Mixed finite element methods with applications to viscoelasticity and gels

Marie E. Rognes

© Marie E. Rognes, 2009

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
Nr. 851

ISSN 1501-7710

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.
Printed in Norway: AiT e-dit AS, Oslo, 2009.

Produced in co-operation with Unipub AS.
The thesis is produced by Unipub AS merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Unipub AS is owned by
The University Foundation for Student Life (SiO)

Preface

This thesis is submitted in partial fulfillment of the requirements for the degree of Philosophiae Doctor, to the Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Oslo. Most of the research has been conducted while residing at the Centre of Mathematics for Applications at the University of Oslo. In the period January – June 2007, I had the opportunity to work at the School of Mathematics, University of Minnesota, Twin Cities. The last two months, January – March 2009, were spent at the Center of Biomedical Computing, Simula Research Laboratory. I acknowledge, and sincerely express my gratitude to the Centre of Mathematics for Applications, the Research Council of Norway and to Simula Research Laboratory for their financial support. I am also grateful for the invitation from the School of Mathematics, University of Minnesota, which allowed for a fruitful and inspiring stay.

I have had the privilege of working with immensely talented people. First and foremost, I am eternally grateful to my advisor, Ragnar Winther, for his excellent guidance, clarity of vision, and soothing presence in moments of panic. My co-advisor, Hans Petter Langtangen, has provided many helpful comments and fabulous encouragement. I am truly grateful to Douglas Arnold for taking the time in the midst of a busy schedule, giving me many late Friday afternoons accompanied by true insight. I wish to thank Carme Calderer especially, for teaching me almost everything I know about continuum mechanics, and for introducing me to a diverse world of fascinating physical phenomena. Sincere thanks also goes to Robert Kirby for enlightening discussions and valuable feedback. Moreover, I do not think I can thank Anders Logg enough, for his everlasting enthusiasm, ingenuity, patience, and for writing the neatest code ever, setting a true example to follow. I also want to express my gratitude to Catherine Micek, for forcing me to explain things better, for great discussions and company, and last minute proofreading.

To my dad Sven Erik Rognes, and my brother John Rognes for always being there, and always believing in me. Finally, to my husband Ragnar, for everything.

Marie Elisabeth Rognes, Oslo, March 2009

Contents

Introduction 1
1. When solids and fluids mix 2
2. When finite elements mix 10
3. When compiling mixed forms 16
4. Overview of papers 18
5. Concluding remarks 22

Paper I: Mixed finite element methods for linear viscoelasticity using weak symmetry 29
1. Introduction 29
2. Viscoelastic models 32
3. Weak formulations and stability estimates 34
4. The semi-discrete problem and stability 39
5. Full-discretization 45
6. Numerical experiments 46
7. Conclusion 49

Paper II: Mixed finite element methods for gels with biomedical applications 55
1. Introduction 55
2. Governing equations of a gel 57
3. Existence of minimizing deformations 61
4. Linearizations of the Euler-Lagrange equations 63
5. Weak formulations and linear stability 65
6. Simulations 70

Paper III: Efficient assembly of H(div) and H(curl) conforming finite elements 79
1. Introduction 79
2. $H(\mathrm{div})$ and $H(\mathrm{curl})$ 80
3. $H(\mathrm{div})$ and $H(\mathrm{curl})$ conforming finite elements 85
4. Representation of $H(\mathrm{div})$ and $H(\mathrm{curl})$ variational forms 85
5. Assembling $H(\mathrm{div})$ and $H(\mathrm{curl})$ elements 90
6. Examples 93
7. Conclusions 97

Paper IV: Stability of Lagrange finite elements for the mixed Laplacian 101
1. Introduction 101
2. Notation and preliminaries 104
3. Eigenvalue problems related to the Babuška-Brezzi constants 106
4. Lower order Lagrange elements for the mixed Laplacian 108

Introduction

Biological and biomedical materials, such as muscles, soft tissue, blood vessels and artificial devices replacing the above, deform under changes in environment and applied traction. The stresses induced by such deformations may be of crucial importance as fractures or device failure can result from high stress concentrations. Given appropriate constitutive models, such scenarios lead themselves to mathematical study. As the resulting systems of equations can be highly complex, numerical approximations must typically be invoked. However, with efficient and accurate numerical methods available, simulations could, for instance, aid biomedical device design.

Most biological materials respond in a manner that is not purely elastic, in the sense that the stresses are not linearly related to the deformation only. Within some regimes, the viscoelastic response; that is, the gradual deformation and relaxation of stresses may be important. Moreover, when implanting artificial or synthetic devices in the human body, the moisture of the environment can affect the properties of the device. In both cases, the effect can be interpreted as a combination of solid-like and fluid-like behavior.

Methods for the approximation of deformations in elastic solids have been extensively studied for at least the last century. Accordingly and in pace with the computational capacities, advanced numerical methods have been developed over the last four decades. This thesis focuses on *mixed finite element methods* approximating the stress to a higher accuracy than the deformation. This viewpoint is motivated by the notion that the stress is the quantity of primary physical interest. Although carefully studied for the linear elasticity equations, such methods are not in widespread use for materials that display both solid and fluid features. However, this perspective constitutes the foundation for this thesis. In particular, extensions, of mixed finite element methods for elasticity, to viscoelasticity and gels are investigated within the linear regime. Moreover, the relative scarcity of these sophisticated numerical methods can also be attributed to their computational and implementational complexity. As a part of this investigation, the development of complexity-reducing computational tools has been an additional point of interest.

The purpose of this introduction is multifarious. First, it aims to motivate the problems and strategies considered. Second, the papers constituting this thesis involve a range of physical, numerical and computational notions and therefore, some background material is included. Also, it intends to underline the shared features of the papers and how these are connected. Thirdly, in order to indicate a bound for the scope of the thesis, limitations and some possibilities for extensions are discussed.

The introduction, preceding the collection of papers, is organized in the following manner. It consists of five sections. The first three sections provide background material for the physical, numerical and computational methods, respectively. More precisely, the continuum modeling of elasticity, viscoelasticity, and gels is reviewed in Section 1. Next, Section 2 treats the general theory behind mixed finite element methods, with particular emphasis on mixed finite element methods for linear elasticity. Section 3 discusses criteria, set by the combination of the mathematical models and numerical methods, on the computational framework, and focuses on a framework fulfilling these criteria at least to some extent. The fourth section provides a terse summary of the interpaper continuity and also a summary of each paper. Finally, a set of limitations and possible extensions are pointed at in Section 5.

Notation. Unless otherwise specified, the following notation is used.

- $\mathbb{V}, \mathbb{M}, \mathbb{S}, \mathbb{K}$**:** The linear spaces of vectors, matrices, symmetric matrices and skew-symmetric matrices, respectively.
- Ω**:** An open, bounded domain in $\mathbb{R}^d$, $d = 1, 2, 3$ with boundary $\partial\Omega$.
- $\mathcal{T}_h$**:** An admissible simplicial tessellation.
- $\mathcal{P}_r^c(\mathcal{T}_h)$**,** $\mathcal{P}_r(\mathcal{T}_h)$**:** Continuous piecewise polynomials and discontinuous piecewise polynomials of order r respectively defined relative to $\mathcal{T}_h$.
- $X(\Omega; Y)$**:** The space of fields $\Omega \mapsto Y$ with regularity properties specified by X. Example: $L^2(\Omega; \mathbb{K})$ is the space of square integrable skew-symmetric matrix fields on Ω.
- $X_0(\Omega; Y)$**:** The subspace of those $x \in X(\Omega; Y)$ such that $x|_{\partial\Omega_0} = 0$ for $\partial\Omega_0 \subseteq \partial\Omega$.
- $||\cdot||_X$**:** The norm defined on the normed space X.
- $\langle \cdot, \cdot \rangle_X$**:** The inner product defined on the inner product space X. Will also be abused to denote duality pairing when appropriate.
- $\nabla, \mathrm{div}, \mathrm{curl}$**:** . The gradient, the divergence, taken row-wise if applied to a matrix field, and the curl operators in two or three dimensions. Interpreted in the weak sense when appropriate.
- $\mathrm{tr}, \det$**:** The trace of a matrix tr; i.e, the sum of the diagonal components, and the determinant of a matrix det. A^T is the transpose of a matrix A. Finally, $\mathrm{cof}\, A = (\det A)A^{-T}$.

In general, norms and inner products without subscripts default to those on $L^2(\Omega)$. (An exception is in Paper II, where $||\cdot||$ defaults to the Frobenius norm on $\mathbb{R}^n$.)

1. When solids and fluids mix

Nature offers a wide spectrum of behavior originating from solid and fluid interactions. Consequently, there also exists a multitude of theories and mathematical models describing solid and fluid interplay within different regimes. For the class of *viscoelastic* materials, the standard models rely on the display of both solid and fluid characteristics on the macroscopic scale. For solid-fluid mixtures, such as polymer-solvent mixtures constituting *gels*, the modeling can be based on the interplay between the solid deformation and the microscopic effects of the mixing between the solid and fluid. The partial differential equations describing the equilibrium behavior of both classes are intrinsically framed by the general setting of continuum mechanics.

Papers I and II of this thesis rely on partial differential equations for viscoelasticity and gels respectively. The material of this section is aimed at providing a mechanical foundation for these equations. The basic concepts and assumptions involved in the modeling of solids are reviewed in Section 1.1. The constitutive

relations for viscoelasticity and gels, and hence the specialization of the general theory to those classes of materials, follow in Sections 1.2 and 1.3 respectively.

1.1. **From continuum mechanics to linear elasticity.** This section is in essence a brief tour, departing from the general continuum mechanics setting and arriving at the equations of linear elasticity. The selection of material is biased by relevance with regard to Papers I and II, and also by a numerical viewpoint rather than a mechanical one. An excellent introduction to classical continuum mechanics, including proofs of the claims presented in this section, can be found in Gurtin [**32**]. Also, Ciarlet [**19**] is an authoritative reference for the more mathematical aspects of elasticity.

Let $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$ be an open and bounded domain occupied by some body or medium with boundary $\partial\Omega$. The issue of interest is the governing equations for the deformation of this body: $\Phi : \Omega \to \Phi(\Omega)$, under applied boundary or body forces. Denoting coordinates in Ω and $\Phi(\Omega)$ by X and x respectively, Cauchy's theorem states that under the balance laws of linear and angular momentum, there exists a *symmetric* tensor field over $\Phi(\Omega)$, the Cauchy stress tensor $\mathcal{T}$. In the static equilibrium with a given body force $\tilde{g}$, $\mathcal{T}$ satisfies

$$\mathcal{T}(x) \cdot n_x = s_n(x), \quad -\operatorname{div}_x \mathcal{T}(x) = \tilde{g}(x), \quad x \in \Phi(\Omega), \tag{1.1}$$

where s_n denotes the stress acting on the plane defined by the normals n and the divergence is defined row-wise. Denote the gradient of deformation by $F = \nabla_X \Phi$. A transform of (1.1) onto the undeformed domain Ω, gives the existence of another stress tensor, the first Piola-Kirchhoff stress tensor $\mathcal{S}$ over Ω such that

$$\mathcal{S}(X) \cdot n_X = s_n(X), \quad -\operatorname{div}_X \mathcal{S}(X) = g(X), \quad X \in \Omega. \tag{1.2}$$

The Cauchy and first Piola-Kirchhoff stress tensors are related by the identity

$$\mathcal{T}(x) = (\det F)^{-1} \mathcal{S}(X) F^T, \quad x = \Phi(X). \tag{1.3}$$

From (1.3) and the symmetry of the Cauchy stress tensor, it is easy to see that $\mathcal{S}$ will not be symmetric in general, but must satisfy the relation

$$\mathcal{S}F^T = F\mathcal{S}^T, \tag{1.4}$$

for any admissible deformation gradient F. However, the stress tensor associated with the identity deformation must be symmetric.

The existence and above properties of these stress tensors are independent of the type of material occupying Ω. Hence, in order to characterize the material response of the body, the governing equations (1.1) or (1.2) are complemented by constitutive equations relating the stress tensor to the deformation. For ideally elastic materials, the stress tensor is assumed to be directly related to the deformation gradient. In the viscous case, the stress tensor is typically linked to the rate of deformation gradient. The remainder of this subsection focuses on elastic constitutive relationships, and in particular on hyperelastic materials. All equations in the following are with reference to the undeformed domain Ω, unless explicitly stated, and so the reference to the coordinates is omitted.

By definition, a material is *hyperelastic* if the first Piola-Kirchhoff stress tensor $\mathcal{S}$ is the Fréchet derivative with respect to the deformation gradient of a real-valued function ω, labelled the strain-energy density:

$$\mathcal{S} = \mathcal{S}(F) = \frac{\partial \omega}{\partial F}(F). \tag{1.5}$$

The elastic energy $\mathcal{E}$ associated with a deformation, can then be expressed as

$$\mathcal{E}(F) = \int_\Omega \omega(F) \, \mathrm{d}X. \tag{1.6}$$

The standard assumption that the response of the material should be independent of a change of observer or frame, implies that ω must depend on the gradient of deformation through the right Cauchy-Green strain tensor $C_G = F^T F$ (or the left FF^T). Moreover, if the material is isotropic; that is, if the response of the material is invariant under all rotations of the body, the energy density must be a function of the invariants of C_G:

$$\omega(F) = \tilde{\omega}(C_G) = \hat{\omega}(\iota_1(C_G), \iota_2(C_G), \iota_3(C_G))$$

where $\iota_1(C) = \operatorname{tr}(C)$, $\iota_2(C) = \operatorname{tr}\operatorname{cof}(C)$ and $\iota_3(C) = \det C$. The additional assumption, if invoked, that the material is incompressible; i.e. that the density of the material must be constant under deformations, yields the local constraint:

$$\det F = 1. \tag{1.7}$$

A variety of strain energy densities proposed for compressible and incompressible elastic materials can be sampled from [**19, 36**]. One example is the compressible neo-Hookean model:

$$\omega(F) = \mu \left(\frac{1}{2} \operatorname{tr}(FF^T - I) + \frac{1}{\beta}((\det F)^{-\beta} - 1) \right), \tag{1.8}$$

where μ and β are material parameters and I denotes the identity matrix.

Formally, the equilibrium deformation, and hence the associated equilibrium position of the body, corresponds to a minimizer of the elastic energy (1.6), augmented by conservative body forces g and boundary forces s, in some space of admissible vector fields. In the pioneering works of Ball [**8**], existence of minimizers was established under the assumption of polyconvexity of ω and under certain growth conditions. A proper presentation is beyond the scope of this introductory chapter. However, in addition to the aforementioned [**19**], an introduction to the topic is provided in [**24**, Section 8]. Assuming sufficient smoothness of ω and the minimizing deformation Φ, the stationary points of the energy will satisfy the Euler-Lagrange equations:

$$\mathcal{S} = \mathcal{S}(\nabla \Phi), \quad -\operatorname{div} \mathcal{S} = g, \tag{1.9}$$

usually complemented by boundary conditions of the form $\Phi|_{\partial\Omega_0} = \Phi_0$, $\mathcal{S}\cdot n|_{\partial\Omega_1} = s$ where $\partial\Omega_0 \cup \partial\Omega_1 = \partial\Omega$ and $\partial\Omega_0 \cap \partial\Omega_1 = \emptyset$.

In the case of small deformation gradients, the linearization of (1.9) is pertinent and useful. The linearized equations can be obtained by letting $F = F_0 + \nabla u$ and taking the affine terms of the Taylor expansion of $\mathcal{S}(F)$ about F_0:

$$\sigma = \mathcal{S}(F_0) + D\mathcal{S}(F_0)\nabla u, \quad -\operatorname{div} \sigma = g. \tag{1.10}$$

Here, $D\mathcal{S}(F_0)$ is the Gateaux derivative of $\mathcal{S}$ at F_0. For hyperelastic materials, the energy density is usually easily scaled to ensure that the reference state is stress-free and linearizations about this state are therefore natural. The elasticity operator $C = D\mathcal{S}(I)$ identifies a fourth order *stiffness* tensor. In general, if the linearization state F_0 is stress-free, it can be shown that the stiffness tensor is completely determined by its restriction to symmetric tensors. Consequently, the standard linear elasticity equations take the following form.

$$\sigma = C\varepsilon(u), \tag{1.11a}$$

$$-\operatorname{div} \sigma = g, \tag{1.11b}$$

where ε is the linearized strain, or more precisely, the symmetrized gradient operator: $2\varepsilon(u) = \nabla u + \nabla u^T$. Observe that (1.11) can also be viewed as the Euler-Lagrange equations of the following constrained minimization problem

$$\min_{\sigma} \frac{1}{2} \int_\Omega C^{-1}\sigma \cdot \sigma \, \mathrm{d}X \quad \text{subject to} \quad \operatorname{div} \sigma = g. \tag{1.12}$$

For hyperelastic materials with sufficient smoothness of the energy density, the tensor C will be symmetric. Hence, the assumption that the linearization state is stress-free, gives that the linearized stress tensor σ is symmetric. For isotropic materials, it can further be shown that the stiffness tensor C reduces to the form:

$$C\varepsilon = 2\mu\varepsilon + \lambda \operatorname{tr}\varepsilon I, \tag{1.13}$$

for scalar coefficients μ and λ, known as the Lamé parameters. In particular, standard calculations give that the linearization of the compressible neo-Hookean strain density (1.8) reduces to (1.13) with $\lambda = \beta\mu$.

If the linearization state is *not* stress-free, then the linearized first Piola-Kirchhoff stress tensor σ will not in general be symmetric. This aspect will be important in Paper II. However, the linearization of the constraint (1.4) about the reference state, gives a skew symmetry constraint:

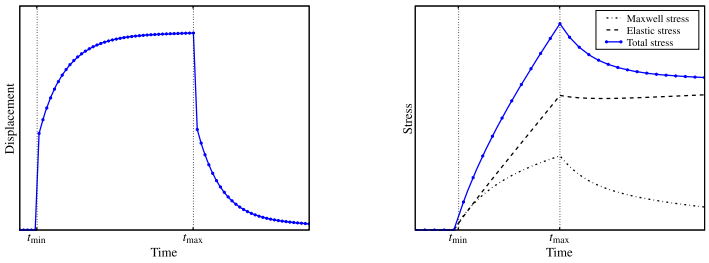
$$\sigma - \sigma^T = \mathcal{S}(I)(\nabla u - \nabla u^T). \tag{1.14}$$

In other words, the skew symmetric part of the stress tensor is proportional to the skew symmetric part of the displacement gradient.

In the subsequent two subsections, the behavior characteristics of viscoelastic materials and gels, respectively, are discussed and constitutive equations for these materials presented. Both types of constitutive equations can be viewed as extensions of the linear elasticity equations above.

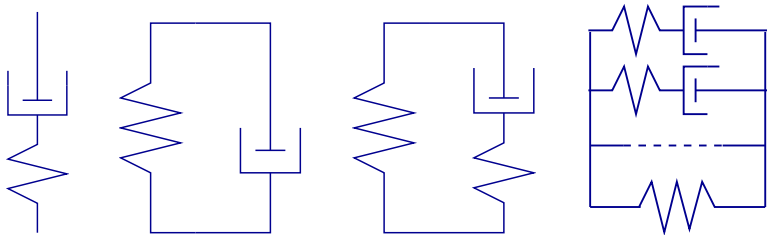
1.2. **Viscoelasticity.** Viscoelastic materials are named by their ability to display both viscous and elastic behavior. Whenever a body is subjected to an external force or deformation, the body responds by rearrangements of its microscopic constituents. In idealized viscous fluids, the time required for the rearrangement is assumed to be infinitely small. For ideally elastic solids, the time is assumed to be infinitely long. However, in any physical material, these rearrangements must take some finite time. Hence, most real-life materials demonstrate some viscoelastic properties. These effects may be particularly important when considering synthetic polymers or biological materials such as muscles or soft tissue. For a more thorough introduction to viscoelastic behavior and modeling than the material presented here, the monographs [**23, 25, 28, 60**] provide ample material. The history of viscoelastic modeling dates back to the last half of the nineteenth century and the works of Boltzmann, Kelvin, Maxwell and Weichert [**22**].

The two main characteristics of viscoelastic behavior are the stress response of the material under an induced strain: *relaxation*, and the strain response of the material under an induced stress: *creep* or *retardation*. Different materials can thus be identified by their relaxation and creep response. To make ideas more concrete, consider a body of some material and the situation where a force is applied to a part of the boundary, kept constant for some time, and then removed. Ideal, linear elastic solids will react with an instantaneous deformation, and only that deformation as long as the force is kept constant, since the strain is directly proportional to the stress cf. (1.11a). On the other hand, the stress of viscous fluids is assumed to be proportional to the rate of change of the strain. Such materials will thus not deform instantaneously, but rather with a constant velocity. Furthermore, after the force is removed, elastic solids will return to their initial state, while viscous fluids tend to display no such behavior. Viscoelastic materials exhibit a combination of both behavioral patterns. For instance, a viscoelastic material might react with a certain instantaneous deformation, and then continue flowing up to a deformation limit. After removal of the force, the material might again react with a certain instantaneous deformation and after some, possibly infinite, time return to the initial configuration. A material model describing this type of behavior is that



(a) (b)

FIGURE 1. Simulation of creep and relaxation behavior for the Standard linear solid. (a) Deformation versus time under a constant boundary force from $t_{\min}$ to $t_{\max}$. (b) Stress versus time under a constant boundary velocity between $t_{\min}$ and $t_{\max}$. For more details, see Paper I.



(a) (b) (c) (d)

FIGURE 2. Schematic representation of viscoelastic models: (a) Maxwell, (b) Kelvin-Voigt, (c) Standard linear solid, and (d) Generalized Maxwell.

of the *Standard linear solid* model and the *creep* and *relaxation* patterns of this material are illustrated in Figure 1.

The viscoelastic response may be interpreted as a sum of elastic and viscous factors. This interpretation motivates the viscoelastic modeling approach based on combinations of elastic and viscous influences, schematically represented as springs and dashpots in some special arrangement. The corresponding constitutive equations may be derived from the compatibility and equilibrium conditions of the physical systems represented. Generalizations to two or three dimensions can be performed by considering deviatoric and volumetric contributions separately. The simplest manner in which to schematically construct a viscoelastic model is to combine one elastic and one viscous component either in series or in parallel. The resulting models are known as the Maxwell and the Kelvin-Voigt model or element respectively. These models are simplicial, but can viewed as building-blocks for more realistic models. The schematical representations of the Kelvin-Voigt, Maxwell, Standard linear solid (or Zener), and generalized Maxwell models are illustrated in Figure 2.

The generalized Maxwell models form an established class of viscoelasticity models. These models, and variations in the form of the generalized Zener models,

have been a common starting-point for several discretization approaches [**9, 39, 52**]. Further, one could argue that any viscoelastic model derived from a linear spring-dashpot combination, can be reduced to an equivalent series-parallel model. Such reductions are discussed in detail by Tschoegl [**60**, p. 135 ff.]. In particular, the standard series-parallel models can be transformed equivalently into this class. Viscoelastic models of this form were therefore considered as an appropriate and sufficiently general framework in Paper I.

Schematically speaking, the models considered are those that can be represented in terms of n parallel branches, each branch either consisting of a spring, a dashpot or a simple Maxwell element. Accordingly, a separate stress component σ_j may be associated with each branch j. The total stress follows as the sum of the stress components. The constitutive equations of these models can be expressed by the following system of differential equations in space and time for the total stress $\sigma = \sigma(X,t)$ and the displacement $u = u(X,t)$, or the velocity $v = \dot{u}$.

$$A_E^j \dot{\sigma}_j + A_V^j \sigma_j = \varepsilon(\dot{u}) \quad j = 1, \ldots, n, \quad \sigma = \sum \sigma_j, \tag{1.15}$$

where n is given by the number of parallel branches and the superimposed dot denotes the time derivative. These are evolution equations for the components of the total stress. Either of the fourth order compliance tensors A_E^j and A_V^j may vanish for some js. Typically, $A = C^{-1}$, where C is a stiffness tensor such as in (1.11a). The compliance tensors may be spatially dependent, but is in this setting assumed to be independent of time. For instance, the constitutive equation for the Maxwell model reads

$$A_E \dot{\sigma} + A_V \sigma = \varepsilon(\dot{u}),$$

and the Standard linear solid model can take the form

$$A_E^1 \dot{\sigma}_1 + A_V^1 \sigma_1 = \varepsilon(\dot{u}), \quad A_E^2 \dot{\sigma}_2 = \varepsilon(\dot{u}), \quad \sigma = \sigma_1 + \sigma_2. \tag{1.16}$$

It is easy to see that (1.16) can be expressed in the equivalent, more traditional form

$$A_E \dot{\sigma} + A_V \sigma = C_V \dot{\varepsilon}(\dot{u}) + C_E \varepsilon(u)$$

cf. [**20**, p. 52].

In addition to being a unifying form for generalized viscoelasticity models, (1.15) is particularly suitable for discretizations approximating the stress and the displacement separately. For the linear elasticity equations, such discretizations typically rely on an inversion of the stress-strain relationship (1.11a). The strain-stress structure of (1.15) gives the desired property. In fact, the specific formulation (1.15) was designed for that purpose in Paper I. A more thorough discussion of these constitutive equations is included in that work. The formulation can also be compared to the popular elastic-viscous stress splitting formulations for discretizations of viscoelastic flow [**7**].

Before continuing, note that there exists a plethora of different, but equivalent, formulations of constitutive equations for linear viscoelasticity. More traditional discretization approaches often seek to eliminate the stress using (1.11b), and approximate the displacement only. For such strategies, other forms of the constitutive equations, such as stress-strain formulations involving multiple strain components [**39**], may be more appropriate. Moreover, the classical, more general interpretation and modeling of viscoelastic materials as *materials with memory* should be mentioned. This approach gives rise to integro-differential constitutive equations for the stress in terms of the displacements:

$$\sigma(t) = C(t)\varepsilon(u(0)) + \int_0^t C(t-s)\,\varepsilon(\dot{u}(s))\,\mathrm{d}s. \tag{1.17}$$

Here, the time dependent tensor C is known as the relaxation tensor. An assumption of fading memory mandates that C is monotonically decreasing in time. The constitutive equation (1.17) can be derived starting from the linear elasticity equation (1.11a), allowing the stiffness tensor C to depend on time, and taking the limit of an incremental strain history. A thorough derivation and discussion of these models can be found in [**25, 55**]. Note that the integro-differential form also allows for strain-stress formulations:

$$\varepsilon(t) = A(t)\sigma(0) + \int_0^t A(t-s)\,\dot{\sigma}(s)\,\mathrm{d}s. \tag{1.18}$$

Both (1.17) and (1.18) take the form of Volterra equations and have been extensively studied from that viewpoint, see for instance the aforementioned [**25**] for a mathematical treatise. The equivalence between the differential models, and the integral models with exponentially decaying relaxation tensors, was carefully studied in [**33**].

