado, D. E., Villarroel, N., Retamal, J., Bugedo, G., & Bruhn, A. (2016, Feb). Improving the accuracy of registration-based biomechanical analysis: A finite element approach to lung regional strain quantification. *IEEE Transactions on Medical Imaging*, *35*(2), 580-588. doi: 10.1109/TMI.2015.2483744

Lee, E., & Gunzburger, M. (2010). An optimal control formulation of an image registration problem. *Journal of Mathematical Imaging and Vision*, *36*(1), 69–80. Retrieved from http://dx.doi.org/10.1007/s10851-009-0172-z doi: 10.1007/s10851-009-0172-z

Lee, E., & Gunzburger, M. (2011). Analysis of finite element discretizations of an optimal control formulation of the image registration problem. *SIAM Journal on Numerical Analysis*, *49*(4), 1321–1349. Retrieved from http://dx.doi.org/10.1137/090767674 doi: 10.1137/090767674

Logg, A., Mardal, K.-A., Wells, G. N., et al. (2012). *Automated solution of differential equations by the finite element method*. Springer. doi: 10.1007/978-3-642-23099-8

Modersitzki, J. (2004). *Numerical Methods for Image Registration*. Oxford University Press, New York. (Oxford Science Publications)

Museyko, O., Stiglmayr, M., Klamroth, K., & Leugering, G. (2009). On the application of the Monge-Kantorovich problem to image registration. *SIAM Journal on Imaging Sciences*, *2*(4), 1068–1097. Retrieved from http://dx.doi.org/10.1137/080721522 doi: 10.1137/080721522

Paquin, D., Levy, D., & Xing, L. (2008). Multiscale deformable registration of noisy medical images. *Mathematical Biosciences and Engineering. MBE*, *5*(1), 125–144.

Retrieved from http://dx.doi.org/10.3934/mbe.2008.5.125 doi: 10.3934/mbe.2008.5.125

Pöschl, C., Modersitzki, J., & Scherzer, O. (2010). A variational setting for volume constrained image registration. *Inverse Problems and Imaging*, *4*(3), 505–522. Retrieved from http://dx.doi.org/10.3934/ipi.2010.4.505 doi: 10.3934/ipi.2010.4.505

Rockafellar, R. T. (1976). Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, *14*(5), 877–898. Retrieved from http://dx.doi.org/10.1137/0314056 doi: 10.1137/0314056

Ruhnau, P., & Schnörr, C. (2007, Jan 01). Optical stokes flow estimation: an imaging-based control approach. *Experiments in Fluids*, *42*(1), 61–78. Retrieved from https://doi.org/10.1007/s00348-006-0220-z doi: 10.1007/s00348-006-0220-z

Schmidt-Richberg, A. (2014). *Registration Methods for Pulmonary Image Analysis: Integration of Morphological and Physiological Knowledge*. Springer Science & Business Media.

Sotelo, J., Urbina, J., Valverde, I., Mura, J., Tejos, C., Irarrazaval, P., ... Uribe, S. (2017). Three-dimensional quantification of vorticity and helicity from 3D cine PC-MRI using finite-element interpolations. *Magnetic Resonance in Medicine*. Retrieved from http://doi.wiley.com/10.1002/mrm.26687 doi: 10.1002/mrm.26687

Sotelo, J., Urbina, J., Valverde, I., Tejos, C., Irarrazaval, P., Andia, M. E., ... Hurtado, D. E. (2016). 3D Quantification of wall shear stress and oscillatory shear index using a finite-element method in 3D CINE PC-MRI data of the thoracic aorta. *IEEE Transactions on Medical Imaging*, *35*(6), 1475–1487. doi: 10.1109/TMI.2016.2517406

Sotiras, A., Davatzikos, C., & Paragios, N. (2013). Deformable medical image registration: A survey. *IEEE Transactions on Medical Imaging*, *32*(7), 1153–1190.

Steinbach, O. (2007). *Numerical Approximation Methods for Elliptic Boundary Value Problems: Finite and Boundary Elements*. Springer Science & Business Media.

Trouvé, A. (1998, Jul 01). Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, *28*(3), 213–221. Retrieved from https://doi.org/10.1023/A:1008001603737 doi: 10.1023/A:1008001603737

Unal, G., & Slabaugh, G. (2005, June). Coupled pdes for Non-Rigid Registration and Segmentation. In *2005 ieee computer society conference on computer vision and pattern recognition (cvpr'05)* (Vol. 1, p. 168-175 vol. 1). doi: 10.1109/CVPR.2005.115

Vese, L. A., & Le Guyader, C. (2016). *Variational Methods in Image Processing*. CRC Press, Boca Raton, FL.