1.3. **Gels.** By definition, *gels* consist of a cross-linked polymer network, or a polymer entanglement, immersed in some solvent. The two components of the gel coexist by balancing each other. The polymer network confines the liquid solvent, while the solvent ensures that the polymer does not collapse into a dry state. Thus, gels typically have the appearance of solids, but may in fact possess radically different properties. A change of shape and volume by expansion or contraction is referred to as a *swelling* of the gel and some gels can swell by several orders of magnitude. Note that these volume changes are typically reversible.

The applications of gels are wide-ranging, and a diverse survey is presented in [**49**, Vol. 3]. Human soft tissue is a prime example of a natural gel, while the term synthetic gels usually refers to gels constructured from synthetic polymers. A common example of the latter is the soft contact lens, which stays flexible while in a moist environment, but becomes brittle upon drying. Other biomedical inventions such as artificial bone replacements rely on synthetic polymer implants. Such may swell upon insertion in the human body with unfavorable consequences. Further, implantable drug-delivery devices take advantage of the stimulus-responsive mechanisms of gels. The latter usage has coined the label *smart materials*, often used for polyelectrolyte gels because of their response to changes in pH, temperature, applied voltage or solvent concentration.

Now, what are these stimulus-responsive mechanisms of gels? The term relates to the effect of gel phase transitions. By phases of a gel, one refers to the configuration of the polymer network. A gel phase transition is a finite change in volume, a change from a swollen to a collapsed state or vice versa, in response to an infinitesimal change in the environment. Note that the change of volume is abrupt only in the sense that tiny environmental changes gives finite changes in volume: the actual deformation is usually gradual. The first experimental evidence of phase transitions in gels was reported in experiments on hydrolyzed, cross-linked polyacrylamide by Tanaka [**58**] in the late 1970s.

Tanaka identified and labelled three forces acting within the gel: *Rubber elasticity*, *polymer-polymer affinity*, and *hydrogen ion pressure*; and referred to the sum of these as the *osmotic pressure*. The balance of these forces, of different directions, ranges and magnitude, gives the possibility of phase transitions. The first force, the rubber elasticity, represents the elasticity of the polymer strands in the coupled network. The second, the polymer-polymer affinity, results from the interaction of the polymer and the solvent. Depending on the specific material properties of these, the polymer strands may or may not prefer interaction with the solvent.

Thus, the energy of the network may be reduced by a realignment due to more favorable conditions in the polymer-solvent configuration. This force is a short-range force and is therefore more prominent as the volume of the gel decreases. The third relates to the ionization of the network. In the process of hydrolysis, the polymer strands become negatively charged by expelling hydrogen ions into the solvent. On the whole, the gel can be neutral, but the distribution of the charges yields an additional pressure due to electric attraction or repulsion.

These behavior characteristics must be accounted for in the modeling of gels. In particular, the elasticity of the polymer network, the mixing of the polymer and solvent and the convection and diffusion of the solvent and ions within the network are important features. In the previous subsection, we saw that although there is a multitude of different formulations for linear viscoelasticity, all mentioned can be viewed as more or less within the same equivalence class. Moreover, the basic formulations have been agreed upon for at least the last five decades. In the modeling of gels, this is not the case. In particular, the dynamical modeling of gels is still an active topic [**17, 21, 37**]. However, for the static equilibrium modeling of non-ionic gels, a common starting-point seem to be the works of Flory and Rehner [**29, 30**], and Huggins [**38**] on polymer solution theory from the 1940s. Such equilibrium equations are the focus of Paper II and some of the very basic concepts are therefore introduced here. More details can be found in Paper II, and the references therein.

The equilibrium modeling of gels as a mixture of polymer and solvent takes the viewpoint of both constituents occupying the same domain in space and time. A scalar field measuring the volume of component i per unit volume of the gel, denoted the volume fraction ϕ_i, is associated with each component i. The volume fraction ϕ_i can also be interpreted as the probability of finding constituent i at a given point in space at a given time. The assumption that there will be no other constituents than polymer and solvent and no voids in the domain, gives the constraint

$$\phi_1(x) + \phi_2(x) = 1 \quad x \in \Phi(\Omega). \tag{1.19}$$

Note that the modeling of biphasic soft tissue can also be based on these notions, but the Flory-Huggins theory is usually not invoked for such materials.

The previous discussion on the forces acting on a non-ionic gel motivates the modeling of its free energy (cf. (1.6)) as the sum of two contributions: an elastic energy associated with the polymer network and a mixing energy accounting for the interaction of polymer and solvent. Assuming that the polymer is hyperelastic and that its energy density depends on the volume fraction of the polymer and the gradient of deformation F, one obtains an elastic energy density of the form:

$$\omega_E = \omega_E(\phi_1 \circ \Phi(X), F(X)). \tag{1.20}$$

The theory of Flory and Huggins gives the following form of the mixing energy density due to interaction between polymer and solvent in the deformed configuration $\Phi(\Omega)$:

$$\omega_{FH}(\phi_1, \phi_2) = a\phi_1 \ln \phi_1 + b\phi_2 \ln \phi_2 + c\phi_1\phi_2 \tag{1.21}$$

where a, b and c are material parameters accounting for the relative difference in size of polymer and solvent molecules, the cross-linking of the polymer, temperature and the propensity of the polymer and solvent to mix. In particular, c regulates the mixing affinity of polymer and solvent. The parameter values also control the convexity of ω_{FH}. If c is sufficient large in comparison with a and b, ω_{FH} is non-convex. In total, the combined gel energy $\mathcal{E}$ reads

$$\mathcal{E}(F, \phi_1, \phi_2) = \int_\Omega \omega_E \, \mathrm{d}X + \int_{\Phi(\Omega)} \omega_{FH} \, \mathrm{d}x \tag{1.22}$$

Paper II is mainly targeted at the case in which (1.22) is polyconvex. However, note that if the Flory-Huggins energy density (1.21) is non-convex, depending on the relative magnitude of the elastic and mixing parameters, the total energy density (1.22) may fail to be polyconvex. In this case, there may be multiple minimizers of the total energy, corresponding to different phases of the gel. This mathematical phenomenon can be interpreted in relation with the physical phase transitions of gels.

The assumption that the polymer is incompressible and that the reference domain Ω is occupied by a volume fraction ϕ_I of polymer, offers a closure of the system of equations. In particular, the balance of mass of polymer yields the local constraint

$$\phi_1(x) \det F(X) = \phi_I(X) \quad X \in \Omega, x = \Phi(X). \tag{1.23}$$

Note that the polymer incompressibility assumption for the mixture does *not* preclude deformations with change of volume, in contrast to the purely elastic case cf. (1.7). This observation enforces the notion that the gel mixture must be viewed as fundamentally distinct from a mere juxtapositioning of the constituents. The problem of minimizing (1.22) subject to (1.19) and (1.23) provides the governing and constitutive equations of Paper II.

2. When finite elements mix

As noted in connection with (1.12), the linear elasticity equations (1.11) can also be viewed as the Euler-Lagrange equations for the minimization of an energy subject to a constraint. The equations define a saddle point problem, where the displacement u can be interpreted as a Lagrange multiplier associated with the divergence constraint (1.11b). Other, physically fundamental, saddle point problems include the equations of Darcy or Stokes flow and Maxwell's equations of electromagnetism. Such saddle point problems provide the motivation for and foundations of mixed finite element methods and their analysis.

The quest for stable, mixed discretizations of saddle point problems, has been a main underlying current for the research on mixed finite element methods over the last three decades. Alternative discretization methods include stabilization techniques or variational crimes in the form of non-conforming methods. However, conforming mixed finite element methods can be viewed as especially attractive, in the sense that properties of the continuous equations are often respected and mimicked in the discretization.

Just as the previous section was a partial introduction to the physics and modeling of elasticity, viscoelasticity and gels, this section surveys the theory of mixed finite element methods with special attention to discretizations of the linear elasticity equations. The material is classical and mainly based on [**11, 16, 26**]. However, special emphasis is placed on aspects of the theory being particularly relevant for Papers I, II and IV: the extended saddle point problem, the idea of reduced stability, weak formulations for linear elasticity, and finally, an overview of stable finite element spaces for the weakly symmetric formulation of linear elasticity.

2.1. Abstract saddle point problems and their discretization.

Let Σ and V be Hilbert spaces with dual spaces Σ^* and V^*, and let a and b be continuous bilinear forms on $\Sigma \times \Sigma$ and $\Sigma \times V$, respectively. The extended abstract saddle point problem takes the form: Given $(f, g) \in \Sigma^* \times V^*$, find $(\sigma, u) \in \Sigma \times V$ satisfying

$$\begin{aligned} a(\sigma, \tau) + b(\tau, u) &= \langle f, \tau \rangle \quad \forall\, \tau \in \Sigma, \\ b(\sigma, v) &= \langle g, v \rangle \quad \forall\, v \in V. \end{aligned} \tag{2.1}$$

For simplicity, assume that a is symmetric. The term *saddle-point* refers to the fact that, if a is positive definite, (σ, u) is indeed a saddle point of the Lagrangian defined by a and b. Anticipating events, define the kernel Z by:

$$Z = \{\tau \in \Sigma \,|\, b(\tau, v) = 0 \quad \forall\, v \in V\}. \tag{2.2}$$

The now classical mixed finite element theory, based on the works of Brezzi in the early 1970s [**13**], relates the existence and stability of solutions to (2.1) to the Brezzi conditions. The two Brezzi conditions prescribe the existence of positive constants α, β satisfying:

$$0 < \alpha \leq \inf_{0 \neq \sigma \in Z} \sup_{0 \neq \tau \in Z} \frac{a(\sigma, \tau)}{||\sigma||_\Sigma ||\tau||_\Sigma}, \tag{2.3a}$$

$$0 < \beta \leq \inf_{0 \neq v \in V} \sup_{0 \neq \tau \in \Sigma} \frac{b(\tau, v)}{||\tau||_\Sigma ||v||_V}. \tag{2.3b}$$

(These conditions are also known in the literature as the Babuska-Brezzi conditions, Ladyshenskaya-Babuska-Brezzi conditions or simply the inf-sup conditions.) If (2.3) is satisfied, Brezzi's splitting theorem states that there exist $\sigma \in \Sigma$ and $u \in V$ satisfying (2.1) and a positive constant c, depending on α, β and the continuity of a, b, such that

$$||\sigma||_\Sigma + ||u||_V \leq c\,(||f|| + ||g||)\,. \tag{2.4}$$

The constants α and β will here be referred to as the continuous Brezzi coercivity and Brezzi inf-sup constant, respectively. Informally speaking, condition (2.3a) ensures that the minimization problem $\min_\sigma \frac{1}{2} a(\sigma, \sigma) - \langle f, \sigma \rangle$ has a stable solution when viewed over the constrained subspace Z. The second condition (2.3b) gives the existence of the Lagrange multiplier u associated with the constraint defining Z.

Next, we consider a *conforming* discretization of the equations defined by (2.1) over Ω. To this end, let $\Sigma_h \subset \Sigma$ and $V_h \subset V$ be finite dimensional spaces defined relative to a tessellation $\mathcal{T}_h$ of Ω. The discretization of (2.1) by $\Sigma_h \times V_h$ reads: Find $(\sigma_h, v_h) \in \Sigma_h \times V_h$ satisfying

$$\begin{aligned} a(\sigma_h, \tau) + b(\tau, u_h) &= \langle f, \tau \rangle \quad \forall\, \tau \in \Sigma_h, \\ b(\sigma_h, v) &= \langle g, v \rangle \quad \forall\, v \in V_h. \end{aligned} \tag{2.5}$$

The discrete analogies of (2.3) reads

$$0 \leq \alpha_h = \inf_{0 \neq \sigma \in Z_h} \sup_{0 \neq \tau \in Z_h} \frac{a(\sigma, \tau)}{||\sigma||_\Sigma ||\tau||_\Sigma}, \tag{2.6a}$$

$$0 \leq \beta_h = \inf_{0 \neq v \in V_h} \sup_{0 \neq \tau \in \Sigma_h} \frac{b(\tau, v)}{||\tau||_\Sigma ||v||_V}, \tag{2.6b}$$

where the discrete kernel is defined as

$$Z_h = \{\tau \in \Sigma_h \,|\, b(\tau, v) = 0 \quad \forall\, v \in V_h\}.$$

We shall refer to the values α_h and β_h as the Brezzi coercivity and Brezzi inf-sup value respectively.

The existence and uniqueness of the discrete solutions (σ_h, u_h) to (2.5) follow from the existence of lower bounds for α_h and β_h [**13**]. This result motivates the notion of stability for a mixed discretization. A family of discretizations $\{\Sigma_h \times V_h\}_h$, parameterized by h, is labelled *stable* in the $\Sigma \times V$ norm if α_h and β_h defined by (2.6) are bounded from below by positive constants uniformly in h. Furthermore, if $\{\Sigma_h \times V_h\}_h$ is stable, the quasi-optimal approximation properties of (σ_h, u_h) follow. Namely, there exists a positive constant c such that

$$||\sigma - \sigma_h||_\Sigma + ||u - u_h||_V \leq c \left(\inf_{\tau \in \Sigma_h} ||\sigma - \tau||_\Sigma + \inf_{v \in V_h} ||u - v||_V \right) \tag{2.7}$$

The error estimate (2.7) can be extensively refined under additional assumptions, but for any details on this matter, the reader is referred to for instance [**16**].

Before proceeding, recall the two canonical examples of partial differential equations for flow problems. First, the mixed formulation of the Laplace equation, also known as the Darcy flow equations, with pure Dirichlet boundary conditions for the scalar variable, reads

$$\begin{aligned} \langle \sigma, \tau \rangle + \langle \operatorname{div} \tau, u \rangle &= 0 \qquad \forall\, \tau \in H(\operatorname{div}, \Omega; \mathbb{V}), \\ \langle \operatorname{div} \sigma, v \rangle &= \langle g, v \rangle \quad \forall\, v \in L^2(\Omega). \end{aligned} \tag{2.8}$$

Here, $H(\operatorname{div}, \Omega; \mathbb{V})$ is the space of square integrable vector fields on Ω with square integrable divergence. In order to satisfy the conditions (2.6), finite element discretizations $\{\Sigma_h \times V_h\}_h$ of these equations usually rely on the use of element spaces Σ_h that are $H(\operatorname{div})$ conforming, but not H^1 conforming. In particular, the classical spaces consist of piecewise polynomial vector fields with continuous normal components [**14, 47, 48, 51**]. The use of piecewise polynomials of one order lower with no continuity restrictions for V_h ensures that $\operatorname{div} \Sigma_h \subseteq V_h$, and (2.6a) follows naturally. The use of H^1-conforming elements for the space Σ_h, that is, continuous piecewise vector polynomials: $\Sigma_h = \mathcal{P}_k^c(\mathcal{T}_h)$, is explored experimentally in Paper IV.

Second, the Stokes equations, with pure Dirichlet boundary conditions for the vector variable, can be phrased as

$$\begin{aligned} \langle \varepsilon(\sigma), \varepsilon(\tau) \rangle + \langle \operatorname{div} \tau, u \rangle &= \langle f, \tau \rangle \quad \forall\, \tau \in H_0^1(\Omega; \mathbb{V}), \\ \langle \operatorname{div} \sigma, v \rangle &= 0 \qquad \forall\, v \in L^2(\Omega),\ \langle v, 1 \rangle = 0. \end{aligned} \tag{2.9}$$

Recall that ε denotes the symmetrized gradient. For discretizations $\{\Sigma_h \times V_h\}_h$ of (2.9), condition (2.6a) will always be fulfilled by Korn's inequality. The condition (2.6b) is however non-trivial. Standard stable discretizations include the (generalized) Taylor-Hood elements $\mathcal{P}_r^c \times \mathcal{P}_{r-1}^c$ or the $\mathcal{P}_r^c \times \mathcal{P}_{r-2}$ elements for $r = 2, 3, \ldots$ [**15, 59**]. A natural question becomes whether there are any conforming discretizations $\{\Sigma_h \times V_h\}_h$, where Σ_h and V_h are associated with the same simplicial tessellation, that are stable for both (2.8) and (2.9). A partial answer is that none is known to exist for polynomial degrees lower than 3. For more details on this matter, see for instance Paper IV and references therein.

The abstract saddle point setting defined by (2.1) can be extended to the perturbed form:

$$\begin{aligned} a(\sigma, \tau) + b(\tau, u) &= \langle f, \tau \rangle \quad \forall\, \tau \in \Sigma, \\ b(\sigma, v) - \delta \langle u, v \rangle_V &= \langle g, v \rangle \quad \forall\, v \in V, \end{aligned} \tag{2.10}$$

where δ is some real parameter. As we shall see in the following, this case arises in connection with nearly incompressible elasticity and nearly symmetric elasticity. For a careful analysis of this and the more general case where the term with coefficient δ is replaced by some general continuous bilinear form on $V \times V$, confer [**16**]. In short, one can show that if δ is positive, or non-positive, but appropriately small; or sufficiently negative, well-posedness of (2.1) ensures well-posedness of (2.10) and hence an estimate of the form (2.4).

Further, assuming that δ is such that the continuous problem is well-posed, the question of stability, and hence approximation properties, of a discretization of (2.10) reduces to that of (2.1).

2.2. **Reduced stability and Brezzi eigenvalue problems.** There are families of discretizations $\{\Sigma_h \times V_h\}_h$ that are not stable in the sense defined above, but have a reduced stability property. In particular, for the Stokes equations (2.9), such elements have been studied exhaustively [**10, 50**]. Paper IV is devoted to such considerations for the mixed Laplacian (2.8).

In order to motivate such a notion consider a saddle point of the form (2.1) as before, and given discrete spaces Σ_h and V_h. The space of *spurious modes* $N_h \subset V_h$ is defined as follows:

$$N_h = \{v \in V_h \,|\, b(\tau, v) = 0 \quad \forall\, \tau \in \Sigma_h\}.$$

It is easy to see that if N_h contains non-zero elements, any solution (σ_h, u_h) to (2.5) could not be unique and $\beta_h = 0$. However, one might hope that the discretization could be stabilized by removal of these spurious modes. Such a filtering could be performed by considering the reduced space formed by the orthogonal complement of N_h in V_h: $N_h^\perp$. By construction, the operator induced by b, mapping $\Sigma_h \mapsto N_h^\perp$, is an isomorphism. Note that the discrete kernel Z_h remains the same even if V_h is replaced by $N_h^\perp$. These considerations lead to the definition of the reduced Brezzi inf-sup value $\tilde{\beta}_h$:

$$\tilde{\beta}_h = \inf_{0 \neq v \in N_h^\perp} \sup_{0 \neq \tau \in \Sigma_h} \frac{b(\tau, v)}{||\tau||_\Sigma ||v||_V}. \tag{2.11}$$

and the definition of reduced stable: A family of discretizations $\{\Sigma_h \times V_h\}_h$ is labelled *reduced stable* if α_h and $\tilde{\beta}_h$ defined by (2.6a) and (2.11) respectively are bounded from below, uniformly in h, by a positive constant.

The Brezzi coercivity and inf-sup constants, and the reduced Brezzi inf-sup constant can be related to a series of eigenvalue problems. First, it can be shown that the Brezzi coercivity constant $\alpha_h = |\lambda_{\min}|$, where $\lambda_{\min}$ is the smallest, in modulus, eigenvalue associated with a eigenfunction component $\sigma \neq 0$ of the following generalized eigenvalue problem:

$$a(\sigma, \tau) + b(\sigma, v) + b(\tau, u) = \lambda \langle \sigma, \tau \rangle_\Sigma \quad \forall\, (\tau, v) \in \Sigma_h \times V_h. \tag{2.12}$$

Second, similar derivations show that the Brezzi inf-sup constant $\beta_h = \sqrt{\lambda_{\min}}$, where $\lambda_{\min}$ is the smallest eigenvalue of the eigenvalue problem below:

$$\langle \sigma, \tau \rangle_\Sigma + b(\sigma, v) + b(\tau, u) = -\lambda \langle u, v \rangle_V \quad \forall\, (\tau, v) \in \Sigma_h \times V_h. \tag{2.13}$$

Finally, it can also be seen that the reduced Brezzi inf-sup constant $\tilde{\beta}_h$ will be the square-root of the smallest non-zero eigenvalue of (2.13). The details behind these derivations can be found in [**50**] or Paper IV and Malkus presented a thorough study of the eigenspectrum associated with the Stokes equations [**46**].

The eigenvalue problems (2.12) and (2.13) can be used to numerically test the stability of a given set of element spaces on a given family of mesh partitions for a given saddle point problem. This idea is explored for a class of finite element spaces on some regular triangulations for the mixed Laplacian in Paper IV. Such a study is greatly enhanced by a flexible and efficient computational finite element framework. One such will be described in connection with Paper III in Section 3.

2.3. **Weak formulations for linear elasticity.** It is now time to turn from the general saddle point framework to the equations of linear elasticity (1.11). The mixed stress-displacement-rotation formulation of linear elasticity, also known as the weak symmetry approach, is a foundation for Papers I and II and provides some underlying motivation for Papers III and IV. In this section, a brief survey of different (mixed) weak formulations and associated stable mixed finite element discretizations are presented. The material considered here can also be compiled from the references [**11, 16, 26**], with the exception of the last nearly symmetric extension. The reader is referred to those for more details and proofs of the claims made in the subsequent paragraphs.

Let Ω be an open, bounded domain in $\mathbb{R}^d$, $d = 1, 2, 3$ with Lipschitz boundary $\partial\Omega$. Further, assume that the boundary splits into two disjoint domains $\partial\Omega_0$ and

$\partial\Omega_1$. Recall the equations of linear elasticity given in (1.11), repeated here for easy reference

$$\sigma = C\varepsilon(u), \quad -\operatorname{div}\sigma = g. \tag{2.14}$$

The equations above are considered over the domain Ω, g is a body force acting on the domain and C is a symmetric fourth-order material tensor. The system of equations is completed by the boundary conditions

$$u|_{\partial\Omega_0} = u_0, \quad \sigma \cdot n|_{\partial\Omega_1} = s_0$$

where n denotes a unit normal. The conditions correspond to a confinement of the body and an applied stress respectively at the indicated parts of the boundary. For notational simplicity, assume that both boundary conditions are homogenous, that is $u_0 = 0$ and $s_0 = 0$. The quantities of physical interest are the displacement u, and, often more importantly, the stress $\sigma \cdot n$, or equivalently the stress tensor σ.

Denote by $H_0^1(\Omega;\mathbb{V})$ those H^1 vector fields on Ω that are zero on $\partial\Omega_0$ in the trace sense. Eliminating σ and taking variations over $H_0^1(\Omega;\mathbb{V})$, integrating by parts, and using the symmetry of C, gives a weak *pure displacement* formulation: Find $u \in H_0^1(\Omega;\mathbb{V})$ satisfying

$$\langle C\varepsilon(u), \varepsilon(v)\rangle = \langle g, v\rangle \quad \forall\, v \in H_0^1(\Omega;\mathbb{V}).$$

There exists a unique solution u under the assumption that $g \in H_0^1(\Omega;\mathbb{V})^*$, that C is sufficiently smooth and non-degenerate, and Ω is sufficiently regular [**24**]. Moreover, these equations can easily be solved by a standard finite element method utilizing H^1conforming finite element spaces. However, for instance when C is isotropic; i.e. as defined by (1.13) with Lamé parameters μ and λ, the bilinear form defining this equation will not be uniformly bounded in λ. In addition, if $\partial\Omega_0$ is very small, Korn's inequality can only provide an unsatisfactory estimate. The consequence of either of these situations is that the resulting discrete linear systems are ill-conditioned and loss of convergence may be encountered. As the case $\lambda \to \infty$ corresponds to the very relevant case of nearly incompressible materials, this is indeed a relevant concern.

A common remedy for the isotropic, nearly incompressible case can be found by seeking inspiration from the equations of incompressible flow, and in particular the Stokes equations. By introducing the auxiliary variable $p = \lambda \operatorname{div} u$, one obtains a *displacement-pressure* formulation: Find $u \in H_0^1(\Omega;\mathbb{V})$ and $p \in L_2(\Omega)$ satisfying

$$\begin{aligned} \langle 2\mu\varepsilon(u), \varepsilon(v)\rangle + \langle \operatorname{div} v, p\rangle &= \langle g, v\rangle \quad \forall\, v \in H_0^1(\Omega;\mathbb{V}), \\ \langle \operatorname{div} u, q\rangle - \langle \frac{1}{\lambda} p, q\rangle &= 0 \quad \forall\, q \in L^2(\Omega). \end{aligned} \tag{2.15}$$

The limiting case $\lambda = \infty$ is well-defined and well-posed as the resulting system of equations are indeed the Stokes equations (2.9), with an additional inconsequential factor 2μ. For finite λ, (2.15) take the form of an extended saddle point problem as discussed in the previous section. A stability estimate of the form (2.4) is ensured if $2\mu + d\lambda > 0$. In correspondence with the previous remarks on the discretizations of extended saddle point problems, stable mixed finite element discretizations of the Stokes equations are appropriate also for (2.15).

However, neither the pure displacement nor the displacement-pressure formulation are entirely satisfactory in the sense that no direct stress approximation is offered. Furthermore, (2.15) is restricted to isotropic materials. The interest in direct stress approximations motives the alternative *stress-displacement* formulation: Find $\sigma \in H_0(\operatorname{div}, \Omega;\mathbb{S})$ and $u \in L^2(\Omega;\mathbb{V})$ satisfying

$$\begin{aligned} \langle A\sigma, \tau\rangle + \langle \operatorname{div}\tau, u\rangle &= 0 \quad \forall\, \tau \in H_0(\operatorname{div}, \Omega;\mathbb{S}), \\ \langle \operatorname{div}\sigma, v\rangle &= -\langle g, v\rangle \quad \forall\, v \in L^2(\Omega;\mathbb{V}), \end{aligned} \tag{2.16}$$

where $A = C^{-1}$ and $H_0(\operatorname{div}, \Omega; \mathbb{S})$ denotes the space of symmetric tensor fields with square integrable row-wise divergence that have zero normal component on $\partial\Omega_1$. Recall from Section 1.1 that the linearized stress tensor must be symmetric if the residual stress vanishes. The derivation of this formulation relies on the invertibility of the elastic stress-strain relationship and the symmetry of the stress tensor. In the isotropic case, the action of A reduces to

$$A\sigma = \frac{1}{2\mu}\left(\sigma - \frac{\lambda}{2\mu + d\lambda}(\operatorname{tr}\sigma)I\right). \tag{2.17}$$

It can be shown that the Brezzi inf-sup condition (2.3b) applied to (2.16) is satisfied. Further, the homogenous, isotropic A of (2.17) is positive-definite if $\mu > 0$ and $2\mu + d\lambda > 0$. Since $\operatorname{div} H(\operatorname{div}, \Omega; \mathbb{S}) \subset L^2(\Omega)$, the Brezzi coercivity condition (2.3a), holds whenever A is positive definite on $L^2(\Omega; \mathbb{S})$. Although the positive-definiteness of A fails on $L^2(\Omega; \mathbb{S})$ as $\lambda \to \infty$, a uniform coercivity estimate can be provided by an additional assumption. Letting $\tau = I$ in (2.16), we see that for homogenous, finite λ and μ, $\int_\Omega \sigma$ must vanish. Under the additional requirement $\int_\Omega \sigma = 0$, a uniform coercivity estimate for A can be established. Since A is also uniformly continuous in λ, this gives the desired robustness in the limit $\lambda \to \infty$.

Alas, in the course of four decades of research, the construction of stable pairs of finite element spaces for the discretization of (2.16) has proven to be nontrivial. In the last decade, stable finite element spaces, associated with a single tessellation family, have been constructed in both two and three dimensions [**1, 3, 6**]. However, the number of degrees of freedom is daunting, especially in three dimensions. The lowest order stress tensor element of Arnold et al. has 156 degrees of freedom on a single tetrahedron [**3**]. In addition, these families are not affine interpolation equivalent, making efficient assembly algorithms more challenging.

These considerations motivate yet another strategy, the *weakly symmetric* or *stress-displacement-rotation* approach, originating from Fraijs de Veubeke [**31**]. Instead of requiring that the stress tensor is an element of a space of symmetric-valued tensor fields, the symmetry can be enforced through a weak symmetry constraint, admitting an additional Lagrange multiplier. More precisely, this formulation reads: Find $\sigma \in H_0(\operatorname{div}, \Omega; \mathbb{M})$, $u \in L^2(\Omega; \mathbb{V})$ and $\gamma \in L^2(\Omega; \mathbb{K})$ satisfying

$$\begin{aligned} \langle A\sigma, \tau\rangle + \langle \operatorname{div}\tau, u\rangle + \langle \tau, \gamma\rangle &= 0 \qquad \forall\, \tau \in H_0(\operatorname{div}, \Omega; \mathbb{M}), \\ \langle \operatorname{div}\sigma, v\rangle + \langle \sigma, \eta\rangle &= -\langle g, v\rangle \quad \forall\, v \in L^2(\Omega; \mathbb{V}), \eta \in L^2(\Omega; \mathbb{K}). \end{aligned} \tag{2.18}$$

Formally, the Lagrange multiplier γ corresponds to the skew component of the gradient of u:

$$\gamma = \frac{1}{2}\left(\nabla u - \nabla u^T\right). \tag{2.19}$$

On the continuous level, (2.16) and (2.18) are equivalent. Furthermore, the Brezzi conditions hold with the same considerations on A for the two formulations. However, on the discrete level, (2.18) offers a relaxation of the symmetry constraint. Consequently, lower order element spaces may be considered for the stress tensor.

Finally, recall (1.14). The constraint $\langle \sigma, \eta\rangle = 0$ for all $\eta \in L^2(\Omega; \mathbb{K})$ in (2.18) corresponds to the case where $\mathcal{S}(I) = 0$. If the residual stress does not vanish; that is, if $\mathcal{S}(I) \neq 0$, the linearized stress tensor is not symmetric. This case is relevant in Paper II. However, the symmetry constraint can be replaced by (1.14) and thus the weakly symmetric formulation can be generalized to the *nearly symmetric* case. Denoting $r = \mathcal{S}(I)$, this observation yields the following formulation: Find $\sigma \in H_0(\operatorname{div}, \Omega; \mathbb{M})$, $u \in L^2(\Omega; \mathbb{V})$ and $\gamma \in L^2(\Omega; \mathbb{K})$ satisfying

$$\begin{aligned} \langle A_r\sigma, \tau\rangle + \langle \operatorname{div}\tau, u\rangle + \langle k_r\tau, \gamma\rangle &= \langle A_r r, \tau\rangle \qquad \forall\, \tau \in H_0(\operatorname{div}, \Omega; \mathbb{M}), \\ \langle \operatorname{div}\sigma, v\rangle + \langle \sigma, \eta\rangle - \langle r\gamma, \eta\rangle &= -\langle g, v\rangle \quad \forall\, v \in L^2(\Omega; \mathbb{V}), \eta \in L^2(\Omega; \mathbb{K}), \end{aligned} \tag{2.20}$$

where A_r and k_r are now functions of r such that $k_r = 1$ and $A_r = A$ for $r = 0$. Thus, if $r = 0$, (2.20) reduces to (2.18). We see that if $r \neq 0$, (2.20) corresponds to an extended saddle point problem of the form (2.10). Thus, under certain assumptions on the parameter r, the stability of a discretization of (2.20) can be deduced from its properties for (2.18). Sufficient assumptions on r in the isotropic case are carefully studied in Paper II.

2.4. **Weakly symmetric finite element discretizations.** The construction of stable, conforming mixed finite element spaces for the strongly symmetric stress-displacement formulation (2.16) was by no means trivial [**6**]. For the weakly symmetric formulation (2.18), matters have been more pleasing. There exists many families of stable triples of element spaces $\Sigma_h \subset H(\mathrm{div}, \Omega; \mathbb{M}), V_h \subset L^2(\Omega; \mathbb{V})$ and $Q_h \subset L^2(\Omega; \mathbb{K})$ including, in chronological order, those of Amara and Thomas [**2**], Arnold, Brezzi and Douglas [**4**], Stenberg [**57**], Farhloul and Fortin [**27**] and extensions of the latter by Arnold, Falk and Winther [**5**].

Stable triples of element spaces $\Sigma_h \times V_h \times Q_h$ for (2.18) have a close relation to stable pairs of element spaces for the mixed Laplacian and the Stokes equations. The following theorem is due to Farhloul and Fortin [**27**], and the recent survey of Falk [**26**] elaborates upon the connection in the finite element exterior calculus setting.

Theorem 2.1 (Farhloul and Fortin, 1997). *A triple of conforming element spaces $\Sigma_h \times V_h \times Q_h$ satisfies the Brezzi conditions* (2.6) *for the weakly symmetric linear elasticity equations* (2.18) *if both*

(1) *$\Sigma_h \times V_h$ satisfy the Brezzi conditions for the mixed Laplacian* (2.8).
(2) *There exists a space W_h such that* $\operatorname{curl} W_h = \Sigma_h$ *and $W_h \times Q_h$ satisfy the Brezzi conditions for the Stokes equations* (2.9).

Recall that discretizations $\Sigma_h \times V_h$ of the mixed Laplacian (2.8) usually rely on the $H(\mathrm{div})$-conforming spaces of Raviart-Thomas ($\Sigma_{h,k} = \mathrm{RT}_k$) [**51**] or Brezzi, Douglas and Marini ($\Sigma_{h,k} = \mathrm{BDM}_k$) [**14**], or their extensions to three dimensions by Nédélec [**47, 48**]. Either of these elements have the convenient property that there exists integers k_0, k_1, such that for any admissible triangulation $\mathcal{T}_h$ of a contractible domain Ω

$$\operatorname{curl} \mathcal{P}^c_{k_1}(\mathcal{T}_h) = \Sigma_{h,k} \quad \operatorname{div} \Sigma_{h,k} = \mathcal{P}_{k_0}(\mathcal{T}_h). \tag{2.21}$$

Further, since stable conforming element pairs for the Stokes equations typically rely on the use of the continuous piecewise polynomials $\mathcal{P}^c_k$, it follows from Theorem 2.1 and (2.21) that triples of element spaces for the weakly symmetric formulation can be easily be constructed. The families of element spaces $\mathrm{BDM}_k \times \mathcal{P}_{k-1} \times \mathcal{P}_k$, described in Arnold et al. [**5**], can be viewed as the combination of the $\mathrm{BDM}_k \times \mathcal{P}_{k-1}$ elements for the mixed Laplacian and the $\mathcal{P}^c_{k+1} \times \mathcal{P}_{k-1}$ elements for the Stokes equations. Furthermore, Farhloul and Fortin [**27**], and later Falk [**26**], suggested the use of the $\mathrm{BDM}_k \times \mathcal{P}_{k-1} \times \mathcal{P}^c_k$ elements. The latter triple correspond to the former elements for the mixed Laplacian, but the generalized Taylor-Hood elements for Stokes. With this connection in mind, the numerical investigations of Paper IV relating to the mixed Laplacian can also be related to a gentle probing of the possible use of continuous stress approximations for the elasticity equations. Both of the families of elements described above are employed in Papers I and II.

3. When compiling mixed forms

This section aims at giving motivation and a framework for the compilation of mixed variational forms in general, and for Paper III in particular. The terminology of this section should be standard according to the computational aspects of

finite element methods, and is based on the monograph [**44**] and the recent survey paper [**45**].

The discretizations of the weakly symmetric stress-displacement-rotation formulation for linear elasticity, described in the previous section, constitute the computational basis for Papers I and II. Specifically, for $\Omega \subset \mathbb{R}^d$, the stress tensor is approximated by d rows of BDM_k elements. Moreover, in Paper I, the proposed strategy for dealing with generalized viscoelasticity models (1.15), involves the introduction of multiple stress tensors. It is evident that the proposed equations potentially generate large element matrices. Consequently, efficient assembly of such element matrices is required. On the other hand, the wide range of possible weak formulations corresponding to different models, and different temporal discretizations, calls for a flexible problem specification framework. Additionally, ease with regard to arbitrary mixed finite element combinations is desirable in order to compare different, possibly nontrivial spatial discretizations, such as for a comparison of different families of elasticity elements, and for the experiments of Paper IV. The combination of the desired flexibility and efficiency is attainable by a *compiled* approach to multilinear form representation and element tensor code generation.

Roughly speaking, the currently available finite element software can be classified as belonging to one of two categories: Those aimed at (only) solving a given physical boundary value problem, and those aimed at providing a framework for new weak formulations and discretizations. For the first category, the underlying discretization is usually not readily available to the user. However, for the second category, the transparency of the discretization is indeed essential. Clearly, the latter category is the more relevant for the development and investigation of new discretizations. However, even for such software projects, implementations of $H(\mathrm{div})$ conforming finite element spaces of arbitrary degree are not entirely common. Moreover, the above criteria with regard to flexibility, generality and efficiency are rarely satisfied. In order to give some credibility to this claim, consider the following examples. FreeFEM [**34**] provides form flexibility through a high-level language interface, but only the lowest-order Raviart-Thomas elements. FEMSTER [**18**] does provide arbitrary degree Nédélec elements, but does not provide automated form evaluation or element tensor assembly. NGSolve [**53**] provides arbitrary order $H(\mathrm{div})$ elements along with automated assembly, but only for a predefined set of bilinear forms by default.

The software project FEniCS, and in particular its components FIAT (FInite element Automated Tabulator), FFC (FEniCS Form Compiler) and DOLFIN (Dynamic Object-oriented Library for FINite element computation), is explicitly aimed at the automation of computational mathematical modeling, emphasizing *generality, efficiency* and *simplicity* [**35**]. Both FIAT and FFC are Python modules. DOLFIN is a c++ library with a python interface PyDOLFIN. Currently, the following functionality is provided by the components listed above. FIAT constructs and tabulates finite element bases of arbitrary order on a reference simplex T, including basis functions for the spaces $\mathcal{P}_k^c, \mathcal{P}_k$, RT_k, BDM_k for $k = 1, 2, \ldots$ in (one,) two and three dimensions [**40, 41**]. FFC provides a high-level interface language for the specification of (mixed) finite element spaces and multilinear forms [**42, 43**]. This form compiler makes extensive use of operator overloading, in order to enable language constructions close to the mathematical notation. Arbitrary mixed finite element spaces can be represented as combinations of the base spaces listed above. The multilinear forms are restricted to those that can be represented as integrals over the algebra of basis functions and their derivatives. After an initial analysis stage, FFC generates low-level c++ code for the evaluation of the element basis functions, the degrees of freedom and the element tensor(s). The generated

```
k = 2

S = FiniteElement("BDM", "triangle", k)
V = VectorElement("DG", "triangle", k-1)
Q = FiniteElement("DG", "triangle", k-1)

AFW = MixedElement([S, S, V, Q])
```

```
k = 2

S = FiniteElement("BDM", "triangle", k)
V = VectorElement("DG", "triangle", k-1)
Q = FiniteElement("CG", "triangle", k)

FFF = MixedElement([S, S, V, Q])
```

TABLE 1. FFC code for the definition of stress-displacement-rotation finite element spaces on triangular meshes. Right: The $\mathrm{BDM}_k^2 \times \mathcal{P}_{k-1}^2 \times \mathcal{P}_{k-1}$ elements. Left: The $\mathrm{BDM}_k^2 \times \mathcal{P}_{k-1}^2 \times \mathcal{P}_k^c$ elements. Note that the only difference is in the definition of the space for the rotation Q both in the code and in the element definitions.

code can be used directly, or for instance through the problem-solving environment DOLFIN, which provides automated element tensor assembly on simplicial meshes. Also, in combination with PyDOLFIN, FFC provides *just-in-time* compilation. A more detailed review of the FEniCS project and a comparison with some other finite element frameworks is presented in [**45**].

The discretization language provided by FFC indeed enables simple and flexible implementations of the discretizations for linear elasticity considered in the previous section, in part due to the considerations of Paper III. In particular, note that the class of multilinear forms covered is quite extensive, and that any of the weak formulations considered in Section 2.3 are easily represented. Further, the generated code is based on the following observation: Under certain assumptions on the finite element spaces and the weak formulations, the element tensor may be represented as a tensor contraction of a *reference tensor* A and a *geometry tensor* G_T. The reference tensor A only relies on the differential formulation, and the reference finite element space, and may thus be precomputed. This observation was formalized by Logg and Kirby for H^1 conforming finite element spaces in connection with the afore class of multilinear forms [**43**]. It was also demonstrated in that work that such as tensor representation could provide a substantial speed-up when compared to standard quadrature-based strategies, thus offering the desired efficiency. Paper III can be viewed as a direct extension of the afore H^1 conforming case to the $H(\mathrm{div})$ and $H(\mathrm{curl})$-conforming cases. The implementation of the ideas of that paper can thus be viewed as greatly enhancing the computations performed in Papers I and II.

For illustration purposes, Table 1 presents FFC code defining the weakly symmetric element spaces of Arnold, Falk and Winter (`AFW`) and Farhloul and Fortin (and Falk) (`FFF`). One can observe that the same number (three) of key-strokes are required to switch between the two mathematical element definitions, and to switch between the associated element implementations, thus providing some justification of the claimed flexibility and simplicity of implementation. More examples are presented in Paper III.

4. Overview of papers

The previous three sections aimed at providing motivation and background material for the four papers constituting the bulk of this thesis. Along the way, some natural relations between the papers have been deliberately pointed out. In this section, some interpaper continuity notions are more concisely summarized

and a brief overview of each paper is presented. The first three papers have been submitted for publication in international journals.

Both Papers I and II investigate extensions, originating from more involved physical phenomena, of weak formulations for the linear elasticity equations. Although the physical settings of the two papers are separate, each of the two scenarios take both solid and fluid aspects into account within the linear regime. As has been indicated, both extensions apply to simulation of deformation and stress in biological tissue and biomedical devices. In Paper I, dealing with viscoelasticity in the quasi-static equilibrium, the extension involves temporal effects. The static equilibrium of gel mixtures is considered in Paper II. Here, the natural extension involves the nearly symmetric case as a result of residual mixing stress.

The first two papers also rely on the same mixed finite element method; namely, the discrete stress-displacement-rotation formulation. Paper III presents an efficient and user-friendly framework for, in particular, the implementation of such mixed finite element discretizations. The strategy presented in the third paper can thus be viewed as an underlying base for the experiments performed in the first two papers.

The fourth paper, investigating the reduced stability of H^1 conforming mixed finite element discretizations of the Laplacian, can be viewed separately from the first three. However, taking the close connection demonstrated by Theorem 2.1 into account, this study can also be considered as a gentle probing into the reduced stability of H^1-conforming mixed finite element discretization of the weakly symmetric formulation.

Paper I: Mixed finite element methods for linear viscoelasticity using weak symmetry. This paper aims at providing a unified discretization framework, approximating the stress directly, for generalized viscoelastic models in two and three dimensions. To the author's knowledge, this is the first work seeking higher order stress approximations on simplicial meshes for the generalized models. The work of Bécache et al. is based on a similar notion, but only stress discretizations on regular grids consisting of squares or cubes, of the generalized Zener model is considered in [**9**].

The starting-point for the study is the strain-stress formulation of generalized viscoelasticity models in differential form (1.15). We derive a weak formulation of these equations, approximating the separate stress contributions, the displacement and the rotation. However, in the analysis, main emphasis is placed on the two fundamental models: the Maxwell and the Kelvin-Voigt model. The purpose of this is two-fold. First, we are able to present the strategy with greater clarity. Second, the two fundamental models can be perceived as two limiting cases, and an analysis of these gives the key features also for the generalized models.

In the Maxwell case, it can easily be seen that the formulation for linear elasticity extends trivially. For the Kelvin-Voigt equations, the introduction of separate elastic and viscous stresses enables the stress-displacement-rotation formulation. In order to demonstrate that the continuous solutions are stable with respect to initial data and body forces, we present stability estimates in the style of (2.4). We continue by considering a spatial mixed finite element discretization of the two models. In particular, for the Kelvin-Voigt equations, both the elastic and the viscous stress tensor are approximated. The analysis mimics the continuous case, and a priori error estimates for the spatial discretization are established. In other words, the approximation errors of the discretization are bounded by the interpolation errors in a suitable temporal sense. The resulting semi-discrete systems of equations are differential-algebraic equations, and can thus be discretized by standard methods for such equations. Finally, due to the relative scarcity of numerical

experiments for stress-displacement-rotation elements in the literature, we report on the convergence rates for both elastic and viscoelastic models and higher order elements.

The analysis and experiments indeed show that the stress approximation methods developed for linear elasticity are also suitable and natural for linear viscoelasticity. No instabilities with regard to material parameters are observed, even in the nearly incompressible or the stiff cases, in accordance with the theoretical predictions.

Paper II: Mixed finite element methods for gels with biomedical applications. In the medical procedure of high tibial osteotomy, a piece is removed from the human shinbone and replaced by a polymer wedge in order to realign the knee. The polymer may absorb moisture from of the human body and swell, in accordance with the discussion of gels in Section 1.3. The stress induced by the swelling of the device is of critical interest. A simulation of such stresses can be viewed as a partial aim of this paper. This view-point also strongly motivates numerical approaches providing higher order stress approximation. Although the gel equilibrium problem can be seen as closely related to an elasticity problem, the use of mixed finite element methods approximating the stress directly have not been previously observed in the literature.

In this paper, a set of equilibrium constitutive equations of a gel are derived, accounting for both swelling and phase transitions. The modeling and derivation is based on the theory of mixtures in combination with nonlinear elasticity. This approach gives the equilibrium equations as the critical equations of an energy minimization problem. Under the polyconvexity assumption, existence of minimizers are established.

The second half of the paper is concerned with small strain, and small environmental perturbation, linearizations of these constitutive equations. The linearized equations take the form of isotropic linear elasticity equations, where the Lamé parameters additionally depend on the polymer volume fraction of the reference domain and polymer-solvent mixing parameters. We consider two standard discretization strategies for these equations, namely the displacement-pressure and the stress-displacement-rotation formulations. Recall that under the assumption of sufficient smoothness of the data and the domain, both discretizations, using the Taylor-Hood elements and the FFF elements, respectively, give second order stress tensor approximations for the lowest order case.

A key point in this study is the observation that the reference polymer fraction may induce a residual stress due to polymer-solvent mixing, and hence a non-symmetric linearized stress tensor, cf. (1.14). In order to apply direct stress approximations in this case, the weakly symmetric stress-displacement-rotation formulation is extended to the nearly symmetric case. The stability of this formulation as a function of the residual stress is studied analytically and numerically. Finally, numerical experiments relating to the above bone implant produce are presented.

In conclusion, the extension of the weakly symmetric formulation to the nearly symmetric case provides a successful method for studying residual stresses induced by polymer-solvent mixing. Although one might argue that direct stress approximations is unnecessary, since the displacement-pressure approach gives the same stress accuracy, we claim otherwise. For applications with internal boundaries, only $H(\mathrm{div})$ stress tensor approximations enable direct measurements of the stresses at these boundaries, since the stresses; that is, the normal components of the stress tensor, are only well-defined over edges/faces in this case.

Paper III: Efficient assembly of H(div) and H(curl) conforming finite elements. As has been emphasized in the previous sections, the space $H(\mathrm{div})$ plays an important role for mixed formulations of elliptic equations in general, and the linear elasticity equations in particular. In order to lower the threshold for the use of such sophistical numerical methods, user-friendly, flexible and efficient implementations of $H(\mathrm{div})$ conforming finite element spaces are required. However, such implementations are relatively scarce. For those projects actually providing $H(\mathrm{div})$ conforming elements in some sense, the implementation often does not provide the desired ease of use. The implementation of the ideas in this paper, in connection with the FEniCS project, claim to satisfy the previous criteria well.

In this paper, we consider some of the aspects required for automated form compilation of $H(\mathrm{div})$ and $H(\mathrm{curl})$ variational formulations and finite element spaces. In particular, the role of the contravariant and covariant Piola mappings is emphasized. A representation theorem for variational forms with Piola mapped basis functions ensures that the class of multilinear forms over H^1 conforming finite element spaces can be extended to $H(\mathrm{div})$ and $H(\mathrm{curl})$ discretizations. We also take advantage of a common numbering scheme for mesh entities, in order to avoid sign issues with regard to the orientation of edges or faces. A set of numerical experiments dealing with standard $H(\mathrm{div})$ and $H(\mathrm{curl})$ convergence and eigenvalue problems in two and three dimensions are presented in order to verify and demonstrate the implementation.

The resulting FFC implementation renders use of $H(\mathrm{div})$ and $H(\mathrm{curl})$ variational formulations no more complicated than standard H^1 formulations. Also note that, the element spaces provided can also easily be used in combination with discontinuous Galerkin methods. The implementation is designed to lower the threshold for the use of, and experiments with, this type of variational formulations, and has been thoroughly used in connection with Papers I and II.

Paper IV: Stability of Lagrange finite elements for the mixed Laplacian.

> *One cannot survive without the Laplacian. In mathematics, that is. In life, that depends on your definition of the Laplacian.*
>
> J. L. Vazquez, 2007

The reduced stability of mixed finite element discretizations for the Stokes equations have been extensively studied [**10, 50**]. However, the corresponding notion for the mixed Laplacian has been less so. This note aims at investigating the reduced stability of H^1-conforming finite element spaces for the latter equations in two dimensions.

The first half of this note summarizes eigenvalue problems associated with the Brezzi (and Babuska) constants for abstract saddle point problems in general. In the case of the mixed Laplacian, a connection between the Brezzi infsup constant and the smallest eigenvalue of the Laplacian is pointed out on the continuous and discrete level.

The second part is concerned with the stability or reduced stability of a class of mixed finite element pairs for the mixed Laplacian. For pairs where weakly divergence free vector fields are indeed divergence free, and thus where the Brezzi coercivity condition holds automatically, the reduced stability of discretizations for the Stokes equations imply reduced stability for the mixed Laplacian. However, the converse is not true. Hence, there may be a wider class of tessellations for which the discretization is reduced stable for the mixed Laplacian. We present numerical evidence of this for the $\mathcal{P}_k^c(\mathbb{V}) \times \mathcal{P}_{k-1}$ family in two dimensions. For $k \geq 4$, this family is known to be stable for both Stokes and the mixed Laplacian on a fairly

generic class of triangulations [**54**]. The cases $k = 1, 2, 3$ are considered in this note. For all triangulation families tested, the family was observed to be at least reduced stable for $k = 2, 3$. For $k = 1$, the situation is more diverse. Stability appears to hold for some meshes, but not for others.

In its current form, this note presents mainly numerical evidence for the above claims. However, the observed reduced stability of $\mathcal{P}_k^c(\mathbb{V}) \times \mathcal{P}_{k-1}(\mathbb{V})$ for $k = 2, 3$ indicates that a theoretical study may be interesting. This work is therefore planned to be supplemented by theoretical investigations.

5. Concluding remarks

In conclusion, mixed finite element methods offering direct stress approximations at a reasonable cost have been investigated for and applied to physical phenomena such as viscoelasticity and gels within the linear regime. The implementation of such methods has been greatly enhanced by an efficient framework for operating with variational forms over $H(\mathrm{div})$. The previous sections aimed at providing background material and a greater perspective at the research constituting this thesis. This section concludes by some remarks on limitations and possible extensions of the work presented.

First, it should be remarked that the analysis of Paper I is carried out for the two basic models and not the generalized model. However, we have seen that the generalized viscoelastic equations in the strain-stress form afford the same weak formulations. Moreover, numerical experiments for the standard linear solid model have been performed. Accordingly, it is the author's belief that the analysis can be extended to the generalized case, using the same techniques and thus not offering extensive new insight. More interestingly perhaps is the temporal discretization of differential-algebraic equations resulting from time-dependent saddle point problems. In Paper I, we rely on the stability of implicit time-stepping schemes for differential-algebraic equations. However, the effect of the spatial discretization, in particular the mesh size, on the stability of these temporal schemes is not entirely clear. Note that no instabilities or loss of convergence is observed in the numerical experiments. Furthermore, it can be shown that for instance a backward Euler scheme in combination with the afore described spatial discretization is stable and convergent for both the Maxwell and Kelvin-Voigt equations. However, this matter may be worth investigating for more general schemes.

Next, with reference to the discussion on the modeling of gels and for instance the dynamical models of gels presented in [**17**], it must be noted that the equilibrium equations of gels only consider a small piece of the dynamical behavior. We have seen that the stress approximation methods are highly germane for the biomedical equilibrium problems aimed at in Paper II. However, for the dynamical systems involving solvent flow and swelling interfaces, displacement or rather velocity-based methods may be more appropriate.

The numerical strategy and simulations in Paper II consider the linear regime. For phase transition of gels, where the volume change can be 10–100 times, the linearity assumption is hopeful at best. Therefore, numerical methods for the nonlinear gel boundary value problem should also be investigated. However, even by close scrutiny of the general form of the linearized equations in Paper II, it is not evident how to formulate either a deformation-pressure or a stress-displacement formulation. The linear stress-displacement formulation rely on the invertibility of the stress-strain relationship, but this inversion is less feasible for the general form. The linearized equations may not be a good starting-point though. Braess and Ming presented a displacement-pressure formulation for a neo-Hookean model by starting at a discretization of the elastic energy [**12**]. Moreover, Steinmann et

al. successfully derived a stress-displacement-rotation method for large deformation under the assumption of an energy principle for the stress [**56**]. This seems to be a more fruitful approach, but has not been within the scope of this thesis.

Finite deformation within viscoelastic behavior is also an interesting topic. Viscoelastic flow problems constitute an active research field, but is focused at other aspects than purely the finite deformation regime. On the other hand, large deformations in viscoelastic solids have been less studied numerically. Note that for nonlinear viscoelastic materials, the established models typically take integral form for which the integral kernel is not separable in space and time. This makes matters much more complicated.

Some more readily feasible and planned extensions are sketched in Paper III. At the time of writing, FFC handles finite element spaces that can be generated by either the affine or the Piola mappings of basis functions on a reference element. More fun element spaces, such as the H^2 conforming Argyris element, the Arnold-Winther strongly symmetric elasticity element and the Mardal-Tai-Winther Darcy-Stokes elements, are not available. However, one may possibly consider mappings that depend more extensively on the geometry. The challenge becomes the automated generation of efficient code for such.

Lastly, Paper IV is presently at large a set of numerical observations with interesting implications. If the numerical observations hold true in a more general sense, they could indeed be of practical importance. A theoretical investigation may therefore also be in order.

Bibliography

[1] S. Adams and B. Cockburn. A mixed finite element method for elasticity in three dimensions. *J. Sci. Comput.*, 25(3):515–521, 2005.

[2] M. Amara and J. M. Thomas. Equilibrium finite elements for the linear elastic problem. *Numer. Math.*, 33:367–383, 1979.

[3] D. N. Arnold, G. Awanou, and R. Winther. Finite elements for symmetric tensors in three dimensions. *Math. Computation*, 77:1229–1251, 2008.

[4] D. N. Arnold, F. Brezzi, and J. Douglas. PEERS: A new mixed finite element for plane elasticity. *Japan J. Appl. Math.*, 1:347–367, 1984.

[5] D. N. Arnold, R. Falk, and R. Winther. Mixed finite element methods for linear elasticity with weakly imposed symmetry. *Math. Computation*, 76:1699–1723, 2007.

[6] D. N. Arnold and R. Winther. Mixed finite elements for elasticity. *Numer. Math.*, 92:401–419, 2002.

[7] F. P. T. Baaijens. Mixed finite element methods for viscoelastic flow analysis: a review. *J. Non-Newtonian Fluid Mech.*, 79:361–385, 1998.

[8] J. M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 63:337–403, 1977.

[9] E. Becache, A. Ezziani, and P. Joly. A mixed finite element appproach for viscoelastic wave propagation. *Computational Geosciences*, 8:255–299, 2004.

[10] D. Boffi, F. Brezzi, and L. Gastaldi. Finite elements for the Stokes problem. In *Mixed Finite Elements, Compatibility Conditions and Applications*. Springer, 2007.

[11] D. Braess. *Finite elements. Theory, fast solvers, and applications in solid mechanics*. Cambridge University Press, second edition, 2001.

[12] D. Braess and P. Ming. A finite element method for nearly incompressible elasticity problems. *Math. Computation*, 74:25–52, 2004.

[13] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *R.A.I.R.O. Anal. Numer.*, 2:129–151, 1974.

[14] F. Brezzi, J. Douglas, and L. D. Marini. Two families of mixed elements for second order elliptic problems. *Numer. Math.*, 47:217–235, 1985.

[15] F. Brezzi and R. Falk. Stability of higher-order Hood–Taylor methods. *SIAM J. Numer. Anal.*, 28:581–590, 1991.

[16] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Springer, 1991.

[17] M. C. Calderer, B. Chabaud, S. Lyu, and H. Zhang. Modeling approaches to the dynamics of hydrogel swelling. *Journal of Computational and Theoretical Nanoscience*. To appear.

[18] P. Castillo, R. Rieben, and D. White. FEMSTER: an object-oriented class library of discrete differential forms. *ACM Trans. Math. Software*, 35:425–457, 2005.

[19] P. Ciarlet. *Mathematical Elasticity, Volume I: Three-Dimensional Elasticity*. North-Holland, Amsterdam, 1987.

[20] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology*, volume 1. Springer, 1990.

[21] J. Dolbow, E. Fried, and H. Ji. Chemically induced swelling of hydrogels. *Journal of Mechanics and Physics of Solids*, 52:51–84, 2004.

[22] D. Doraiswamy. The origins of rheology: A short historical excursion. *The Society of Rheology, Bulletin*, 71(1), January 2002.

[23] A. D. Drozdov. *Mechanics of Viscoelastic Solids*. Wiley, 1998.

[24] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 2002.
[25] M. Fabrizio and A. Morro. *Mathematical Problems in Linear Viscoelasticity*. SIAM, 1992.
[26] R. S. Falk. Finite element methods for linear elasticity. In *Mixed Finite Elements, Compatibility Conditions and Applications*. Springer, 2008.
[27] M. Farhloul and M. Fortin. Dual hybrid methods for the elasticity and the Stokes problem: a unified approach. *Numer. Math.*, 76:417–440, 1997.
[28] J. D. Ferry. *Viscoelastic Properties of Polymers*. John Wiley and Sons, 2002.
[29] P. Flory. *Principles of Polymer Chemistry*. Cornell University Press, 1953.
[30] P. J. Flory and J. Rehner. Statistical mechanics of cross-linked polymer networks I. rubberlike elasticity. *Journal of Chemical Physics*, 11, 1943.
[31] B. M. Fraijs de Veubeke. Stress function approach. In *Proc. of the World Congress on Finite Element Methods in Structural Mechanics*, volume 1, pages J.1–J.51, 1975.
[32] M. E. Gurtin. *An introduction to Continuum Mechanics*. Academic Press, 1981.
[33] M. E. Gurtin and E. Sternberg. On the linear theory of viscoelasticity. *Arch. Rational Mech. Anal*, 11:291–356, 1962.
[34] F. Hecht, O. Pironneau, A. L. Hyaric, and K. Ohtsuka. *FreeFEM++ manual*.
[35] J. Hoffman, J. Jansson, C. Johnson, M. G. Knepley, R. C. Kirby, A. Logg, L. R. Scott, and G. N. Wells. *FEniCS*. `http://www.fenics.org/`.
[36] G. A. Holzapfel. *Nonlinear Solid Mechanics. A continuum approach for Engineering*. Wiley, first edition, 2000.
[37] W. Hong, X. H. Zhao, J. Zhou, and Z. Suo. A theory of coupled diffusion and large deformation in polymeric gels. *Journal of the Mechanics and Physics of Solids*, 56:1779–1793, 2008.
[38] M. L. Huggins. Solutions of long chain compounds. *Journal of Chemical Physics*, 9, 1941.
[39] A. Idesman, R. Niekamp, and E. Stein. Finite elements in space and time for generalized viscoelastic Maxwell model. *Computational Mechanics*, 27:49–60, 2001.
[40] R. C. Kirby. Algorithm 839: FIAT, a new paradigm for computing finite element basis functions. *ACM Transactions on Mathematical Software*, 30:502–516, 2004.
[41] R. C. Kirby. Optimizing FIAT with Level 3 BLAS. *ACM Transactions on Mathematical Software*, 32:223–235, 2006.
[42] R. C. Kirby and A. Logg. A compiler for variational forms. *ACM Transactions on Mathematical Software*, 32:417–444, 2006.
[43] R. C. Kirby and A. Logg. Efficient compilation of a class of variational forms. *ACM Transactions on Mathematical Software*, 33, 2007.
[44] H. P. Langtangen. *Computational Partial Differential Equations. Numerical Methods and Diffpack Programming*. Springer, second edition, 2003.
[45] A. Logg. Automating the finite element method. *Arch. Comput. Methods Eng*, 14:93–138, 2007.
[46] D. S. Malkus. Eigenproblems associated with the discrete LBB condition for incompressible finite elements. *Int. J. Engng. Sci.*, 19:1299–1310, 1981.
[47] J.-C. Nédélec. Mixed finite elements in $\mathbf{R}^3$. *Numer. Math.*, 35:315–341, 1980.
[48] J. C. Nédélec. A new family of mixed finite elements in $\mathbf{R}^3$. *Numer. Math.*, 50:57–81, 1986.
[49] Y. Osada and K. Kajiwari, editors. *Gels Handbook*. Academic Press, 2001.
[50] J. Qin. *On the Convergence of Some Simple Finite Elements for Incompressible Flows*. PhD thesis, Penn State, 1994.
[51] P. A. Raviart and J. M. Thomas. A mixed finite element method for second order elliptic problems. *Mathematical Aspects of the Finite Element Meth. Lecture Notes in Math., Springer*, 606:292–315, 1977.
[52] B. Rivière, S. Shaw, and J. R. Whiteman. Discontinuous Galerkin finite element methods for dynamic linear solid viscoelasticity problems. *Numerical Methods for Partial Differential Equations*, 23:1149–1166, 2007.
[53] J. Schöberl. `NGS`olve. URL: `http://www.hpfem.jku.at/ngsolve/index.html/`.
[54] L. R. Scott and Vogelius. Norm estimates for a maximal right inverse of the divergence operator. *R.A.I.R.O. Anal. Numer.*, 19:111–143, 1985.
[55] S. Shaw. *Finite element and discrete time methods for continuum problems with memory and applications to viscoelasticity*. PhD thesis, Department of Mathematics and Statistics, Brunel University, 1993.
[56] P. Steinmann, U. Brink, and E. Stein. On a mixed finite element formulation involving large rotations for geometrically nonlinear elasticity. *Rev. Européenne des éléments finis*, 4:577–596, 1995.
[57] R. Stenberg. A family of mixed finite elements for the elasticity problem. *Numer. Math.*, 53:513–538, 1988.
[58] T. Tanaka. Collapse of gels and the critical endpoint. *Phys. Rev. Lett.*, 40:820–823, 1978.

[59] C. Taylor and P. Hood. A numerical solution of the Navier–Stokes equations using the finite element technique. *Internat. J. Comput. and Fluids*, 1:73–100, 1973.

[60] N. Tschoegl. *The Phenomenological theory of linear viscoelastic behaviour*. Springer, first edition, 1989.

Paper I

MIXED FINITE ELEMENT METHODS FOR LINEAR VISCOELASTICITY USING WEAK SYMMETRY

MARIE E. ROGNES AND RAGNAR WINTHER

Abstract. Small deformations of a viscoelastic body are considered through the linear Maxwell and Kelvin-Voigt models in the quasi-static equilibrium. A robust mixed finite element method, enforcing the symmetry of the stress tensor weakly, is proposed for these equations on simplicial tessellations in two and three dimensions. A priori error estimates are derived and numerical experiments presented. The approach can be applied to general models for linear viscoelasticity and thus offers a unified framework.

1. Introduction

Viscoelastic materials are characterized by their ability to display both viscous and elastic behaviour. Most real-life solids demonstrate some viscoelastic properties, and these effects may be particularly important when considering synthetic polymers or biological materials such as muscles or soft tissue. In this paper, we revisit the fundamental models for small deformation viscoelasticity in the quasi-static equilibrium, with the purpose of deriving a robust and flexible mixed finite element method. In particular, we consider the analysis of, and numerics for, two basic models and provide arguments for how the setup extends to generalized viscoelastic models.

The established theory for linear elasticity provides a sound starting point for the study of linear viscoelasticity. In this work, we shall rely on known results on the stability and robustness of mixed finite element methods for linear elasticity. The classical approach to linear isotropic elasticity consists of solving Navier's equations for the displacement u over a domain Ω:

$$\operatorname{div}(2\mu\varepsilon(u) + \lambda \operatorname{div} u I) = g \quad \text{in } \Omega, \tag{1.1}$$

where g is some prescribed body force. The Lamé coefficients μ and λ relate to the stiffness and compressibility of the material respectively. These equations can be solved numerically by for instance a standard finite element method giving optimal order error estimates for u using continuous piecewise vector polynomials.

However, the quantity of primary physical interest is often the stress and pure displacement methods will yield stress approximations of lower order accuracy. Furthermore, it is well-known that standard discretizations based on the formulation of (1.1) is not robust with regard to the material parameters. The effect is that the method performs poorly in the incompressible and nearly incompressible case, i.e. as $\lambda \to \infty$. Alternative approaches are therefore eligible. A mixed formulation involving the stress tensor in addition to the displacement address the afore issues. For this approach, the main obstacle has been the construction of stable pairs of finite element spaces. Non-composite families of such elements have been established only in the recent years, cf. [**1, 4, 8**] for simplicial tesselations in two and three dimensions. However, the complexity of these may seem prohibitive: The lowest order element space on a simplex has 21 degrees of freedom in two dimensions and 156 in three dimensions.

Partially as a remedy when suitable finite element spaces were not known, and partially in order to avoid the cost of such elements, a further extension has been considered. Instead of enforcing the symmetry of the stress through the element space directly, it can be enforced weakly by an additional equation and an associated Lagrange multiplier. This idea dates back to the 1970's, originally suggested by Fraijs de Veubeke [**24**], and various stable element spaces have been presented in later works, including [**3, 5, 6, 20, 21, 39**]. This approach enables the use of simpler finite element spaces for the stresses. Hence, since the multiplier associated with the symmetry constraint can be approximated in a relatively small element space, the total complexity can be reduced. We shall pursue this approach in this exposition.

For linear viscoelasticity, where the stress-strain relationship may be non-local, stress-displacement methods are especially advantageous. For instance, for the linear viscoelastic Maxwell model, the quasi-static equations can take the form:

$$\begin{aligned} A_1\dot{\sigma}(t) + A_0\sigma(t) &= \varepsilon(\dot{u}(t)) \\ \operatorname{div}\sigma(t) &= g(t), \end{aligned}$$

where the superimposed dot gives the time-derivative and A_0, A_1 are fourth-order material tensors. These equations are suited for a mixed stress-displacement method. In contrast, an elimination of the stress relies on an inversion of the stress-strain relation. The resulting formulation would involve an integro-differential equation for the displacements. The equivalence of the differential and the integral models was discussed by Gurtin and Sternberg in [**25**].

Most of the numerical work on linear viscoelasticity in the late 1980's and 1990's focused on such *hereditary integral* formulations, reflecting the inherent interpretation of viscoelastic materials as *materials with memory*. These terms refer to the property that the stress does not depend on the strain, or its rate of change, pointwise in time, but rather on the history of the strain evolution. For the linear theory, this idea, along with a Boltzmann superposition principle, gives integro-differential models in the form of Volterra integrals. The stress can then be expressed as an integral operator of the strain, typically of the form

$$\sigma(t) = C(t)\,\varepsilon(u(0)) + \int_0^t C(t-s)\,\varepsilon(\dot{u}(s))\,\mathrm{d}s$$

where C is a time-dependent fourth order material tensor. For the afore Maxwell model in one dimension, C takes the form $C(t) = A_1^{-1}e^{-A_1 t/A_0}$. The stress can then again be eliminated to yield a pure displacement integral formulation. Numerical methods for such problems thus involve approximations in time and space of u.

We shall not pursue these formulations further in this paper, and therefore simply remark that it has been extensively studied, including from a numerical

point of view. A detailed mathematical review for the integral formulations can be found in the monograph [**19**]. Shaw et al. presented a series of papers, [**35, 36, 37, 38**], including a priori and a posteriori error estimates, using continuous and discontinuous Galerkin finite elements for the spatial and time discretization respectively. Discontinuous Galerkin methods for the spatial discretization have been studied by Rivière et al. [**33**]. The integral kernel can be extended to fractional order time derivatives and treated by similar techniques [**2**].

The differential form of the constitutive equations has regained some of its popularity over the last decade. These formulations typically require the introduction of *internal variables* corresponding to internal state variables, such as elastic or viscous contributions to the stress or strain. The generalized Maxwell models have been a common starting point for most of the studies. In [**34**], Rivière et al. follow Johnson and Tessler [**29**] with regard to introducing internal stresses in $L^2(\Omega;\mathbb{R}^{d\times d})$ while seeking the displacement in $H^1(\Omega;\mathbb{R}^d)$, $d=2,3$. Idesman et al. [**28**] consider a converse approach introducing internal variables for the strains, in contrast to the stresses.

To our knowledge, there is only a handful of papers concerned with the analysis of mixed finite element methods for this type of formulations. Le Tallec and Ravachol considered a mixed finite element method inspired by the Stokes equations for the Maxwell model and its non-linear extensions including extensions for viscoelastic flow [**30**]. The more recent paper [**10**], thoroughly treats the dynamic (generalized) Zener model by approximating the symmetric internal stresses in addition to the displacements. However, the element spaces of the latter require regular cubical partitions and thus lay restrictions on the computational domain. No study for general simplicial partitions in two and three dimensions, flexibly treating both basic and generalized models in their differential form, seems to have been undertaken.

1.1. Main results. The main aim of this paper is to propose a robust and accurate mixed finite element method for generalized linear viscoelasticity models on general domains in two and three dimensions.

To this end, we revisit the basic Kelvin-Voigt and Maxwell models in their differential form. We consider a mixed finite element method for the spatial discretization inspired by the family of elements introduced by Arnold et al. [**6**]. For each polynomial degree k, these consist of piecewise discontinuous polynomials of order $k-1$ for the displacements and the auxiliary Lagrange multiplier, and the BDM_k [**14, 31**] elements, that is, polynomials of order k with inter-element normal continuity, for the stresses. We prove stability of the continuous solutions and their spatially discrete counterparts with regard to data along with deriving error estimates of the order $\mathcal{O}(h^k)$ for sufficiently smooth solutions. Furthermore, we indicate how the discretization and analysis can be extended to generalized linear viscoelasticity models of the form:

$$A_V^j\sigma_j + A_E^j\dot{\sigma}_j = \varepsilon(\dot{u}) \quad j=0,\dots,n-1,$$
$$\operatorname{div}\sum_{j=0}^{n-1}\sigma_j = g.$$

In all, this intends to show how mixed finite elements for linear elasticity are available and suitable also for linear viscoelasticity.

1.2. Outline. The organization of this paper is as follows: We provide a brief derivation of, along with equations for, the viscoelastic models of interest in Section 2. In Section 3, we carefully derive weak formulations of the models, explain the

relation to the linear elasticity equations and provide energy estimates demonstrating stability in time. Section 4 deals with the spatial discretization and a priori error estimates for this discretization, while the fully discrete system is considered in Section 5. For the time-discretization, we mainly refer to previous results on implicit differentiation schemes for differential-algebraic equations. Finally, we provide numerical examples in Section 6 before we conclude in Section 7.

2. Viscoelastic models

In this section, we shall describe a class of models representing viscoelastic behaviour. The material presented here is classical in many senses. However, this review is targeted at providing motivation for the choice of viewpoint in Section 3. In particular, we shall focus on the derivation of stress-strain relations and point out less classical, alternative formulations. The reader can find a more thorough discussion of viscoelastic behaviour and modelling among the references [**17, 19, 23, 32, 40**].

The classical continuum modelling of solids is based on balance of linear momentum in combination with an appropriate constitutive equation. For the quasi-static state, the former equation takes the form

$$\operatorname{div} \sigma = g, \tag{2.1}$$

where σ is the *stress tensor* and g a body force. The material characteristics of the solid must be reflected in a constitutive model relating the stress tensor to the strain (and possibly rates of strain) of the body. In the small-deformation framework, the linearized strain tensor ε is defined in terms of the displacement u by the relation: $2\varepsilon(u) = \nabla u + \nabla u^T$. We shall focus on constitutive equations for viscoelastic materials in the subsequent paragraphs.

Various viscoelastic behaviour patterns may be illustrated in one dimension by combinations of springs and dashpots, representing elastic and viscous factors respectively. The corresponding constitutive laws may be derived from the compatibility and equilibrium conditions of the physical systems represented. Generalizations to two or three dimensions can be performed by considering deviatoric and volumetric contributions separately.

The simplest manner in which to schematically construct a viscoelastic model is to combine one elastic and one viscous component either in series or in parallel. The resulting models are known as the Maxwell and the Kelvin-Voigt model or element respectively. Note that these are clearly simplicial and only capable of describing viscoelastic behaviour partially. However, these basic models provide building-blocks for the construction of more realistic models. For instance, an arbitrary number of Maxwell elements in parallel, or analogously, a serial combination of Kelvin-Voigt elements could model the complex behaviour displayed by viscoelastic materials within the linear regime. With this in mind, we turn to the precise differential formulations of the two basic models.

Let σ denote the total stress, u the total displacement and $\varepsilon(u)$ the linearized strain as before. The subscripts E and V will denote elastic and viscous components respectively in this section. We begin by considering the Maxwell model. In this case, the following equations relate the elastic and viscous components of the displacement to the stress:

$$\sigma = C_E \varepsilon(u_E), \quad \sigma = C_V \varepsilon(\dot{u}_V), \quad u = u_E + u_V, \tag{2.2}$$

where C_E and C_V are fourth-order material tensors. Assuming isotropy, these take the reduced form:

$$C\tau = 2\mu\tau + \lambda \operatorname{tr} \tau I$$

where μ and λ are the scalar, possibly spatially varying, Lamé coefficients. We will not restrict our attention to the isotropic case in the following, though we shall assume that the tensors C are invertible and independent of time. The notation $A = C^{-1}$ is used throughout.

The classical formulation of the Maxwell model, given in (2.3) below, may be obtained by inverting the stress-strain relations, differentiating the equation for the elastic contributions with respect to time and adding the resulting two constitutive equations to replace $\dot{u}_E + \dot{u}_V$ by $\dot{u}$.

$$A_E\dot{\sigma} + A_V\sigma = \varepsilon(\dot{u}). \tag{2.3}$$

The equation (2.3) should be equipped with the initial condition

$$\sigma(0) = \zeta \text{ where } \operatorname{div}\zeta = g(0) \tag{2.4}$$

and g is as in (2.1). Note that the purely elastic case can be viewed as a special case of the Maxwell model, with $A_V = 0$.

For the Kelvin-Voigt model, the starting point is similar, but the compatibility conditions take a different form. We consider the same constitutive equations for the separate elastic and viscous contributions, but now the combined system reads:

$$\sigma_E = C_E\varepsilon(u), \quad \sigma_V = C_V\varepsilon(\dot{u}), \quad \sigma = \sigma_E + \sigma_V. \tag{2.5}$$

The classical formulation for the Kelvin-Voigt model can be formed by summing the contributions to give the equation

$$\sigma = C_E\varepsilon(u) + C_V\varepsilon(\dot{u}).$$

Eliminating the stress, using (2.1), results in a pure displacement formulation:

$$\operatorname{div}\left(C_E\varepsilon(u) + C_V\varepsilon(\dot{u})\right) = g. \tag{2.6}$$

An alternative would be to imitate to procedure used for Maxwell: Keep the separate stress components, invert the stress-strain relations and differentiate the equation for the elastic components with respect to time. This strategy yields the equations:

$$A_E\dot{\sigma}_E = \varepsilon(\dot{u}), \quad A_V\sigma_V = \varepsilon(\dot{u}), \quad \sigma = \sigma_E + \sigma_V. \tag{2.7}$$

The natural initial condition is now

$$\sigma_E(0) = \zeta_E, \tag{2.8}$$

where ζ_E could be calculated from an initial displacement $u(0)$ using (2.5).

The latter approach has the immediate disadvantage that additional tensor-valued variables are introduced to the system of equations. However, there are also several advantages. First, we would expect discretizations of the displacement formulation (2.6) to be wrought with the same, or more severe, problems in the nearly incompressible case as discretizations of (1.1). In contrast, the strain-stress form of the Kelvin-Voigt equations (2.7) will enable the use of robust mixed finite element methods for the stress and the displacement. Second, the mixed approach enables a unified formulation of the two models, and thus lays the foundation for flexibly treating a wide range of viscoelastic models. For these reasons, we shall pursue the formulation in (2.7). We shall conclude this section by some comments on the treatment of the generalized models.

Equations for generalized viscoelastic models may be formulated in a variety of ways; One classical formulation [**16**, p. 52] is the constitutive relationship

$$\sum_{l=0}^{k} A_l \frac{\mathrm{d}^l\sigma}{\mathrm{dt}^l} = \sum_{l=0}^{k} B_l \frac{\mathrm{d}^l\varepsilon}{\mathrm{dt}^l},$$

where A_l and B_l are again fourth-order material tensors. However, any viscoelastic model derived from a linear spring-dashpot combination, can be reduced to an equivalent series-parallel model [**40**, p. 135]. Further, the standard series-parallel models, can be reduced to a number of Maxwell elements, springs and dashpots in parallel. Such combinations, with n elements in parallel, can be expressed in the form:

$$\begin{gathered} A_E^j \dot{\sigma}_j + A_V^j \sigma_j = \varepsilon(\dot{u}) \quad j = 0, \dots, n-1, \\ \sigma = \sum_{j=0}^{n-1} \sigma_j, \end{gathered} \tag{2.9}$$

where A_E^j, A_V^j may be zero for some j, along with the initial conditions $\sigma_k(0) = \zeta_k$ for k such that $A_E^k \neq 0$. If $A_E^j \neq 0$ for all $j = 0, \dots, n-1$, the compatibility condition $\operatorname{div} \sum_j \zeta_j = g(0)$ applies. Note that these initial conditions are consistent with the initial conditions introduced above for the Maxwell and Kelvin-Voigt models. For instance, the standard linear solid model[1] takes the form:

$$\begin{gathered} A_E^0 \dot{\sigma}_0 + A_V^0 \sigma_0 = \varepsilon(\dot{u}), \quad A_E^1 \dot{\sigma}_1 = \varepsilon(\dot{u}), \\ \sigma = \sigma_0 + \sigma_1. \end{gathered} \tag{2.10}$$

The discretization approach studied in this paper can, in principle, be applied for any model in the class defined by (2.9). However, the analysis will be focused on the basic models (2.3) and (2.7). Furthermore, in Section 6, we present some numerical experiments relating to the model (2.10).

3. Weak formulations and stability estimates

The main focus of this section is to derive weak formulations for the Maxwell and the Kelvin-Voigt models. We preface this derivation by introducing appropriate notation and provide an analogy to the Hellinger-Reissner formulation for linear elasticity. We conclude the section by giving stability estimates, thus demonstrating that the weak solutions are stable with regard to the initial conditions and data.

We start by introducing some notation. Let Ω be an open, bounded domain in $\mathbb{R}^d$, $d = 2, 3$, with Lipschitz boundary $\partial\Omega$. We denote the space of square integrable functions on Ω by $L^2(\Omega)$ with inner product $\langle \cdot, \cdot \rangle$ and norm $|| \cdot ||_{0,\Omega}$. The standard Sobolev spaces $H^m(\Omega)$ have norm $|| \cdot ||_{H^m}$ and semi-norm $| \cdot |_{H^m}$ for $m = 1, 2, \dots$. The reference to the domain Ω will be omitted when context makes it superfluous. Further, a norm without subscripts will default to the L^2 norm. If $\langle A\cdot, \cdot \rangle$ is a coercive bilinear form on $L^2(\Omega)$, we shall denote the induced norm by $|| \cdot ||_A$. The linear spaces of vectors, matrices, symmetric matrices and skew-symmetric matrices are denoted by $\mathbb{V}$, $\mathbb{M}$, $\mathbb{S}$ and $\mathbb{K}$ respectively. The space of fields on Ω with square integrable components and values in X, is denoted $L^2(\Omega; X)$. Moreover, the subspace of $L^2(\Omega; \mathbb{V})$ of vector fields with square integrable divergence is denoted by $H(\operatorname{div}, \Omega; \mathbb{V})$ with the associated norm $|| \cdot ||_{\operatorname{div}}$. Accordingly, the space of matrix fields with rows in $H(\operatorname{div}, \Omega; \mathbb{V})$ is denoted $H(\operatorname{div}, \Omega; \mathbb{M})$ and the analogous subspace of symmetric matrix fields is denoted $H(\operatorname{div}, \Omega; \mathbb{S})$.

For the formulation of the Kelvin-Voigt model and the generalized models, we will require a product space of matrix fields, denoted by H^+, having the property that the sum is in $H(\operatorname{div}, \Omega; \mathbb{M})$. More precisely,

$$H^+ = \{(\tau_0, \tau_1) \in L^2(\Omega; \mathbb{M})^2 \,|\, \tau_0 + \tau_1 \in H(\operatorname{div}, \Omega; \mathbb{M})\}. \tag{3.1}$$

[1]The standard linear solid model is equivalent to the Zener model, and was studied, along with its generalized version, in [**10**].

We shall also require spaces involving time [**18**, p.285 ff.]. Let X be a Hilbert space with norm $||\cdot||_X$. $L^2(0,T;X)$ will denote the space of strongly measurable functions $u : [0,T] \to X$ with norm

$$||u||_{L^2(0,T;X)} = \left(\int_0^T ||u(t)||_X^2 \right)^{1/2} < \infty.$$

Differentiation with respect to time is denoted by a superposed dot. Finally, $H^1(0,T;X)$ denotes the space of fields $u \in L^2(0,T;X)$ such that $\dot{u} \in L^2(0,T;X)$ exists in the weak sense.

3.1. Weak formulations for linear elasticity. The Hellinger-Reissner formulation of linear elasticity, with pure Dirichlet boundary conditions for the displacement, takes the form of finding $\sigma \in H(\mathrm{div}, \Omega; \mathbb{S})$ and $u \in L^2(\Omega)$ satisfying

$$\begin{aligned} \langle A\sigma, \tau \rangle + \langle \mathrm{div}\, \tau, u \rangle &= 0 \quad \forall\, \tau \in H(\mathrm{div}, \Omega; \mathbb{S}), \\ \langle \mathrm{div}\, \sigma, w \rangle &= \langle g, w \rangle \quad \forall\, w \in L^2(\Omega), \end{aligned} \tag{3.2}$$

where $g \in L^2(\Omega)$ is a given body force and $A = A(x) : \mathbb{S} \to \mathbb{S}$ is a uniformly positive definite operator. For isotropic, homogenous elastic materials with shear modulus μ and stiffness λ, the action of A reduces to

$$A\sigma = \frac{1}{2\mu} \left(\sigma - \frac{\lambda}{2\mu + d\lambda} (\mathrm{tr}\ \sigma) I \right). \tag{3.3}$$

The derivation of this formulation relies on the invertibility of the elastic stress-strain relationship and the symmetry of the stress tensor.

Following [**13**], the existence and uniqueness of solutions to these equations depend on the existence of a positive constant α such that

$$\langle A\tau, \tau \rangle \geq \alpha ||\tau||_{\mathrm{div}} \quad \text{for all } \tau \in Z, \tag{3.4}$$

where

$$Z = \{ \tau \in H(\mathrm{div}, \Omega; \mathbb{S}) \,|\, \langle \mathrm{div}\, \tau, w \rangle = 0 \text{ for all } w \in L^2(\Omega) \}.$$

Since $\mathrm{div}\, H(\mathrm{div}, \Omega; \mathbb{S}) \subset L^2(\Omega)$, the uniform positive definiteness of the operator A on $L^2(\Omega; \mathbb{S})$ guarantees condition (3.4). In addition, there must, and does, exist a positive constant β such that

$$||v||_0 \leq \beta \sup_{\tau \in H(\mathrm{div}, \Omega; \mathbb{S})} \frac{|\langle \mathrm{div}\, \tau, v \rangle|}{||\tau||_{\mathrm{div}}}. \tag{3.5}$$

The latter condition is usually referred to as the inf-sup condition.

The positive definiteness of the isotropic A, defined by (3.3), fails as $\lambda \to \infty$. However, a uniform coercivity estimate can be established under the additional requirement that $\int \mathrm{tr}\, \sigma = 0$. Since A is also uniformly continuous in λ, this formulation makes the desired robustness in the case $\lambda \to \infty$ attainable.

Unfortunately, the construction of stable pairs of finite element spaces for the discretization of these equations has proven, in the course of four decades of research, to be nontrivial. In the last decade, stable finite element spaces, associated with a single triangulation family, have been constructed in both two and three dimensions [**1, 4, 8**]. These families of element spaces are advantageous in the sense that the approximation error of the stress can be separated from that of the displacement. On the other hand, their complexity, including complexity of implementation, makes other approaches eligible.

In this work, we shall pursue the alternative weak symmetry approach [**24**]. Instead of restricting the stress tensor space to tensor fields with symmetric values, the symmetry can be enforced by the introduction of a constraint and a Lagrange

multiplier ρ. More precisely, the weak symmetry formulation of (3.2) reads: Find $\sigma \in H(\mathrm{div}, \Omega; \mathbb{M})$, $u \in L^2(\Omega)$ and $\rho \in L^2(\Omega; \mathbb{K})$ satisfying

$$\langle A\sigma, \tau\rangle + \langle \mathrm{div}\,\tau, u\rangle + \langle \tau, \rho\rangle = 0 \qquad \forall\, \tau \in H(\mathrm{div}, \Omega; \mathbb{M}), \tag{3.6a}$$

$$\langle \mathrm{div}\,\sigma, w\rangle + \langle \sigma, \eta\rangle = \langle g, w\rangle \quad \forall\, w \in L^2(\Omega), \eta \in L^2(\Omega; \mathbb{K}). \tag{3.6b}$$

Formally, the variable ρ corresponds to the skew component of the gradient of u: $2\rho = \nabla u - \nabla u^T$. On the continuous level, the formulations (3.2) and (3.6) are equivalent. However, the weakening of the symmetry constraint opens up the possibility for simpler element spaces for the stress. In the following, we shall frequently require the trilinear form of (3.6b) and therefore label it here for later reference:

$$b(\tau, w, \eta) = \langle \mathrm{div}\,\tau, w\rangle + \langle \tau, \eta\rangle \tag{3.7}$$

The inf-sup condition for this formulation, corresponding to (3.5) for the strong symmetry formulation, guarantees the existence of a $\beta > 0$ such that for any $v \in L^2(\Omega)$ and $\eta \in L^2(\Omega; \mathbb{K})$,

$$||v||_0 + ||\eta||_0 \leq \beta \sup_{\tau \in H(\mathrm{div}, \Omega; \mathbb{M})} \frac{|b(\tau, v, \eta)|}{||\tau||_{\mathrm{div}}}. \tag{3.8}$$

The reader is referred to for example [**11**] for details and proofs of the afore stability assertions relating to (3.2) and (3.6).

3.2. Weak formulations for linear viscoelasticity. We now find ourselves in the position to derive weak formulations with weak symmetry for the Maxwell and Kelvin-Voigt models. The constitutive relationships are defined by (2.3) and (2.7) respectively and the system is closed by the balance of linear momentum (2.1), the initial conditions of (2.4) and (2.8) and boundary conditions. In the subsequent analysis, we shall assume natural homogenous boundary conditions, i.e. $u(t) = 0$ on $\partial\Omega$ for simplicity. We introduce the velocity $v = \dot{u}$ and the rotation of the velocity $\gamma = \dot{\rho}$ as these are the more natural variables for the formulations. Both $u(t)$ and $\rho(t)$ can clearly be post-calculated given an additional initial condition for u.

The strain-stress form of the Maxwell constitutive equation (2.3), in combination with (2.1), yields the following weak equations after an integration by parts of the strain term.

$$\langle A_0\sigma, \tau\rangle + \langle A_1\dot{\sigma}, \tau\rangle + b(\tau, v, \gamma) = 0 \qquad \forall\, \tau \in H(\mathrm{div}, \Omega; \mathbb{M}), \tag{3.9a}$$

$$b(\sigma, w, \eta) = \langle g, w\rangle \quad \forall\, w \in L^2(\Omega), \eta \in L^2(\Omega; \mathbb{K}), \tag{3.9b}$$

where g is given and b is defined in (3.7). We have suppressed the dependency on time in the notation of (3.9).

We proceed to consider the Kelvin-Voigt model. Multiplying (2.7) and (2.1) by test functions, integrating by parts and enforcing the symmetry of the *full* stress tensor weakly, give the weak formulation of the Kelvin-Voigt model in (3.10) below. The original subscripts V and E have been replaced by 0 and 1 respectively.

$$\langle A_1\dot{\sigma}_1, \tau_1\rangle + \langle A_0\sigma_0, \tau_0\rangle + b(\tau_0 + \tau_1, v, \gamma) = 0 \quad (\tau_0, \tau_1) \in H^+, \tag{3.10a}$$

$$b(\sigma_0 + \sigma_1, w, \eta) = \langle g, w\rangle \quad w \in L^2(\Omega), \eta \in L^2(\Omega; \mathbb{K}), \tag{3.10b}$$

where H^+ is as defined by (3.1).

We further observe that a weak formulation for the generalized models defined by (2.9) can be derived in the entirely analogous manner. We enforce the symmetry of the total stress tensor $\sigma = \sum_j \sigma_j$ weakly, i.e. $\langle \sigma, \eta\rangle = 0$ for all $\eta \in L^2(\Omega; \mathbb{K})$, as

for the Kelvin-Voigt model. The weak equations then take the form:

$$\langle A_E^j \dot{\sigma}_j, \tau_j \rangle + \langle A_V^j \sigma_j, \tau_j \rangle + b(\tau_j, v, \gamma) = 0 \quad \forall\, \tau_j, \quad j = 0, \ldots, n-1,$$
$$b(\sum_j \sigma_j, w, \eta) = \langle g, w \rangle \quad \forall\, w, \eta.$$

These derivations demonstrate that this type of formulation is well-suited for both the basic and the generalized models. We therefore have a unified framework in place. At this point however, we let the generalized models rest in order to avoid notational overflow.

3.3. Existence and regularity of solutions. We shall give arguments for the existence of solutions to the systems of equations (3.9) and (3.10) in order to motivate the assumptions to follow.

First, consider the Maxwell model of (3.9) with the initial condition $\sigma(0) = \zeta \in H(\mathrm{div}, \Omega; \mathbb{M})$. Assume that $g \in H^1(0,T; L^2(\Omega))$ and that $\mathrm{div}\, \zeta = g(0)$. Then, there exist $\sigma_e \in H^1(0,T; H(\mathrm{div}, \Omega; \mathbb{M}))$, $u_e \in H^1(0,T; L^2(\Omega))$ and $\rho_e \in H^1(0,T; L^2(\Omega; \mathbb{K}))$ solving the elasticity equations (3.6) with $A = A_0$ for each $t \in [0,T]$. Next, let $H_0 = \{\tau \in H(\mathrm{div}, \Omega; \mathbb{M}) \,|\, \mathrm{div}\, \tau = 0\}$. Since A_0, A_1 are bounded on H_0, there exists a $\sigma_0 \in H^1(0,T; H_0)$ satisfying the ordinary differential equation:

$$\langle A_1 \dot{\sigma}_0, \tau \rangle + \langle A_0 \sigma_0, \tau \rangle = -\langle A_1 \dot{\sigma}_e, \tau \rangle \quad \tau \in H_0$$

with the initial condition $\sigma_0(0) = \zeta - \sigma_e(0) \in H_0$. Further, the inf-sup condition (3.8), gives the existence of $v_0(t) \in L^2(\Omega)$, $\gamma_0(t) \in L^2(\Omega; \mathbb{K})$ for a.e t such that

$$\langle A_1 \dot{\sigma}_0, \tau \rangle + \langle A_0 \sigma_0, \tau \rangle + b(\tau, v_0, \gamma_0) = -\langle A_1 \dot{\sigma}_e, \tau \rangle \quad \tau \in H(\mathrm{div}, \Omega; \mathbb{M}),$$
$$b(\sigma_0, w, \eta) = 0 \quad w \in L^2(\Omega), \eta \in L^2(\Omega; \mathbb{K}).$$

It follows that $\sigma = \sigma_e + \sigma_0$, $v = u_e + v_0$ and $\gamma = \rho_e + \gamma_0$ solve the Maxwell equations (3.9) for a.e t and the initial condition $\sigma(0) = \zeta$ with $\mathrm{div}\, \sigma(0) = g$. Further, $\sigma \in H^1(0,T; H(\mathrm{div}, \Omega; \mathbb{M}))$, $v \in L^2(0,T; L^2(\Omega))$ and$\gamma \in L^2(0,T; L^2(\Omega; \mathbb{K}))$. This existence argument motivates Definition 3.1.

Definition 3.1 (Weak solutions of the Maxwell equations). *Assume that*

$$g \in H^1(0,T; L^2(\Omega)),\ \zeta \in H(\mathrm{div}, \Omega; \mathbb{M}),\ \mathrm{div}\, \zeta = g(0).$$

The fields $\sigma \in H^1(0,T; H(\mathrm{div}, \Omega; \mathbb{M}))$, $v \in L^2(0,T; L^2(\Omega))$, $\gamma \in L^2(0,T; L^2(\Omega; \mathbb{K}))$ constitute a weak solution of the Maxwell equations provided (3.9) *is satisfied for a.e. $t \in (0,T]$ and the initial condition $\sigma(0) = \zeta$ holds.*

Uniqueness of these solutions is a consequence of Theorem 3.3 below.

We proceed to consider the Kelvin-Voigt model with the initial condition $\sigma_1(0) = \zeta_1 \in L^2(\Omega; \mathbb{M})$ and assume that $g \in L^2(0,T; L^2(\Omega))$. Solving the stationary elasticity equation

$$\langle C_1 \varepsilon(\nu), \varepsilon(v) \rangle = \langle \zeta_1, \nabla\, v \rangle \quad \forall v \in H_0^1(\Omega; \mathbb{V}),$$

yields $\nu \in H_0^1(\Omega; \mathbb{V})$. Next, from a weak formulation of (2.6):

$$\langle C_1 \varepsilon(u(t)) + C_0 \varepsilon(\dot{u}(t)), \varepsilon(v) \rangle = \langle g(t), v \rangle, \quad \forall\, v \in H_0^1(\Omega; \mathbb{V}),$$

equipped with the initial condition $u(0) = \nu$, we can deduce the existence of a solution $u \in H^1(0,T; H^1(\Omega; \mathbb{V}))$. In accordance with (2.5), we define

$$\sigma_0(t) = C_0 \varepsilon(\dot{u}(t)) \in L^2(0,T; L^2(\Omega; \mathbb{M})),$$
$$\sigma_1(t) = C_1 \varepsilon(u(t)) + \zeta_1 - C_1 \varepsilon(\nu) \in H^1(0,T; L^2(\Omega; \mathbb{M})),$$
$$\gamma(t) = \frac{1}{2} \left(\nabla\, \dot{u} - \nabla\, \dot{u}^T \right) \in L^2(0,T; L^2(\Omega; \mathbb{K})).$$

It follows that $\mathrm{div}(\sigma_0 + \sigma_1) = g$ and $\sigma_1(0) = \zeta_1$.

DEFINITION 3.2 (Weak solutions of the Kelvin-Voigt equations). *Assume that*

$$g \in L^2(0,T;L^2(\Omega)),\ \zeta_1 \in L^2(\Omega;\mathbb{M}).$$

The fields $\sigma_0 \in L^2(0,T;L^2(\Omega;\mathbb{M}))$, $\sigma_1 \in H^1(0,T;L^2(\Omega;\mathbb{M}))$ *satisfying* $(\sigma_0(t),\sigma_1(t)) \in H^+$ *for a.e.* t, $v \in L^2(0,T;L^2(\Omega))$ *and* $\gamma \in L^2(0,T;L^2(\Omega;\mathbb{K}))$ *constitute a weak solution of the Kelvin-Voigt equations provided* (3.10) *is satisfied for a.e* $t \in (0,T]$, *and the initial condition* $\sigma_1(0) = \zeta_1$ *holds.*

These solutions are indeed unique cf. Theorem 3.4 below.

3.4. Stability estimates. The two formulations (3.9), (3.10) can be viewed as evolutionary problems subject to a constraint. The two theorems closing this section provide stability estimates for the Maxwell and the Kelvin-Voigt equations respectively, thus demonstrating stability in time. The estimates rely on two main factors. First, the inf-sup condition (3.8) holds for the spaces $H(\mathrm{div},\Omega;\mathbb{M})$ and $L^2(\Omega)\times L^2(\Omega;\mathbb{K})$. Second, the assumption that the operators $A_j : L^2(\Omega;\mathbb{M}) \to L^2(\Omega;\mathbb{M})$ are uniformly positive definite and continuous, i.e. there exist positive constants α_j and c_j such that

$$\langle A_j\sigma,\sigma\rangle \geq \alpha_j||\sigma||_0^2, \quad \langle A_j\sigma,\tau\rangle \leq c_j||\sigma||_0||\tau||_0.$$

We shall make this assumption here and throughout. As a consequence, each A_j induces a norm equivalent to the L^2 norm: $\langle A_j\tau,\tau\rangle = ||\tau||_{A_j}^2$.

For the Maxwell equations (3.9), we start by observing that if there is no applied body force, i.e. $g = 0$, then

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}||\sigma||_{A_1}^2 + ||\sigma||_{A_0}^2 = 0.$$

Using the equivalence of the norms induced by A_0 and A_1 and Grönwall's inequality, we obtain the estimate:

$$||\sigma(t)||_{A_1}^2 \leq e^{-\frac{2\alpha_0}{c_1}t}||\sigma(0)||_{A_1}^2$$

In other words, the energy, measured in the A_1-norm of the stress, decays exponentially from its initial state. The situation with a constant body force, $\dot{g} = 0$, can also be reduced to a system with $g = 0$, and we can therefore derive a similar estimate in that case. The case $g \neq 0$ is covered in the following theorem.

THEOREM 3.3 (Stability estimates for Maxwell). *Let* (σ,v,γ) *be a weak solution of the Maxwell equations. There exists a positive constant* c *such that*

$$||\sigma(t)||_{A_1}^2 \leq e^{-\frac{\alpha_0}{c_1}t}||\zeta||_{A_1}^2 + c\int_0^t e^{-\frac{\alpha_0}{c_1}(t-s)}\left(||g(s)||^2 + ||\dot{g}(s)||^2\right)\mathrm{d}s \tag{3.11}$$

and

$$||v(t)||^2 + ||\gamma(t)||^2 + ||\dot{\sigma}(t)||^2 \leq c\left(||\dot{g}(t)||^2 + ||\sigma(t)||^2\right) \tag{3.12}$$

for a.e. $t \in [0,T]$.

PROOF. Let $(\sigma(t),v(t),\gamma(t))$ satisfy (3.9) for a.e. t. In order to obtain estimates for the stress, we use $\tau = \sigma(t)$ in (3.9a) and apply (3.9b) to find that

$$||\sigma(t)||_{A_0}^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}||\sigma(t)||_{A_1}^2 = -\langle g,v\rangle. \tag{3.13}$$

Clearly, in order to bound the stress, we need a bound for the velocity v. Using Cauchy-Schwartz for the norms induced by A_0 and A_1 and the inf-sup condition of (3.8) with constant β, it follows that

$$||v(t)||_0 + ||\gamma(t)||_0 \leq \beta\left(c_0^{1/2}||\sigma||_{A_0} + c_1^{1/2}||\dot{\sigma}||_{A_1}\right). \tag{3.14}$$

Next, using $\tau = \dot{\sigma}(t)$ in (3.9a), and (3.9b) differentiated with respect to time, give an analogy to (3.13) for $\dot{\sigma}$:

$$\langle A_0\sigma, \dot{\sigma}\rangle + ||\dot{\sigma}(t)||^2_{A_1} = -\langle \dot{g}, v\rangle.$$

Together with (3.14), this gives the desired estimate for $\dot{\sigma}$ in terms of σ.

$$||\dot{\sigma}(t)||^2_{A_1} \leq c\left(||\sigma(t)||^2_{A_0} + ||\dot{g}(t)||^2\right).$$

This estimate and (3.14) give (3.12). Furthermore, combining this estimate with (3.13) and (3.14), we obtain

$$\frac{\alpha_0}{2c_1}||\sigma(t)||^2_{A_1} + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}||\sigma(t)||^2_{A_1} \leq c\left(||g(t)||^2 + ||\dot{g}||^2\right).$$

Hence, (3.11) follows by Grönwall's inequality. □

Corresponding estimates can be derived for the Kelvin-Voigt equations and are summarized in the following theorem.

THEOREM 3.4 (Stability estimates for Kelvin-Voigt). *Let $(\sigma_0, \sigma_1, v, \gamma)$ be a weak solution of the Kelvin-Voigt equations. There exists a positive constant c such that*

$$||\sigma_1(t)||^2_{A_1} + \int_0^t ||\sigma_0(s)||^2_{A_0} + ||v(s)||^2 + ||\gamma(s)||^2\,\mathrm{d}s \leq ||\zeta_1||^2_{A_1} + c\int_0^t ||g(s)||^2\,\mathrm{d}s.$$

for $t \in [0, T]$.

PROOF. The proof is straightforward: Letting $\tau_0 = \sigma_0$ and $\tau_1 = \sigma_1$ in (3.10a) and using (3.10b), we obtain the relation

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}||\sigma_1||^2_{A_1} + ||\sigma_0||^2_{A_0} = -\langle g, v\rangle. \tag{3.15}$$

An estimate for the velocity and the rotation follows from (3.8):

$$||v||_0 + ||\gamma||_0 \leq \beta \sup_\tau ||\tau||^{-1}_{\mathrm{div}} |b(\tau, v, \gamma)| = \beta \sup_\tau ||\tau||^{-1}_{\mathrm{div}} |\langle A_0\sigma_0, \tau\rangle| \leq c||\sigma_0||_{A_0}.$$

Combining this with (3.15) gives the final result. □

We pause to remark that although we focus on the weak symmetry formulation in this paper, the strong symmetry approach also yields a meaningful base for discretization. The previous energy estimates clearly also hold for the latter approach with γ ignored, due to the equivalence between the formulations at the continuous level. For the discrete stability and error estimates in the subsequent sections, the path of the proofs would be the same, and would to some extent be simplified with a strong, rather than the weak, symmetry constraint.

Now, having established the desired stability properties of the formulations, we move on to consider discretization strategies in space and time. We shall focus on the semi-discrete problem and conforming finite element discretizations in space in Section 4. The full discretizations in time and space will be considered in Section 5.

4. The semi-discrete problem and stability

The scope of this section is to consider conforming finite element spatial discretizations of the two systems of equations (3.9) and (3.10). In particular, we demonstrate that the same discretization will be appropriate for both models. We shall again start by drawing inspiration from suitable finite element discretizations of the linear elasticity equations and comment on the properties of such, before embarking on the discretization of the viscoelasticity models.

Assume that $\{\mathcal{T}_h\}_h$ is a shape-regular family of admissible, simplicial tessellations of Ω, where h measures the mesh size. We are interested in finite element

spaces $\Sigma_h \subset H(\mathrm{div}, \Omega; \mathbb{M})$, $V_h \subset L^2(\Omega)$ and $Q_h \subset L^2(\Omega; \mathbb{K})$ subordinate to this tessellation. From the stationary theory of mixed finite element methods [**13**], we know that the discrete spaces yield a stable discretization of a given weak formulation if the Brezzi conditions hold. For the weak symmetry formulation of the linear elasticity equations (3.6), these conditions are the discrete equivalents of (3.4) and (3.8), and take the following form: For the spaces Σ_h, V_h and Q_h, there must exist positive constants α and β, independent of h, such that

$$\langle A\tau, \tau \rangle \geq \alpha ||\tau||^2_{\mathrm{div}} \quad \forall\, \tau \in Z_h, \tag{4.1}$$

$$||w|| + ||\eta|| \leq \beta \sup_{\tau \in \Sigma_h} \frac{|b(\tau, w, \eta)|}{||\tau||_{\mathrm{div}}} \quad \forall\, w \in V_h, \eta \in Q_h, \tag{4.2}$$

where $b(\tau, w, \eta) = \langle \mathrm{div}\, \tau, w \rangle + \langle \tau, \eta \rangle$ as before and the kernel Z_h is defined as

$$Z_h = \{\tau \in \Sigma_h \,|\, b(\tau, w, \eta) = 0 \quad \forall\, w \in V_h, \eta \in Q_h\}.$$

There exists a multitude of finite element spaces satisfying these conditions. A selection of such can be sampled from the references [**3, 5, 6, 24, 21, 39**] and the survey [**20**]. In the following analysis, we shall assume that the spaces Σ_h, V_h and Q_h satisfy conditions (4.1) and (4.2) and additionally are such that

$$\mathrm{div}\, \Sigma_h \subset V_h. \tag{4.3}$$

Let P_h denote the L^2 projection from $L^2(\Omega; \mathbb{K})$ onto Q_h and, with a minor abuse of notation, from $L^2(\Omega)$ onto V_h. We let Π_h be a projection onto Σ_h such that

$$\langle \mathrm{div}(\tau - \Pi_h \tau), w \rangle = 0 \quad \text{for all } w \in V_h. \tag{4.4}$$

We shall assume that the projection Π_h is bounded, i.e. that there exists a constant c such that

$$||\Pi_h \tau||_{\mathrm{div}} \leq c ||\tau||_{\mathrm{div}} \quad \text{for all } \tau \in H(\mathrm{div}, \Omega; \mathbb{M}). \tag{4.5}$$

Note that the canonical projection onto Σ_h, defined by the degrees of freedom, is typically *not* bounded on $H(\mathrm{div}, \Omega; \mathbb{M})$ and thus does not satisfy (4.5). However, bounded projections, satisfying (4.5), can be constructed through smoothing, cf. [**7, 15**].

4.1. Two stable sets of element spaces. We shall describe two families of element spaces, namely those introduced by Arnold et al. [**6**] and Falk [**20**], both satisfying the conditions (4.1), (4.2) and (4.3) above. The lowest order element spaces of these families were also suggested by Farhloul and Fortin [**21**]. The former family of spaces will mainly be used for the numerical experiments in Section 6.

The particular finite element spaces Σ_h, V_h and Q_h introduced by Arnold et al. are as follows: The lowest order elements are the combination of linear vector polynomials with continuity of normal components over inter-element facets for the stress, and piecewise constants for the velocity and the rotation approximations. The element spaces generalize to arbitrary polynomial degree; Let $\mathcal{P}_k(\mathcal{T}_h)$ denote the space of discontinuous k'th order polynomials defined on the tessellation $\mathcal{T}_h$ of $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, and BDM_k denote the k'th order vector polynomials with continuous normal components over inter-element facets as introduced by Brezzi et al. [**14**] and extended by Nedelec [**31**] to three dimensions. With this notation, the k'th order elasticity elements, for $k \geq 1$, are:

$$\Sigma_{h,k} = \mathrm{BDM}_k(\mathcal{T}_h; \mathbb{V}), \quad V_{h,k} = \mathcal{P}_{k-1}(\mathcal{T}_h; \mathbb{V}), \quad Q_{h,k} = \mathcal{P}_{k-1}(\mathcal{T}_h; \mathbb{K}) \tag{4.6}$$

For these element spaces, we have the following interpolation estimates for the projections P_h^k and Π_h^k:

$$||\tau - \Pi_h^k \tau||_0 \leq ch^{m+1} \, |\tau|_{H^{m+1}},$$
$$||\operatorname{div}(\tau - \Pi_h^k \tau)||_0 \leq ch^m \, |\operatorname{div} \tau|_{H^m},$$
$$||p - P_h^k p||_0 \leq ch^m \, |p|_{H^m},$$

for $1 \leq m \leq k$.

For the weak symmetry formulation of the elasticity equations, these interpolation properties carry over to the approximation properties of the finite element discretization, with the exception that the error of the stress approximation cannot be split from that of the rotation, since $\operatorname{skw} \Sigma_h \not\subset Q_h$. Hence, the higher-order L^2 interpolation error of the stress is not expected to be conserved.

Another stable family of element spaces, introduced by Falk [**20**] for the lowest order case, uses the same spaces Σ_h and V_h as of (4.6), but the space of piecewise constants for the rotation Q_h is replaced by continuous piecewise linears $\tilde{Q}_h$. In general, we let

$$\tilde{Q}_{h,k} = \mathcal{P}_k(\mathcal{T}_h; \mathbb{K}) \cap C^0(\mathcal{T}_h; \mathbb{K}) \tag{4.7}$$

In this case, the interpolation error of the rotation is of the same order as the L^2 error of the stress interpolation. We can thus expect to retain the higher order L^2 estimates for the stress approximation.

Having discussed stable and accurate finite element spaces for discretizations of the elasticity equations, we now turn to the questions of stability and spatial a priori error estimates for the Maxwell and the Kelvin-Voigt models. Attention is paid to the Maxwell model in Section 4.2 and to the Kelvin-Voigt model in Section 4.3. The techniques involved in the following are fairly standard and we rely on the results for the stationary elasticity equations.

For ease of reading, we give the results applied to the element spaces (4.6) here. Let the subscript h indicate discrete solutions and assume sufficient smoothness of the domain and the data, all which will be made precise in the subsequent sections. Then, for the Maxwell model, we will show that

$$||\sigma_h(t) - \Pi_h^k \sigma(t)||_0 + ||v_h(t) - P_h^k v(t)||_0 + ||\gamma_h(t) - P_h^k \gamma(t)||_0 \leq c_T h^k,$$

for a.e. $t \in [0, T]$. For the Kelvin-Voigt model,

$$||\sigma_{1,h}(t) - \Pi_h^k \sigma_1(t)||_0 + \int_0^t ||\sigma_{0,h}(s) - \Pi_h^k \sigma_0(s)||_0 \, \mathrm{d}s + \int_0^t ||v_h(s) - P_h^k v(s)||_0 + ||\gamma_h(s) - P_h^k \gamma(s)||_0 \, \mathrm{d}s \leq c_T h^k.$$

4.2. Semi-discretization of the Maxwell model. For the Maxwell equations, the natural finite element spaces correspond directly to those of the elasticity equations: $\Sigma_h \subset H(\operatorname{div}, \Omega; \mathbb{M})$, $V_h \subset L^2(\Omega)$ and $Q_h \subset L^2(\Omega; \mathbb{K})$. The spatially discretized equations follow immediately:

$$\langle A_0 \sigma_h(t), \tau \rangle + \langle A_1 \dot{\sigma}_h(t), \tau \rangle + b(\tau, v_h(t), \gamma_h(t)) = 0, \tag{4.8a}$$
$$b(\sigma_h(t), w, \eta) = \langle g(t), w \rangle, \tag{4.8b}$$

for all $\tau \in \Sigma_h$, $w \in V_h$ and $\eta \in Q_h$ and for a.e $t \in (0, T]$. We also enforce the discrete initial condition $\sigma_h(0) = \zeta_h \in \Sigma_h$ Assuming that $\langle \operatorname{div} \zeta_h - g(0), w \rangle = 0$ for all $w \in V_h$, we have semi-discrete solutions $\sigma_h \in H^1(0, T; \Sigma_h)$, $v_h \in L^2(0, T; V_h)$ and $\gamma_h \in L^2(0, T; Q_h)$. We observe that, if the spaces Σ_h, V_h and Q_h are such that the discrete Brezzi conditions (4.1) and (4.2) hold, the energy estimates for the

continuous formulation carry over to the semi-discrete formulation. Thus, Theorem 3.3 also holds, with the obvious modifications, for the semi-discrete solutions.

In the subsequent error analysis however, stability estimates for the discrete equations with additional source terms will be required:

$$\langle A_0\sigma_h(t),\tau\rangle + \langle A_1\dot\sigma_h(t),\tau\rangle + b(\tau,v_h(t),\gamma_h(t)) = \langle F(t),\tau\rangle \tag{4.9a}$$

$$b(\sigma_h(t),w,\eta) = \langle G(t),\eta\rangle \tag{4.9b}$$

For clarity of presentation, such estimates are presented in the following lemma. Clearly, the lemma also holds for the continuous solutions with additional source terms.

LEMMA 4.1 (Discrete stability with source terms for Maxwell). *Assume that* $G \in H^1(0,T;L^2(\Omega;\mathbb{M}))$, $F \in L^2(0,T;L^2(\Omega;\mathbb{M}))$ *and that* (σ_h,v_h,γ_h) *solve* (4.9) *for a.e.* $t\in(0,T]$. *Then there exists a constant* $c>0$ *such that*

$$||\sigma_h(t)||_{A_1}^2 \le e^{-ct}||\sigma_h(0)||_{A_1}^2 + c\int_0^t e^{-c(t-s)}K(s)\,\mathrm{d}s,$$

$$||v_h(t)||^2 + ||\gamma_h(t)||^2 + ||\dot\sigma_h(t)||^2 \le c\left(||F(t)||^2 + ||\dot G(t)||^2 + ||\sigma_h(t)||^2\right),$$

where $K(s) = ||F(s)||^2 + ||G(s)||^2 + ||\dot G(s)||^2$.

PROOF. Let $\tau=\sigma_h(t), w=v_h(t), \eta=\gamma_h(t)$ in (4.9). We obtain the equation

$$||\sigma_h(t)||_{A_0}^2 + \frac{\mathrm{d}}{\mathrm{dt}}\frac{1}{2}||\sigma_h(t)||_{A_1}^2 = \langle F(t),\sigma_h(t)\rangle - \langle G(t),\gamma_h(t)\rangle$$

The discrete inf-sup condition (4.2) gives a bound for the velocity and rotation:

$$||v_h|| + ||\gamma_h|| \le c\left(||\sigma_h|| + ||\dot\sigma_h|| + ||F||\right).$$

Further, let $\tau=\dot\sigma_h(t), w=v_h(t), \eta=\gamma_h(t)$ in (4.9) after differentiating the second equation in time. We obtain the equation

$$\langle A_0\sigma_h(t),\dot\sigma_h(t)\rangle + ||\dot\sigma_h(t)||_{A_1}^2 = \langle F(t),\dot\sigma_h(t)\rangle - \langle\dot G(t),\gamma_h(t)\rangle.$$

The same techniques as employed in the proof of Theorem 3.3 give the final estimates. □

We are now in the position to easily derive error estimates for the semi-discrete equations. To this aim, we introduce the following standard notation: The discretization error is split into a projection error $\mathcal{E}$ and an approximation error E. For the stress this takes the form

$$\sigma-\sigma_h = \mathcal{E}^\sigma - E^\sigma = (\sigma-\Pi_h\sigma) - (\sigma_h-\Pi_h\sigma). \tag{4.10}$$

and we have the analogies for the velocity and the rotation, in terms of the projection P_h, with superscripts v and γ respectively. Observe that

$$b(\mathcal{E}^\sigma,w,\eta) = \langle\mathcal{E}^\sigma,\eta\rangle, \tag{4.11}$$

for all $w\in V_h$, $\eta\in Q_h$ by the definition of the projection Π_h onto Σ_h (4.4). Moreover, due to (4.3), we have that

$$b(\tau,\mathcal{E}^v,\mathcal{E}^\gamma) = \langle\tau,\mathcal{E}^\gamma\rangle \tag{4.12}$$

for all $\tau\in\Sigma_h$.

Let $K(t) = ||\mathcal{E}^\sigma(t)||^2 + ||\mathcal{E}^{\dot\sigma}(t)||^2 + ||\mathcal{E}^\gamma(t)||^2$. By definition, the function K is bounded by the approximation properties of the projections Π_h and P_h and of the smoothness of the exact solution (σ,v,γ). As a result of the theorem below, we obtain that the full error, which is bounded by $||E||+||\mathcal{E}||$, is of the same order as long as the initial approximation is sufficiently accurate.

THEOREM 4.2 (A priori error estimates for Maxwell). *The approximation errors for the stress, velocity and rotation satisfy the following bounds for a.e $t \in [0,T]$ in terms of the projection errors*

$$||E^\sigma(t)||^2 \leq e^{-ct}||\sigma_h(0) - \Pi_h\sigma(0)||^2_{A_1} + c\int_0^t e^{-c(t-s)}K(s)\,\mathrm{d}s$$
$$||E^{\dot\sigma}(t)||^2 + ||E^v(t)||^2 + ||E^\gamma(t)||^2 \leq c\left(||E^\sigma(t)||^2 + K(t)\right)$$

PROOF. We omit the reference to the time t for brevity in the following. We begin by subtracting the discrete Maxwell equations from the continuous ones to obtain the error equations:

$$\langle A_0(\sigma - \sigma_h), \tau\rangle + \langle A_1(\dot\sigma - \dot\sigma_h), \tau\rangle + b(\tau, v - v_h, \gamma - \gamma_h) = 0 \quad \forall\, \tau \in \Sigma_h,$$
$$b(\sigma - \sigma_h, w, \eta) = 0 \quad \forall\, w \in V_h, \eta \in Q_h.$$

Inserting the error decompositions (4.10), we have:

$$\langle A_0E^\sigma, \tau\rangle + \langle A_1E^{\dot\sigma}, \tau\rangle + b(\tau, E^v, E^\gamma) = \langle A_0\mathcal{E}^\sigma, \tau\rangle + \langle A_1\mathcal{E}^{\dot\sigma}, \tau\rangle + \langle \tau, \mathcal{E}^\gamma\rangle$$
$$b(E^\sigma, w, \eta) = b(\mathcal{E}^\sigma, w, \eta) = \langle \mathcal{E}^\sigma, \eta\rangle.$$

where the last equality follows from the definition of the projection onto Σ_h (4.11). The term $\langle \tau, \mathcal{E}^\gamma\rangle$ is a result of (4.12). It is now easy to observe that E^σ, E^v and E^γ solve the semi-discrete equations (4.9) with the right-hand side(s):

$$F = A_0\mathcal{E}^\sigma + A_1\mathcal{E}^{\dot\sigma} + \mathcal{E}^\gamma, \quad G = \mathcal{E}^\sigma,$$

and the initial condition $E^\sigma(0) = \sigma_h(0) - \Pi_h\sigma(0)$. The stability estimate with additional source terms, Lemma 4.1, thus gives the error estimates. □

4.3. Semi-discretization of the Kelvin-Voigt model. This subsection aims at introducing a spatial discretization of the Kelvin-Voigt model corresponding to the one introduced for the Maxwell model in the previous. Moreover, error estimates for the semi-discrete solutions are established through an analogous extended energy estimate.

Recall that for the Kelvin-Voigt equations, we seek two components of the stress: σ_0 and σ_1 such that $\sigma_0 + \sigma_1 = \sigma \in H(\mathrm{div}, \Omega; \mathbb{M})$, or with the notation introduced in (3.1), $(\sigma_0, \sigma_1) \in H^+$. For discretization purposes, we are therefore interested in a finite dimensional product space Σ_h^2 approximating H^+. A natural choice would be to let

$$\Sigma_h^2 = \Sigma_h \times \Sigma_h. \tag{4.13}$$

The resulting spatially discretized equations are presented below in (4.14).

$$\langle A_1\dot\sigma_{1,h}(t), \tau_1\rangle + \langle A_0\sigma_{0,h}(t), \tau_0\rangle + b(\tau_0 + \tau_1, v_h(t), \gamma_h(t)) = 0, \tag{4.14a}$$
$$b(\sigma_{0,h}(t) + \sigma_{1,h}(t), w, \eta) = \langle g(t), w\rangle, \tag{4.14b}$$

for all $(\tau_0, \tau_1) \in \Sigma_h^2$, $w \in V_h$ and $\eta \in Q_h$ and for a.e $t \in (0,T]$. As usual, we enforce the discretized initial condition: $\sigma_{1,h}(0) = \zeta_{1,h} \in \Sigma_h$. Again, we comment that, provided that the spaces Σ_h, V_h and Q_h are such that (4.1) and (4.2) hold, the energy estimates for the continuous formulations carry over to the semi-discrete formulations. Thus, Theorem 3.4 also holds, with the obvious modifications, for the semi-discrete solutions. We also note that the projections constructed in [**15**] are bounded in $L^2(\Omega;\mathbb{M}) \mapsto L^2(\Omega;\mathbb{M})$ and hence the projections of the separate stress components can be defined even if each component is not in $H(\mathrm{div}, \Omega; \mathbb{M})$.

As in the case of the Maxwell model, a stability estimate for the equations with additional source terms is a key step in deriving error estimates. Such a result

is considered separately in the following lemma. We shall however, restrict our attention to source terms F_0, F_1 yielding the following alternative to (4.14a):

$$\langle A_1\dot\sigma_{1,h},\tau_1\rangle + \langle A_0\sigma_{0,h},\tau_0\rangle + b(\tau_0+\tau_1,v_h,\gamma_h) = \langle \dot F_1,\tau_1\rangle + \langle F_0,\tau_0\rangle, \tag{4.15a}$$

$$b(\sigma_{0,h}+\sigma_{1,h},w,\eta) = \langle G,\eta\rangle. \tag{4.15b}$$

Recall that in the derivation of the Kelvin-Voigt model, the elastic stress-strain relation corresponding to σ_1 was differentiated in time. The $\dot F_1$ term of (4.15a) takes this aspect into account.

LEMMA 4.3 (Stability with source terms for Kelvin-Voigt). *Let*

$$F_0 \in L^2(0,T;L^2(\Omega;\mathbb{M})),\ F_1 \in H^1(0,T;L^2(\Omega;\mathbb{M})),\ G \in L^2(0,T;L^2(\Omega;\mathbb{M})),$$

and assume that $(\sigma_{0,h},\sigma_{1,h},v_h,\gamma_h)$ *solve* (4.15) *for a.e.* $t\in(0,T]$. *Then there exists a positive constant* c *such that*

$$||\sigma_{1,h}(t)||^2 + \int_0^t ||\sigma_{0,h}(s)||^2 + ||v_h(s)||^2 + ||\gamma_h(s)||^2\,\mathrm{d}s \le c\left(I + ||F_1(t)||^2 + \int_0^t K(s)\,\mathrm{d}s\right),$$

for $t\in[0,T]$, *where* $I = ||\sigma_{1,h}(0)||^2 + ||F_1(0)||^2$ *and* $K(s) = ||G(s)||^2 + ||F_0(s)||^2 + ||F_1(s)||^2 + ||\dot F_1(s)||^2$.

PROOF. We omit the subscripts h for notational brevity in the following. First, observe that letting $\tau_1 = -\tau_0 = \tau$ in (4.15a) gives the identity

$$\langle A_1\dot\sigma_1,\tau\rangle - \langle A_0\sigma_0,\tau\rangle = \langle \dot F_1 - F_0,\tau\rangle$$

for any $\tau\in\Sigma_h$. In particular, letting $\tau = \dot\sigma_1$, we find that

$$||\dot\sigma_1||_{A_1}^2 \le c\left(||\sigma_0||_{A_0}^2 + ||\dot F_1||^2 + ||F_0||^2\right). \tag{4.16}$$

Second, let $\tau_0 = \sigma_0$, $\tau_1 = \sigma_1$, $w = v$ and $\eta = \gamma$ in (4.15). We are left with:

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}||\sigma_1||_{A_1} + ||\sigma_0||_{A_0} = -\langle G,\gamma\rangle + \langle \dot F_1,\sigma_1\rangle + \langle F_0,\sigma_0\rangle.$$

Since, $||v|| + ||\gamma|| \le c\,(||\sigma_0||_{A_0} + ||F_0||)$ by (4.2), it follows that

$$||\sigma_1(t)||_{A_1}^2 + \int_0^t ||\sigma_0(s)||_{A_0}^2\,\mathrm{d}s \le ||\sigma_1(0)||_{A_1}^2 + c\int_0^t ||G(s)||^2 + ||F_0(s)||^2\,\mathrm{d}s + \int_0^t \langle \dot F_1(s),\sigma_1(s)\rangle\,\mathrm{d}s \tag{4.17}$$

We integrate the last term by parts to see that

$$\int_0^t \langle \dot F_1(s),\sigma_1(s)\rangle\,\mathrm{d}s \le \langle F_1(s),\sigma_1(s)\rangle|_{s=0}^t + \frac{1}{4\epsilon}\int_0^t ||F_1(s)||^2\,\mathrm{d}s + \epsilon\int_0^t ||\dot\sigma_1(s)||^2\,\mathrm{d}s$$

for any $\epsilon > 0$. Thus, combining (4.17) with (4.16), gives the stated result. □

As before, cf. (4.10), we introduce the approximation errors: $E_h^\tau = ||\Pi_h\tau - \tau_h||$ and the projection errors: $\mathcal{E}^\tau = ||\tau - \Pi_h\tau||$ for $\tau\in\{\sigma_0,\sigma_1\}$ and analogously for $\tau\in\{v,\gamma\}$ with Π_h replaced by P_h. As a consequence of Lemma 4.3, we obtain the following error estimates.

THEOREM 4.4 (A priori error estimates for Kelvin-Voigt). *Let* $(\sigma_0, \sigma_1, v, \gamma)$ *solve the continuous Kelvin-Voigt equations* (3.10) *and* $(\sigma_{0,h}, \sigma_{1,h}, v_h, \gamma_h)$ *be approximations satisfying the discrete equations* (4.14). *Then there exists* $c > 0$ *such that*

$$||E^{\sigma_1}(t)||^2 + \int_0^t ||E^{\sigma_0}(s)||^2 + ||E^v(s)||^2 + ||E^\gamma(s)||^2 \,\mathrm{d}s \\ \leq c||\sigma_{1,h}(0) - \Pi_h\sigma_1(0)||^2 + c\left(I(0) + I(t) + \int_0^t L(s)\,\mathrm{d}s\right)$$

where $I(t) = ||\mathcal{E}^{\sigma_1}(t)||^2 + ||\mathcal{E}^\rho(t)||^2$, *and*

$$L = ||\mathcal{E}^{\sigma_0}||^2 + ||\mathcal{E}^{\sigma_1}||^2 + ||\mathcal{E}^{\dot{\sigma}_1}||^2 + ||\mathcal{E}^\rho||^2 + ||\mathcal{E}^\gamma||^2.$$

Again we observe that this result allows us to bound the full error in terms of the projections Π_h and P_h and the properties of the exact solution.

In conclusion, this section has treated a unified mixed finite element method for the spatial discretization of the Maxwell and the Kelvin-Voigt models. We have given a priori error estimates for the spatial discretization under the assumptions of saddle point stability and (4.3). We have focused on these spatial error estimates in order to present the techniques and results involved with greater clarity.

5. FULL-DISCRETIZATION

The aim of this section is to consider full-discretizations of the Maxwell and Kelvin-Voigt equations. The starting points are the semi-discrete formulations (4.8) and (4.14). After spatial discretization, the systems can be viewed as linear constant coefficient differential-algebraic equations (DAEs) in time. The theory of linear constant coefficient DAEs is well-developed and the monographs [**12, 26**] give a thorough summary. As this is the case, we shall not carry out an explicit analysis for different temporal discretization schemes, but rather rely on the known results for DAEs. However, we point out that it is not obvious that these time discretizations will be uniformly stable with respect to the spatial discretizations.

With regard to choice of time-discretization, there are some factors that deserve special attention. First, since the material parameters for the viscous and the elastic contributions may vary greatly, we face possibly stiff systems. Second, in thread with the previous emphasis on robustness, we aim to avoid stability conditions for the discretization parameters. These aspects make implicit time-stepping schemes attractive. Also note that since we can use an arbitrary, up to computation time, high order scheme in space, higher order schemes in time are relevant.

We shall briefly summarize the relevant concepts of and results for linear constant coefficient DAEs needed in the following, but refer to [**12**] for details. Linear constant coefficient DAEs take the form

$$\mathbf{D_1}\dot{\mathbf{y}} + \mathbf{D_0}\mathbf{y} = \mathbf{f} \tag{5.1}$$

where $\mathbf{D_i} \in \mathbb{M}^{n\times n}$ for $i = 0, 1$. Such systems are *solvable* if and only if $\lambda\mathbf{D_1} + \mathbf{D_0}$ is a *regular* pencil, i.e. if $\det(\lambda\mathbf{D_1} + \mathbf{D_0})$ is not identically zero as a function of λ. Further, assume that the system in (5.1) is of *index* ν [**12**, p. 17]. A k-step implicit backward difference scheme of local order k applied to (5.1) gives convergence of order k after a possible initial boundary layer of thickness $(\nu - 1)k + 1$ time steps [**12**, Theorem 3.1.1]. For implicit one-step Runge-Kutta methods, matters are a bit more elaborate, and we shall again refer the reader to [**12**, Section 4]. In the following, we shall see how the stability conditions for the spatial discretization of the Maxwell and Kelvin-Voigt systems yields regularity and thus solvability in the DAE sense.

Consider the spatial discretizations (4.8) and (4.14) of the Maxwell and the Kelvin-Voigt equations respectively. Inserting bases for the element spaces Σ_h, V_h and Q_h, letting $y_h = (\sigma_{j,h}, v_h, \gamma_h)$ and denoting the corresponding vector of finite element expansion coefficients by $\mathbf{y}$, yield following form of the equations:

$$\mathbf{D_1}\dot{\mathbf{y}} + \mathbf{D_0}\mathbf{y} = \mathbf{f}.$$

In particular, let $\{\psi_i\}_i$, $\{\phi_k\}_k$ and $\{p_m\}_m$ be bases for Σ_h, V_h and Q_h respectively and define the element matrices

$$\mathbf{A}_{\iota ij} = \langle A_\iota \psi_j, \psi_i \rangle, \quad \mathbf{B}_{kj} = \langle \operatorname{div} \psi_j, \phi_k \rangle, \quad \mathbf{C}_{mj} = \langle \psi_j, p_m \rangle.$$

The spatial discretization of the Maxwell model takes the form

$$\mathbf{D_1} = \begin{pmatrix} \mathbf{A_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \quad \mathbf{D_0} = \begin{pmatrix} \mathbf{A_0} & \mathbf{B^T} & \mathbf{C^T} \\ \mathbf{B} & \mathbf{0} & \mathbf{0} \\ \mathbf{C} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{5.2}$$

while for the Kelvin-Voigt model, we have:

$$\mathbf{D_1} = \begin{pmatrix} \mathbf{A_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \quad \mathbf{D_0} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{B^T} & \mathbf{C^T} \\ \mathbf{0} & \mathbf{A_0} & \mathbf{B^T} & \mathbf{C^T} \\ \mathbf{B} & \mathbf{B} & \mathbf{0} & \mathbf{0} \\ \mathbf{C} & \mathbf{C} & \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{5.3}$$

The Maxwell system is a differential algebraic equation of index two, while the Kelvin-Voigt system in its current form is of index one.

The discrete inf-sup conditions (4.1), (4.2) guarantee that the matrices $\mathbf{A_0}$, $\mathbf{A_1}$ and $\mathbf{B}, \mathbf{C}$ are such that $\det(\lambda \mathbf{D_1} + \mathbf{D_0}) \not\equiv 0$ as a function of λ for either system. Hence, with the spatial discretizations of Section 4, the systems will be regular, and therefore solvable. The already well-developed theory of discretizations of DAEs therefore carries over to these systems.

6. Numerical experiments

We now turn to consider a series of numerical experiments for both models in two dimensions. We shall focus on two aspects. First, we demonstrate convergence rates for different material parameter values on the unit square with constructed exact solutions. Second, we give examples of viscoelastic behaviour through looking at simple creep and relaxation scenarios. All simulations have been run using the DOLFIN library [**27**] from the FEniCS project [**22**].

The spatial discretization is based on the family of elasticity elements defined by (4.6), and its variant (4.7). The temporal discretization is carried out using the L-stable, second-order accurate, TR-BDF$_2$ scheme, that is, a trapezoidal rule followed by a 2-step backward difference scheme at each time step [**9**]. For the general DAE (5.1), this scheme takes the following form for $n = 0, 1, \ldots$:

$$\mathbf{D_1}\left(\mathbf{y}^{n+\frac{1}{2}} - \mathbf{y}^n\right) = \frac{1}{2}\frac{\Delta t}{2}\left(\mathbf{f}^{n+\frac{1}{2}} + \mathbf{f}^n - \mathbf{D_0}\left(\mathbf{y}^{n+\frac{1}{2}} + \mathbf{y}^n\right)\right) \tag{6.1a}$$

$$\left(\frac{\Delta t}{2}\right)^{-1} \mathbf{D_1}\left(\frac{3}{2}\mathbf{y}^{n+1} - 2\mathbf{y}^{n+\frac{1}{2}} + \frac{1}{2}\mathbf{y}^n\right) + \mathbf{D_0}\mathbf{y}^{n+1} = \mathbf{f}^{n+1} \tag{6.1b}$$

The linear systems of equations that result from the application of this scheme to $\mathbf{D_0}, \mathbf{D_1}$ defined by (5.2) and (5.3), share a common matrix structure and again the discrete inf-sup conditions guarantee the solvability at each time step for any fixed $\Delta t > 0$. On the other hand, it is not immediately evident that the time discretizations are stable uniformly in h. However, there are no indications in the following convergence experiments of that stability not holding uniformly.

6.1. Convergence. The families of finite element spaces defined by (4.6) and (4.7) were suggested and analyzed for the elasticity equations by Arnold et al [**6**] and Falk [**20**] respectively. As these have not been widely used for numerical simulations, we commence by examining their convergence properties for the elasticity equations separately. We continue by considering the finite element spaces (4.6) with $k = 2$ in space and the scheme (6.1) in time for the Maxwell and Kelvin-Voigt models.

6.1.1. *Elasticity.* Consider a discretization of the isotropic elasticity equations (3.6) using the element spaces $\Sigma_{h,k}, V_{h,k}$ and $Q_{h,k}$ as defined by (4.6). We let $\Omega = [0,1]^2$ and consider a uniform, regular triangulation of Ω. We solve for the following displacement field over Ω:

$$u(x_0, x_1) = \begin{pmatrix} -x_1 \sin(\pi x_0) \\ \frac{1}{2}\pi x_1^2 \cos(\pi x_0) \end{pmatrix} \tag{6.2}$$

with corresponding elastic stress $\sigma = 2\mu\varepsilon(u) + \lambda \operatorname{div} u I$ and force $g = \operatorname{div} \sigma$. Observe that u, as defined by (6.2), is divergence-free and hence that the stress σ and force g are independent of λ. The convergence rates for the displacement and stress approximations using $\mu = 1$, $\lambda = 100$, $k = 1, 2, 3$ are given in Table 1 and are in agreement with the predicted rates.

k	$\log_2 \frac{\|\|u-u_h\|\|_0}{\|\|u-u_{h/2}\|\|_0}$	$\log_2 \frac{\|\|\sigma-\sigma_h\|\|_0}{\|\|\sigma-\sigma_{h/2}\|\|_0}$	$\log_2 \frac{\|\|\sigma-\sigma_h\|\|_{\mathrm{div}}}{\|\|\sigma-\sigma_{h/2}\|\|_{\mathrm{div}}}$	$\log_2 \frac{\|\|\gamma-\gamma_h\|\|_0}{\|\|\gamma-\gamma_{h/2}\|\|_0}$
1	1	0.997	1	1.01
2	2	1.99	2	2
3	3	3	3	3

TABLE 1. Convergence rates for elasticity using $\Sigma_{h,k} \times V_{h,k} \times Q_{h,k}$, $h = (16)^{-1}$, $k = 1, 2, 3$, for $\lambda = 100$, $\mu = 1$. The exact displacement u is defined by (6.2).

Furthermore, if the displacement is small in comparison to the stress, improved convergence in the L^2-norm of stress has been observed. With the exact solution,

$$u(x_0, x_1) = \begin{pmatrix} \sin(\pi x_0) \sin(\pi x_1) \\ \sin(\pi x_0) \sin(\pi x_1) \end{pmatrix} \tag{6.3}$$

and $\mu = 1$, $\lambda = 1000$, the convergence rates for $k = 1$ are given in Table 2. Observe that the L^2 error rate of the stress is consistently close to 2, that is, of one order higher than predicted by the analysis. This might be attributed to the inconsistently high convergence rates for the rotation and also to the large difference in magnitude of the stress and the rotation; The stress is many orders larger than the rotation in this test case and so the error contribution from the rotation will be less significant.

h^{-1}	$\log_2 \frac{\|\|u-u_h\|\|_0}{\|\|u-u_{h/2}\|\|_0}$	$\log_2 \frac{\|\|\sigma-\sigma_h\|\|_0}{\|\|\sigma-\sigma_{h/2}\|\|_0}$	$\log_2 \frac{\|\|\sigma-\sigma_h\|\|_{\mathrm{div}}}{\|\|\sigma-\sigma_{h/2}\|\|_{\mathrm{div}}}$	$\log_2 \frac{\|\|\gamma-\gamma_h\|\|_0}{\|\|\gamma-\gamma_{h/2}\|\|_0}$
4	2.82	1.91	0.98	2.80
8	2.45	1.97	0.995	2.88
16	1.42	1.99	0.999	2.68
32	1.04	2.00	1.00	1.74

TABLE 2. Convergence rates for elasticity using $\Sigma_{h,1} \times V_{h,1} \times Q_{h,1}$, for $\lambda = 1000$, $\mu = 1$. The exact displacement u is defined by (6.3).

In order to demonstrate the convergence properties of the element spaces of Falk (4.7), we replace $Q_{h,k}$ by $\tilde{Q}_{h,k}$. Solving for the divergence free displacement of

(6.2) with $\mu = 1$, $\lambda = 100$, gives the convergence rates in Table 3. The anticipated improvement of the L^2 convergence for the stress and rotation is observed.

k	$\log_2 \frac{\|\|u-u_h\|\|_0}{\|\|u-u_{h/2}\|\|_0}$	$\log_2 \frac{\|\|\sigma-\sigma_h\|\|_0}{\|\|\sigma-\sigma_{h/2}\|\|_0}$	$\log_2 \frac{\|\|\sigma-\sigma_h\|\|_{\mathrm{div}}}{\|\|\sigma-\sigma_{h/2}\|\|_{\mathrm{div}}}$	$\log_2 \frac{\|\|\gamma-\gamma_h\|\|_0}{\|\|\gamma-\gamma_{h/2}\|\|_0}$
1	1	2	1	2
2	2	3	2	3.17
3	3	4.03	3.04	4.05

TABLE 3. Convergence rates for elasticity using $\Sigma_{h,k} \times V_{h,k} \times \tilde{Q}_{h,k}$, $h = (16)^{-1}$, $k = 1, 2, 3$, for $\lambda = 100$, $\mu = 1$. The exact displacement u is defined by (6.2).

6.1.2. *Maxwell.* We continue letting $\Omega = [0, 1]^2$, consider a regular triangulation $\mathcal{T}_h$ of Ω and the associated second order finite element spaces of the family (4.6): $\Sigma_{h,2}, V_{h,2}$ and $Q_{h,2}$. Moreover, we let $T = 1$ and consider a uniform partition of the time domain $[0, T]$ and the TR-BDF$_2$ difference scheme defined by (6.1). We shall let μ_0 dominates μ_1 and solve for a known smooth velocity in order to demonstrate robustness with regard to these parameters.

Define

$$v(x,t) = e^{t-1}((2\mu_1)^{-1}2t + (2\mu_0)^{-1}t^2) \begin{pmatrix} -x_1 \sin(\pi x_0) \\ \frac{1}{2}\pi x_1^2 \cos(\pi x_0) \end{pmatrix} \tag{6.4}$$

and let $\lambda_0 = \lambda_1 = 100$, $\mu_0 = 100$ and $\mu_1 = 1$. Observe that the consistency condition at the initial time is satisfied, in particular that $\sigma(x, t) = 0$ at $t = 0$. The convergence rates at $T = 1.0$ are given in Table 4. We conclude that there is no loss of convergence even though $\mu_0 >> \mu_1$.

h^{-1}	$\log_2 \frac{\|\|v-v_h\|\|_0}{\|\|v-v_{h/2}\|\|_0}$	$\log_2 \frac{\|\|\sigma-\sigma_h\|\|_0}{\|\|\sigma-\sigma_{h/2}\|\|_0}$	$\log_2 \frac{\|\|\sigma-\sigma_h\|\|_{\mathrm{div}}}{\|\|\sigma-\sigma_{h/2}\|\|_{\mathrm{div}}}$	$\log_2 \frac{\|\|\gamma-\gamma_h\|\|_0}{\|\|\gamma-\gamma_{h/2}\|\|_0}$
4	1.96	2.03	1.99	1.98
8	1.98	2	2	1.99
16	1.99	1.99	2	1.99

TABLE 4. Maxwell convergence rates for $\lambda_0 = \lambda_1 = 100$, $\mu_0 = 100$, $\mu_1 = 1$ at $T = 1$. $\Delta t = h$. Exact velocity v as defined by (6.4)

.

6.1.3. *Kelvin-Voigt.* For the Kelvin-Voigt model, we consider the previous domain in time and space, but solve for v given by (6.5).

$$v(x,t) = (2t + t^2)e^{t-1} \begin{pmatrix} \sin(\pi x_0)\sin(\pi x_1) \\ \sin(\pi x_0)\sin(\pi x_1) \end{pmatrix} \tag{6.5}$$

and enforce the initial condition $\sigma_1(x, 0) = 0$. The L^2 error of the velocity and rotation and $H(\mathrm{div})$ error of the stresses, are measured at each time step and included in Table 5. We observe that the convergence rates point-wise in time is of the order $\mathcal{O}(h^2)$.

6.2. Relaxation and creep. We now turn to illustrate the two main viscoelastic behaviour characteristics, creep and relaxation. The Maxwell model exhibits stress relaxation as response to non-zero displacement and the Kelvin-Voigt model exhibits creep as response to applied traction. Thus, we shall qualitatively demonstrate a relaxation pattern using the Maxwell model and a creep pattern with the Kelvin-Voigt model. However, the simplest model that predicts both adequate stress relaxation and creep, is the standard linear solid (2.10). For this reason, and

h^{-1}	$\log_2 \frac{\|\|v-v_h\|\|_0}{\|\|v-v_{h/2}\|\|_0}$	$\log_2 \frac{\|\|\sigma_0-\sigma_{0,h}\|\|_{\mathrm{div}}}{\|\|\sigma_0-\sigma_{0,h/2}\|\|_{\mathrm{div}}}$	$\log_2 \frac{\|\|\sigma_1-\sigma_{1,h}\|\|_{\mathrm{div}}}{\|\|\sigma_1-\sigma_{1,h/2}\|\|_{\mathrm{div}}}$	$\log_2 \frac{\|\|\gamma-\gamma_h\|\|_0}{\|\|\gamma-\gamma_{h/2}\|\|_0}$
4	1.98	1.98	1.93	2.14
8	1.99	1.99	1.97	2.06
16	2	2	1.98	2.02

TABLE 5. Kelvin-Voigt convergence rates for $\lambda_0 = \lambda_1 = 100$, $\mu_0 = 100$, $\mu_1 = 1$ at $T = 1.0$. $\Delta t = h$. Exact solutions corresponding to exact velocity defined by (6.5).

Maxwell	Kelvin-Voigt	Standard linear solid
$\mu_0 = 20.0$, $\lambda_0 = 100.0$	$\mu_0 = 20.0$, $\lambda_0 = 100.0$	$\mu_0^0 = 5.0$, $\lambda_0^0 = 100.0$
$\mu_1 = 80.0$, $\lambda_1 = 100.0$	$\mu_1 = 80.0$, $\lambda_1 = 100.0$	$\mu_1^0 = 30.0$, $\lambda_1^0 = 100.0$
		$\mu_1^1 = 20.0$, $\lambda_1^1 = 100.0$

TABLE 6. Parameters values used in the creep and relaxation simulations. The parameters $\mu_i^{(j)}$ and $\lambda_i^{(j)}$ correspond to the compliance tensors $A_i^{(j)}$ of the separate models.

in order to show the numerical method applied to a more general model, we also include experiments for these equations.

For the experiments, we consider the usual domain in space: $(x_0, x_1) \in \Omega = [0,1]^2$ and $t \in [0,T]$. We let the body be clamped at the left boundary $x_0 = 0$ and stress-free at the top and bottom $x_1 = 0, 1$. On the right boundary, a velocity is prescribed for the relaxation experiments and a prescribed traction is used for the creep experiments. The material parameters used are listed in Table 6.

6.2.1. *Relaxation.* The Maxwell and standard linear solid models exhibit stress relaxation as response to non-zero displacement. To illustrate this behaviour, we apply a unit boundary velocity in the x_0 direction, on the right boundary $x_0 = 1$ at a given time $t_{\min}$ and remove the velocity at $t_{\max}$. This produces a displacement, in particular a constant displacement for $t > t_{\max}$. The Maxwell model predicts that the stresses increase from $t_{\min}$ to $t_{\max}$ and after that decrease to the zero limit. For the standard linear solid model, the stresses will decrease toward an elastic limit for $t > t_{\max}$.

The displacement and stress approximations for the Maxwell model are plotted in Figure 1. Observe in particular the exponential relaxation of the stress for $t > t_{\max}$. The same experiment for the standard linear solid is plotted in Figure 2.

6.2.2. *Creep.* The Kelvin-Voigt and standard linear solid models predict creep as response to applied traction. To demonstrate this, we apply a unit boundary traction in the x_0 direction at the right boundary $x_0 = 1$ at a given time $t_{\min}$ and remove at $t_{\max}$. The $L^2(\Omega)$ norms of the approximated displacements and stresses are displayed in Figure 3 for the Kelvin-Voigt model and Figure 4 for the standard linear solid. Observe the initial increase towards the elastic limit and subsequent decrease for the displacement for the Kelvin-Voigt model. Also note the instantaneous displacements for the standard linear solid.

7. CONCLUSION

We have presented a unified mixed finite element framework for linear viscoelasticity. The framework relies on mixed finite element discretizations originating from linear elasticity. These discretizations have been evaluated by numerical experiments both for linear elasticity and viscoelasticity. The analytical predictions

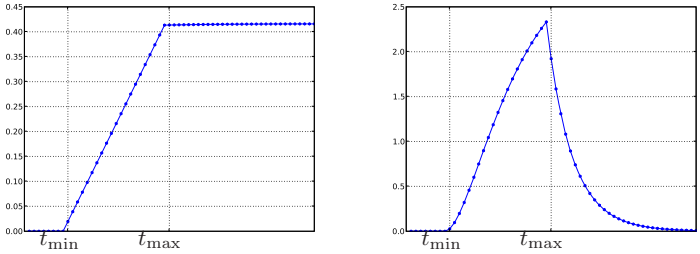


FIGURE 1. Maxwell response to applied velocity. Left: Displacement u versus time t. Right: Stress $||\sigma(t)||_{A_1}$ versus time t. Discretization parameters: $\Delta t = 1.0/30$, $h = 1.0/32$, $t_{\min} = 0.3$, $t_{\max} = 1.0$, $T = 2.0$.

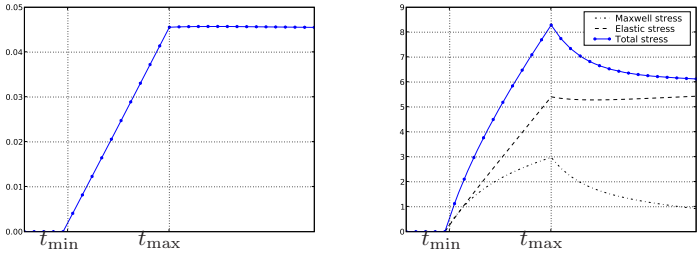


FIGURE 2. Standard linear solid response to applied velocity. Left: Displacement u versus time t. Right: Stresses $||\sigma_0(t)||$, $||\sigma_1(t)||$, and $||\sigma(t)||_0$ versus time t. Discretization parameters: $\Delta t = 1.0/30$, $h = 1.0/32$, $t_{\min} = 0.3$, $t_{\max} = 1.0$, $T = 2.0$. (The applied velocity is one tenth of the one applied for the Maxwell model.) The standard linear model can, in term of springs and dashpots, be viewed as a Maxwell element in parallel with an elastic spring. The corresponding components of the stress are therefore labelled *Maxwell* and *Elastic* according to this interpretation.

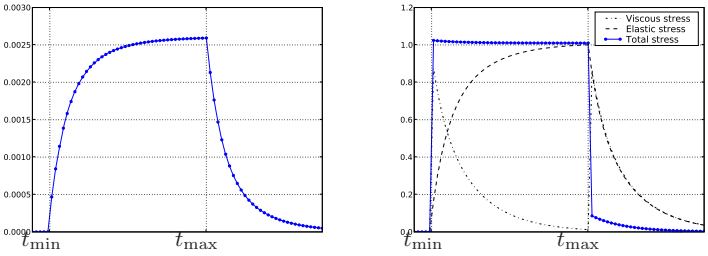


FIGURE 3. Kelvin-Voigt response to applied traction. Left: Displacement u versus time t. Right: Viscous and elastic stress, $||\sigma_0(t)||_0$ and $||\sigma_1(t)||_0$ versus time t. Parameters: $\Delta t = 1.0/30$, $h = 1.0/32$, $t_{\min} = 0.3$, $t_{\max} = 3.0$, $T = 5.0$.

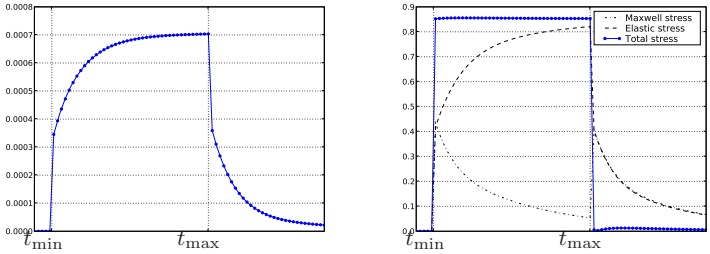


FIGURE 4. Standard linear solid response to applied traction. Left: Displacement u versus time t. Right: Maxwell, elastic and total stress, $||\sigma_0(t)||_0$, $||\sigma_1(t)||_0$ and $||\sigma(t)||_0$ versus time t. Parameters: $\Delta t = 1.0/30$, $h = 1.0/32$, $t_{\min} = 0.3$, $t_{\max} = 3.0$, $T = 5.0$. (The applied traction is ten times the one applied for the Kelvin-Voigt model.)

are confirmed and the anticipated robustness with regard to material parameters demonstrated.

Acknowledgments. The authors acknowledge and are grateful for the initiative and valuable advice provided by Hans Petter Langtangen.

CENTRE OF MATHEMATICS FOR APPLICATIONS, P.O.BOX 1053, 0316 OSLO, NORWAY. meg@cma.uio.no

CENTRE OF MATHEMATICS FOR APPLICATIONS, P.O.BOX 1053, 0316 OSLO, NORWAY. ragnar.winther@cma.uio.no

Bibliography

[1] S. Adams and B. Cockburn. A mixed finite element method for elasticity in three dimensions. *J. Sci. Comput.*, 25:515–521, 2005.

[2] K. Adolfsson, M. Enelund, S. Larsson, and M. Racheva. Discretization of integro-differential equations modeling dynamic fractional order viscoelasticity. In *Proceedings of Large-Scale Scientific Computations, 2005, Sozopol, Bulgaria*, pages 76–83. Springer, 2006.

[3] M. Amara and J. M. Thomas. Equilibrium finite elements for the linear elastic problem. *Numer. Math.*, 33:367–383, 1979.

[4] D. N. Arnold, G. Awanou, and R. Winther. Finite elements for symmetric tensors in three dimensions. *Math. Computation*, 77:1229–1251, 2008.

[5] D. N. Arnold, F. Brezzi, and J. Douglas. PEERS: A new mixed finite element for plane elasticity. *Japan J. Appl. Math.*, 1:347–367, 1984.

[6] D. N. Arnold, R. Falk, and R. Winther. Mixed finite element methods for linear elasticity with weakly imposed symmetry. *Math. Computation*, 76:1699–1723, 2007.

[7] D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus, homological techniques and applications. *Acta numerica*, 2006.

[8] D. N. Arnold and R. Winther. Mixed finite elements for elasticity. *Numer. Math.*, 92:401–419, 2002.

[9] R. Bank, W. Coughran, W. Fichtner, E. Grosse, D. Rose, and R. Smith. Transient simulation of silicon devices and circuits. *IEEE Transactions on Computer-Aided Design on Integrated Circuits and Systems*, 4:436–451, 1985.

[10] E. Becache, A. Ezziani, and P. Joly. A mixed finite element appproach for viscoelastic wave propagation. *Computational Geosciences*, 8:255–299, 2004.

[11] D. Braess. *Finite elements. Theory, fast solvers, and applications in solid mechanics*. Cambridge University Press, second edition, 2001.

[12] K. E. Brenan, S. L. Campbell, and L. R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. SIAM, second edition, 1996.

[13] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *R.A.I.R.O. Anal. Numer.*, 2:129–151, 1974.

[14] F. Brezzi, J. Douglas, and L. D. Marini. Two families of mixed elements for second order elliptic problems. *Numer. Math.*, 47:217–235, 1985.

[15] S. H. Christiansen and R. Winther. Smoothed projections in finite element exterior calculus. *Math. computation*, 77:813–829, 2008.

[16] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology*, volume 1. Springer, 1990.

[17] A. D. Drozdov. *Mechanics of Viscoelastic Solids*. Wiley, 1998.

[18] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 2002.

[19] M. Fabrizio and A. Morro. *Mathematical Problems in Linear Viscoelasticity*. SIAM, 1992.

[20] R. S. Falk. Finite element methods for linear elasticity. In *Mixed Finite Elements, Compatibility Conditions and Applications*. Springer, 2008.

[21] M. Farhloul and M. Fortin. Dual hybrid methods for the elasticity and the stokes problem: a unified approach. *Numer. Math.*, 76:417–440, 1997.

[22] FEniCS. FEniCS project. URL: http//www.fenics.org/.

[23] J. D. Ferry. *Viscoelastic Properties of Polymers*. John Wiley and Sons, 2002.

[24] B. M. Fraijs de Veubeke. Stress function approach. In *Proc. of the World Congress on Finite Element Methods in Structural Mechanics*, volume 1, pages J.1–J.51, 1975.
[25] M. E. Gurtin and E. Sternberg. On the linear theory of viscoelasticity. *Arch. Rational Mech. Anal*, 11:291–356, 1962.
[26] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*. Springer, second edition, 1996.
[27] J. Hoffman, J. Jansson, A. Logg, G. N. Wells, et al. DOLFIN. URL: http//www.fenics.org/wiki/DOLFIN.
[28] A. Idesman, R. Niekamp, and E. Stein. Finite elements in space and time for generalized viscoelastic Maxwell model. *Computational Mechanics*, 27:49–60, 2001.
[29] A. R. Johnon and A. Tessler. A viscoelastic higher-order beam finite element. In J. R. Whiteman, editor, *The Mathematics of Finite Elements and Applications*, pages 333–345. Wiley, 1997.
[30] P. Le Tallec and M. Ravachol. Finite element approximation of viscoelastic progressively incompressible flows. *Numerische Mathematik*, 56:43–66, 1989.
[31] J. C. Nédélec. A new family of mixed finite elements in $\mathbf{R}^3$. *Numer. Math.*, 50:57–81, 1986.
[32] N. Phan-Thien. *Understanding Viscoelasticity. Basics of Rheology*. Springer, 2002.
[33] B. Rivière, S. Shaw, and J. R. Whiteman. Discontinuous Galerkin finite element methods for linear elasticity and quasistatic linear viscoelasticity. *Numer. Math*, 95:347–376, 2003.
[34] B. Rivière, S. Shaw, and J. R. Whiteman. Discontinuous Galerkin finite element methods for dynamic linear solid viscoelasticity problems. *Numerical Methods for Partial Differential Equations*, 23:1149–1166, 2007.
[35] S. Shaw. *Finite element and discrete time methods for continuum problems with memory and applications to viscoelasticity*. PhD thesis, Department of Mathematics and Statistics, Brunel University, 1993.
[36] S. Shaw, M. K. Warby, J. R. Whiteman, C. Dawson, and M. F. Wheeler. Numerical techniques for the treatment of quasistatic solid viscoelastic stress problems. *Comput. Methods Appl. Mech. Engrg*, 118:211–237, 1994.
[37] S. Shaw and J. R. Whiteman. Numerical solution of linear quasistatic hereditary viscoelasticity problems. *Siam J. Numer. Anal.*, 38:80–97, 2000.
[38] S. Shaw and J. R. Whiteman. A posteriori error estimates for space-time finite element approximation of quasistatic hereditary linear viscoelasticity problems. *Comput. Methods Appl. Mech. engrg.*, 193:5551–5572, 2004.
[39] R. Stenberg. A family of mixed finite elements for the elasticity problem. *Numer. Math.*, 53:513–538, 1988.
[40] N. Tschoegl. *The Phenomenological theory of linear viscoelastic behaviour*. Springer, first edition, 1989.

Paper II

MIXED FINITE ELEMENT METHODS FOR GELS WITH BIOMEDICAL APPLICATIONS

MARIE E. ROGNES, M. CARME CALDERER AND CATHERINE A. MICEK

Abstract. A set of equilibrium equations for a biphasic polymer gel are considered with the end purpose of studying stress and deformation in confinement problems encountered in connection with biomedical implants. The existence of minimizers for the gel energy is established first. Further, the small-strain equations are derived and related to the linear elasticity equations with parameters dependent on the elasticity of the polymer and the mixing of the polymer and solvent. Two numerical methods are considered, namely a two-field displacement-pressure formulation and a three-field stress-displacement formulation with weak symmetry. The symmetry of the stress tensor is affected by the residual stress induced by the polymer-solvent mixing. A novel variation of the stress-displacement formulation of linear elasticity with weak symmetry is therefore proposed and analyzed. Finally, the numerical methods are used to simulate the stresses arising in a confined gel implant.

1. Introduction

Since the development and commercialization of the pacemaker in 1957, a wide variety of biomedical devices have been designed to address chronic health conditions. In addition to the pacemaker, other examples of *body implantable devices* include artificial bone tissue and cardiovascular stents. When designing such devices, their long-term use must be taken into account. Many of these devices are made of synthetic polymers, and, upon insertion in the human body, the absorption of moisture can cause the polymer to swell. The combination of swelling and confinement causes stresses that may compromise the intended life-time of the device. For example, in devices such as artificial bones that are attached with a chemical glue, the stress buildup along the attachment can loosen the glue and destroy device performance. Hence, the impact of the body environment on a device is an important aspect of body implantable devices. In particular, the ability to predict the resulting stresses may be crucial to effective design.

Polymers that have absorbed moisture classify as gels, and many of the biomedical devices use polymers in a gel form. The term *gel* refers to cross-linked polymer networks or entanglements enclosing a solvent such as water. The two components of the gel coexist by balancing each other. The polymer network confines

the liquid solvent, and the solvent ensures that the polymer does not collapse into a dry state. Thus, gels typically have the appearance of solids and may in many situations behave as such. However, the pioneering work of Tanaka in the 1970s revealed some surprising physical properties [**24**]. One such is the ability of the gel to display phase transitions. A gel phase transition is a finite change in volume of the polymer network – either a large expansion or contraction – that is induced by an infinitesimal change in the external environment. This and other properties of gels give an immensely rich array of applications [**21**, Vol 3].

Due to their complex physical properties, the mathematical modelling of gels is a non-trivial task. Tanaka identified the following three forces acting on the polymer network and labelled their sum the osmotic pressure: the elasticity of the polymer network, the interaction of the polymer and solvent (also called the polymer-polymer affinity), and the ionization of the polymer network [**24**]. For the dynamics of gels, additional mechanisms, such as the diffusion of solvent and ions, the dynamics of the polymer network, and the dissipation of energy must be accounted for. The modelling of gel dynamics has been an active field since the 1970s. Recent approaches in gel modelling include [**8, 11, 19**]. Also, the mechanics and modelling of biphasic soft tissue share common features with that of polymer gels. Mathematical models for gels may comprise fields such as deformation of the polymer network, velocities of polymer and solvent, pressures, stresses, and electrical and chemical concentrations. Clearly, the dynamical equations become a complicated system. Numerical methods are therefore essential for the solution and simulation of these equations.

The wide range of different modelling approaches makes direct comparison of simulation approaches challenging, but it is possible to identify some general trends. The flexibility of finite element methods with regard to complicated domain geometries has made it a natural choice for numerical simulations. In particular, mixed finite element methods approximating the velocities and pressure, possibly in combination with other variables, such as in [**22**], have been widespread. However, numerical stability considerations with respect to the material parameters and the mixed finite element spaces are rare in the literature. Some exceptions include the comparison of stabilized and stable velocity-pressure formulations for the deformation of biphasic soft tissue by Almeida and Spilker [**1**] and the study of an extended finite element method for the transitional interface dynamics of hydrogels by Dolbow et al. [**12**]. In this paper, the stability of the numerical methods with reference to the material parameters is carefully considered.

For biomedical devices such as artificial bone implants, the main question is the equilibrium state of the gel and the stress distribution in the gel and at interfaces. Accordingly, in this paper, we focus on the static equilibrium problem of non-ionic gels. The starting-point is taken from the equilibrium equations of [**8**], based on balance laws and the theory of mixtures, which again can be related to the series of papers of Doi et al. [**26, 27, 28**] and the classical work of Flory [**16**]. The polymer network and polymer-solvent interaction contributions to the osmotic pressure are accounted for by the combination of an elastic and a mixing energy. Similar equations were treated by a deformation-based finite element method by Hong et al. [**18**]. Further, for this type of applications, the deformations are typically small and thus the linearized regime may be relevant. Consequently, we shall derive the small strain gel equilibrium equations. These reduce to the equations of linear elasticity, with Lamé coefficients additionally depending on the volume fraction of the polymer and the mixing energy. As the theory of finite element and mixed finite element methods for linear elasticity is very well-studied, we refer to [**13**]

for a recent survey. Since the stress is of the main interest, we consider two numerical approaches, offering higher order stress approximations, though in different ways. Both a two-field displacement-pressure and a three-field stress-displacement-rotation mixed finite element method are presented and subsequently invoked.

To our knowledge, there are very few, if any, other works having studied mixed finite element methods of the stress-displacement-rotation type in the biphasic gel setting. The presence of the additional mixing energy gives additional features when compared to the purely elastic case. In particular, the state in which the residual stress vanishes is less intrinsic. The effect of non-vanishing residual stress due to non-equilibrium fractions of polymer is of direct physical interest, and a key aspect of this work. Both the displacement-pressure and stress-displacement-rotation methods we present allow for the possibility of non-vanishing residual stress. However, if the residual stress does not vanish, the linearized first Piola-Kirchhoff stress tensor is not symmetric, and thus the standard symmetry constraint of the stress tensor does not apply. Therefore, we propose a strategy for extending the weak symmetry stress-displacement-rotation formulation to the case where the stress tensor is not symmetric. The new weak formulation can be viewed as a perturbation of the stress-displacement-rotation formulation, and we characterize the stability of this formulation with respect to the elastic and mixing parameters.

The organization of this paper is as follows. The governing equations of a gel are reviewed and the free energy assumptions are discussed in Section 2. The minimization of the energy problem and the existence of minimizers are considered in Section 3. In Section 4, the Euler-Lagrange equations are derived and linearizations of these stated for further study in Section 5. There, we first discuss the stability of the resulting displacement-pressure and the stress-displacement-rotation formulations with regard to the gel material parameters. The main component of Section 5 though is the derivation and analysis of a novel variation of the stress-displacement-rotation formulation. In Section 6, we investigate the new formulation numerically, and then provide simulations of a confined polymer gel representing an artificial bone implant.

2. Governing equations of a gel

This paper focuses on the static equilibrium equations of a biphasic gel. In order to frame these equations and to introduce notation, we begin by presenting a brief survey of the governing equations of a gel and their derivation. This material is mostly classical, and we refer to [**21, 25**] and [**8, 9**] for a more thorough exposition. Under certain additional assumptions, the model reduces to the stress-diffusion coupling model proposed by Yamaue and Doi [**27**]. Next, assumptions regarding the free energy of the gel, and hence the constitutive equations, are discussed. Adopting the viewpoint of Flory [**16**] and subsequent modelling approaches [**8**], we consider the sum of an elastic and a Flory-Huggins energy. These two terms account for the elasticity of the gel and the mixing of the polymer and solvent, respectively.

2.1. Governing equations of a biphasic gel. Following the approach in [**8**], we consider the gel to be an immiscible, incompressible mixture of polymer and solvent. In an immiscible polymer-solvent mixture, the constitutive equations explicitly depend on the volume fractions ϕ_1 and ϕ_2 of the polymer and solvent, respectively. The volume fraction ϕ_i of each component is canonically defined as the volume of the component per unit volume of the gel. We denote by ρ_i the *mass* density of component i per unit volume in space for $i = 1, 2$. These are related to the *intrinsic* densities γ_i, by $\rho_i = \gamma_i \phi_i$. A mixture is incompressible if the intrinsic

density of each component is constant. Without loss of generality, we take the intrinsic densities of polymer and solvent to be equal to 1, and thus the mass density and volume fraction of each component coincide. We remark that the incompressibility of the mixture does not preclude deformations with change of volume. This point will be discussed further. Finally, we assume that there are neither voids nor additional material components in the system.

Let $\Omega \subset \mathbb{R}^n$ $(n = 1, 2, 3)$ describe a reference domain, with coordinates X, and assume that Ω is open and bounded with boundary $\partial\Omega$. Let

$$x : \Omega \to \Omega_x,$$

denote a smooth deformation map satisfying $\det \nabla x > 0$. Throughout this work, $F = \nabla x$ denotes the gradient of deformation. The divergence operator, taken rowwise when applied to matrices, is labelled div, $\frac{\partial}{\partial t}$ denotes the time derivative. These assumptions, together with the laws of balance of mass and linear momentum, and the second law of thermodynamics, yield the following set of governing equations over the Eulerian frame, that is, over the domain Ω_x.

$$\phi_1 + \phi_2 = 1, \tag{2.1a}$$

$$\frac{\partial \phi_i}{\partial t} + (v_i \cdot \nabla)\phi_i + \phi_i \operatorname{div} v_i = 0 \quad i = 1, 2, \tag{2.1b}$$

$$\phi_i \frac{\partial v_i}{\partial t} + \phi_i (v_i \cdot \nabla) v_i = \operatorname{div} \mathcal{T}_i + f_i \quad i = 1, 2, \tag{2.1c}$$

$$f_1 = \pi \nabla \phi_1 - \beta(v_1 - v_2) = -f_2. \tag{2.1d}$$

Here, for each component i, v_i is the velocity, $\mathcal{T}_i$ is the Cauchy stress tensor, and f_i is a force due to friction. Moreover, β is the polymer drag coefficient, and π is a Lagrange multiplier associated with the constraint (2.1a). In addition to (2.1), a kinematic compatibility condition between the time derivative of the deformation gradient and the velocity gradient results from the chain rule:

$$F_t + (v_1 \cdot \nabla)F = (\nabla v_1)F. \tag{2.2}$$

The governing system of equations becomes (2.1) combined with (2.2) and constitutive equations for the stress tensors, $\mathcal{T}_i$.

The boundary conditions for the system are formulated as mixed displacement-traction conditions for the gel and permeability conditions at the interface between the gel and a fluid environment. These can be represented as boundary conditions for the balance of linear momentum of the polymer and solvent, that is, (2.1c) for $i = 1, 2$, respectively. The derivation of these equations and more details can be found in [**8**].

The static equilibrium equations are obtained by setting $v_1 = v_2 = 0$. The total stress $\mathcal{T}$ is the sum of the component stresses $\mathcal{T}_i$. Furthermore, we assume that the polymer is elastic and the solvent is Newtonian. Hence, the polymer and solvent stresses only depend on the deformation and the rate of deformation, respectively. As a result, the equilibrium stress reduces to that of the polymer only, so $\mathcal{T} = \mathcal{T}_1$. The additional assumption that the polymer stress depends on the volume fraction of the polymer gives the relation $\mathcal{T}_1 = \mathcal{T}_1(F, \phi_1)$. The resulting problem is an elastic equilibrium problem, which may be more naturally handled in the reference configuration.

Introducing the first Piola-Kirchoff stress tensor $\mathcal{S} = \det(F)F^{-1}\mathcal{T}$, the equilibrium reference equation becomes:

$$\operatorname{div} \mathcal{S}(F, \phi_1) = 0 \quad \text{in } \Omega. \tag{2.3}$$

We have here, and will throughout, made a standard abuse of notation by labeling $\phi_1(X) := \phi_1 \circ x(X)$ for $X \in \Omega$. Next, assume that the body in the reference

configuration has a reference polymer volume fraction, $\phi_I : \Omega \to (0,1)$. Using the incompressibility assumption of the mixture and the resulting identification of the mass density and volume fraction, the balance of mass constraint of the polymer component reads

$$\int_\Omega \phi_I \,\mathrm{d}X = \int_{\Omega_x} \phi_1 \,\mathrm{d}x = \int_\Omega \phi_1 \,\det F \,\mathrm{d}X.$$

With the assumption that the above constraint is satisfied for all parts of the body Ω, we obtain the following local constraint:

$$\phi_1 \det F = \phi_I \quad \text{in } \Omega. \tag{2.4}$$

REMARK 2.1. *We emphasize that the balance of mass constraint* (2.4) *allows deformations with change in volume. In this regard, the mixture incompressibility assumption differs from classical incompressibility assumptions. Further, by definition,* ϕ_1 *must take values in* $[0,1]$. *This observation, in combination with* (2.4)*, yields the constraint* $\det F \geq \phi_1$.

2.2. Elastic and mixing energies. To provide a complete physical description of the gel, the governing equations in equilibrium, (2.3) and (2.4), must be augmented with appropriate constitutive equations for the polymer-solvent mixture. In this exposition, keeping with [**9**], we shall assume that the gel is hyperelastic and model the free energy for the gel as a sum of an elastic and a Flory-Huggins mixing energy [**16**]. This approach aims to account for the osmotic pressure contributions from the elasticity of the polymer network and the mixing of polymer and solvent. Furthermore, recall that the physical manifestation of gel phase transitions implies that there exists a critical set of physical parameters, corresponding to the onset of phase separation. Mathematically, this can be represented by a convexity transition of the free energy. The Flory-Huggins energy offers such a transition for given critical parameter values.

Due to (2.1a), the volume fraction of the solvent can be trivially eliminated in terms of the volume fraction of the polymer: $\phi_2 = 1 - \phi_1$. Hereafter, we omit the subscript and let $\phi = \phi_1$. The elastic energy per unit undeformed volume, denoted $\mathcal{W}_E$, is naturally formulated over the reference domain Ω. However, the Flory-Huggins energy, denoted $\mathcal{W}_{FH}$, is traditionally formulated per unit deformed volume [**16**]. This necessitates a change of variables on the Flory-Huggins energy. The resulting total energy $\mathcal{E}$ now follows:

$$\mathcal{E}(x,\phi) = \int_\Omega \mathcal{W}_E(F,\phi) + \det F \,(\mathcal{W}_{FH}(\phi) + c_{FH}) \,\mathrm{d}X, \tag{2.5}$$

where c_{FH} is a nonnegative scalar field on Ω. The term with coefficient c_{FH} penalizes growth of $(\det F)$, and thus growth of the physical gel volume. We postpone further discussion of this term momentarily, and instead provide some discussion of assumptions on the energy potentials $\mathcal{W}_E$ and $\mathcal{W}_{FH}$.

For elastic energy potentials, we restrict our attention to potentials of the form:

$$\mathcal{W}_E(F,\phi) = \phi_I \,\mathcal{W}_P(F), \tag{2.6}$$

where the potential is separable in its arguments and depends linearly on the reference volume fraction. We observe that, as a consequence of the local balance of mass (2.4), the dependency on the reference volume fraction in the energy per undeformed volume naturally translates to a dependency on the volume fraction in the energy per deformed volume. This restriction is thus in accordance with that of [**9**]. In Section 4 and onwards, we consider $\mathcal{W}_P$ taking the following, isotropic

form:

$$\mathcal{W}_P(F) = \mu_E \left(\frac{1}{2p} \left(||F||^{2p} - ||I||^{2p} \right) + \frac{||I||^{2(p-1)}}{\beta} ((\det F)^{-\beta} - 1) \right), \tag{2.7}$$

for $p \geq 1$, where $||F||^2 = \operatorname{tr}(FF^T) = F : F$, I is the identity matrix in $\mathbb{R}^{n \times n}$, μ_E is an elastic stiffness modulus, and β a parameter related to polymer compressibility. The constants are such that the identity state has zero energy and is also stress-free. This elastic energy density (2.7) has been proposed and studied in the context of compressible elasticity (see [**23**], and the discussion in [**10**, Section 4.10]). Note that the energy potential reduces to a compressible neo-Hookean potential when $p = 1$.

REMARK 2.2. *In view of Remark 2.1, the traditional interpretation of the term* $(\det F)^{-\beta}$ *does not apply within the mixture framework. In particular, the interpretation of the limiting case* $\beta \to \infty$ *is less clear. In the following, we shall view this term as a contraction penalty, ensuring that the domain does not degenerate under deformation.*

The Flory-Huggins mixing energy quantifies the energy available in the gel for polymer-solvent mixing and the energy per deformed volume reads:

$$\mathcal{W}_{FH}(\phi) = a\phi \ln \phi + b(1-\phi) \ln (1-\phi) + c\phi(1-\phi), \tag{2.8}$$

for positive parameters a, b, c. The parameters a, b, and c can be related to the specific polymer, solvent, and gel environment as follows [**8, 16**]:

$$a = \frac{K_B T}{V_m N_1}, \quad b = \frac{K_B T}{V_m N_2}, \quad c = \frac{K_B T \chi}{2 V_m}, \quad \chi = \frac{\Delta w}{K_B T} \tag{2.9}$$

where T is the absolute temperature, K_B is the Boltzmann constant, χ is the Flory interaction parameter, Δw is the change in energy per monomer-solvent interaction, V_m is the volume occupied by one monomer, and N_1 and N_2 are the number of lattice sites occupied by the polymer and solvent, respectively. The energy potential given by (2.8) is well-defined for $\phi \in [0, 1]$. In particular, $\mathcal{W}_{FH}(0) = \mathcal{W}_{FH}(1) = 0$, and $\mathcal{W}_{FH}$ is smooth for $\phi \in (0, 1)$. We note, however, that the energy given by (2.8) can take both positive and negative values. In order to ensure a lower bound of the total energy (2.5), we shall here introduce a lifting of the Flory-Huggins energy potential by letting $c_{FH} = c_{FH}(a, b, c) = -\inf_{0<\phi<1} \mathcal{W}_{FH}(\phi)$. The lifting parameter c_{FH} is clearly bounded in terms of the other Flory-Huggins parameters: $0 \leq c_{FH} \leq (a+b)$. The change of variables from deformed to reference configuration accounts for the $(\det F)$ factor in (2.5).

The convexity of $\mathcal{W}_{FH}$, or the lack of such, plays a central role in the following analysis. Changes in the convexity of $\mathcal{W}_{FH}$ may impact the polyconvexity of the total energy potential. Furthermore, changes in the polyconvexity of the total energy potential impact the existence of energy minimizers. The convexity of $\mathcal{W}_{FH}$ depends on the values of the coefficients a, b and c, and, in particular, on the interaction parameter χ. It can easily be seen that W_{FH} will be convex at ϕ if a, b and c are such that

$$\frac{a}{\phi} + \frac{b}{1-\phi} \geq 2c.$$

Our final observation is that the constraint (2.4) allows for the elimination of ϕ in terms of $\det F$. This substitution reduces the total free energy to an entirely mechanical energy with the following potential:

$$\mathcal{E}(x) = \int_\Omega \mathcal{W}_E(F, \phi_I (\det F)^{-1}) + (\det F) \left(\mathcal{W}_{FH}(\phi_I (\det F)^{-1}) + c_{FH} \right) \mathrm{d}X. \tag{2.10}$$

The problem of minimizing (2.10) can be viewed as a non-linear, compressible elasticity problem. One advantage of this formulation is that known techniques provide a framework for its mathematical and numerical analysis.

3. Existence of minimizing deformations

The physical deformation x of a gel defined over the domain Ω will be a minimizer in some admissible function space of the energy (2.10) augmented by body and boundary forces, if such a minimizer exists. This section, in particular the main theorem, Theorem 3.2, treats the existence of energy minimizers in the polyconvex case. A remark on a strategy for the non-convex case closes the section.

We begin by introducing some notation in order to define the space of admissible functions and for later use.

- Let Ω be an open, bounded subset of $\mathbb{R}^n$, with Lipschitz boundary $\partial\Omega$, and let $\partial\Omega = \partial\Omega_0 \cup \partial\Omega_1$ with $\partial\Omega_0 \cap \partial\Omega_1 = \emptyset$. Assume that $\partial\Omega_0$ has positive measure. The unit outward normal on $\partial\Omega$ is denoted n.
- We use the linear spaces of n vectors $\mathbb{V}$, $n \times n$ matrices $\mathbb{M}$, symmetric matrices $\mathbb{S}$ and skew-symmetric matrices $\mathbb{K}$. The inner product on M is denoted : and $||\cdot||$ is the Frobenius norm. The adjugate of an invertible matrix F is denoted $\operatorname{adj} F$, that is, $\operatorname{adj} F = (\det F)F^{-1}$.
- Let $\{\iota_1, \iota_2, \iota_3\}$ denote the invariants of the left Cauchy-Green strain tensor FF^T. Recall that $\iota_1 = \operatorname{tr}(FF^T)$, $\iota_2 = \operatorname{tr}\operatorname{adj}(FF^T)$ and $\iota_3 = \det(FF^T)$.
- The space of p-integrable fields on Ω with values in X is denoted $L^p(\Omega; X)$ with inner product $\langle\cdot,\cdot\rangle$ and norm $||\cdot||_0$. For notational ease, we will frequently omit the domain and range.
- $W^{k,p}(\Omega;\mathbb{V})$ denotes the Sobolev space of vector fields in $L^p(\Omega;\mathbb{V})$ such that the k'th derivatives exist in the distributional sense and belong to L^p. The associated norm is denoted by $||\cdot||_{k,p}$. For $p = 2$, we use the standard abbreviation $H^k = W^{k,2}$ with norm $||\cdot||_k$. Further, we let
$$W_0^{1,p}(\Omega;\mathbb{V}) = \{x \in W^{1,p}(\Omega;\mathbb{V}),\, x = 0 \text{ on } \partial\Omega_0\}.$$
Also, $H_0^1 = W_0^{1,2}$.
- The space of square integrable matrix fields with square integrable divergence (taken row-wise) is denoted $H(\operatorname{div}, \Omega; \mathbb{M})$. We also have the constrained space:
$$H_0(\operatorname{div}, \Omega; \mathbb{M}) = \{\sigma \in H(\operatorname{div}, \Omega; \mathbb{M}),\, \sigma \cdot n|_{\partial\Omega_1} = 0\}.$$

Assume that $\phi_I \in L^\infty(\Omega; (0,1))$. For $2p > n$ and a prescribed boundary condition $x_0 \in W^{1,2p}(\Omega)$, we define the space of admissible vector fields $\mathcal{A}$ as follows:

$$\mathcal{A} = \{x : \Omega \to \mathbb{R}^n, x \in x_0 + W_0^{1,2p},\, \operatorname{adj}\nabla x \in L^{2q}(\Omega),\, \det\nabla x \geq \phi_I \text{ a.e. in } \Omega\}.$$

We shall assume that the space $\mathcal{A}$ is non-empty; more specifically, that the constraint on the deformation gradient determinant can be fulfilled under the given boundary condition x_0. Now consider the following isotropic energy density:

$$\mathcal{E}(x) = \int_\Omega \mathcal{W}(F)\,\mathrm{d}X = \int_\Omega \mathcal{G}(\iota_1,\iota_2) + \mathcal{H}(\iota_3)\,\mathrm{d}X. \tag{3.1}$$

Clearly, the gel energy defined by (2.10) can be expressed as a special case of (3.1), using the following identifications:

$$\mathcal{G}(\iota_1,\iota_2) = \frac{\mu_E}{2p}\left(\iota_1^p - ||I||^{2p}\right), \tag{3.2}$$

$$\mathcal{H}(\iota_3) = \begin{cases} \iota_3\left(\mathcal{W}_{FH}(\phi_I\iota_3^{-1}) + c_{FH}\right) + C(\iota_3^{-\beta} - 1), & \phi_I < \iota_3 \\ +\infty, & \text{otherwise} \end{cases}, \tag{3.3}$$

where $C = C(\beta, \mu_E, \phi_I, p)$ is (implicitly) defined by (2.7) and the Flory-Huggins energy has been extended to the real line.

REMARK 3.1. *If $\mathcal{W}_{FH}$ is convex, $\mathcal{H}''(\iota_3) > 0$ for all $\iota_3 \geq \phi_I$. However, $\mathcal{H}''(\iota_3) < 0$ for some $\iota_3 \geq \phi_I$, if the interaction parameter χ, defined in* (2.9)*, is sufficiently large.*

Moreover, we assume that there exist constants $A_1, A_2 > 0, B \geq 0$ (or, if $\mathcal{G}(\iota_1, \iota_2) = \mathcal{G}(\iota_1)$, $A_2 \geq 0$) and $p > \frac{n}{2}$, $q \geq \frac{p}{2p-1}$ such that $\mathcal{G}(\iota_1, \iota_2)$ satisfies the growth condition

$$\mathcal{G}(\iota_1, \iota_2) \geq A_1 \iota_1^p + A_2 \iota_2{}^q - B. \tag{3.4}$$

Clearly, (3.2) satisfies (3.4) with $A_2 = 0$, if $2p > n$ and $\mu_E > 0$. Note that the neo-Hookean potential, corresponding to $p = 1$, does not satisfy the growth condition for $n \geq 2$.

The existence of minimizers of the energy (3.1), with $\mathcal{H}$ convex, over the space of admissible vector fields $\mathcal{A}$ is established in Theorem (3.2) below. We note that the assumption that $\mathcal{G}$ and $\mathcal{H}$ are convex give the polyconvexity of $\mathcal{W}$.

THEOREM 3.2. *Assume that $\mathcal{G}(\cdot, \cdot)$ is convex, satisfies the growth inequality* (3.4)*, and that $\mathcal{H}(\cdot)$ is convex. Suppose that $x_0 \in W^{1,2p}$ and that $\det \nabla x_0 > \phi_I$. Then, there exists $x \in \mathcal{A}$ that minimizes the energy* (3.1)*.*

PROOF. We first observe that, by the construction of c_{FH}, there exists an $m > -\infty$ such that $m \equiv \inf_{\mathcal{A}} \mathcal{E}(x)$. Let $\{x_k\}_{k \geq 1} \in \mathcal{A}$ denote a minimizing sequence of $\mathcal{E}$, that is, $\{x_k\}$ has the property $\lim_{k \to \infty} \mathcal{E}(x_k) = m$. To show that the problem of minimizing $\mathcal{E}$ in the set $\mathcal{A}$ has a solution, we proceed along the following steps.
Step 1. Prove that there exists a subsequence of $\{x_k\}$ (still denoted by x_k) such that

$$\lim_{k \to \infty} x_k = \bar{x}, \quad \text{weakly in } W^{1,2p}.$$

Step 2. Show that $\mathcal{E}$ is weakly lower semicontinuous in $W^{1,2p}$, that is, for any minimizing sequence $\{x_k\} \in \mathcal{A}$,

$$\lim_{k \to \infty} \inf \mathcal{E}(x_k) \geq m. \tag{3.5}$$

Step 3. Show that $\bar{x} \in \mathcal{A}$. This combined with the weak lower semicontinuity property (3.5) allows us to conclude that

$$\min_{y \in \mathcal{A}} \mathcal{E}(y) = \mathcal{E}(\bar{x}).$$

We outline steps 1 and 2 following [**10**, Theorem 7.7.1] and [**4**]. First, we find a lower bound of the energy. It follows from (3.4) and the form of $\mathcal{H}$, that there exists a constant γ such that

$$\mathcal{E}(x) \geq \int_\Omega A_1 \| \nabla x \|^{2p} + A_2 \| \operatorname{adj} \nabla x \|^{2q} + A_3 (\det \nabla x) \, \mathrm{d}X + \gamma |\Omega|,$$

for all $x \in \mathcal{A}$. An application of the generalized Poincaré inequality [**10**, Theorem 6.1-8(b)] allows us to conclude that there exist constants $a_1 > 0$ and a_2 such that

$$\mathcal{E}(x) \geq a_1 \left(\|x\|_{1,2p}^{2p} + \| \operatorname{adj} \nabla x \|_{0,2q}^{2q} \right) + a_2, \tag{3.6}$$

for all $x \in \mathcal{A}$. In particular, it follows from (3.6) that the the sequence $\{x_k, \operatorname{adj} \nabla x_k\}$ is bounded in $W^{1,2p} \times L^{2q}$. By the assumptions on p and q, this space is reflexive. Therefore, there exists a subsequence $\{x_k, \operatorname{adj} \nabla x_k\}$ that converges weakly to an element $(\bar{x}, \operatorname{adj} \nabla \bar{x})$ in the space $W^{1,2p} \times L^{2q}$.

Second, by the weak lower semicontinuity of the determinant function, we have that

$$\det \nabla x_k \rightharpoonup \det \nabla \bar{x} \ \text{ in } \ L^{\frac{2p}{n}}. \tag{3.7}$$

The weak lower semicontinuity of $\mathcal{E}$ is then a consequence of the convexity of $\mathcal{G}$ together with the convexity of $\mathcal{H}$.

Third, the property $\det \nabla \bar{x} \geq \phi_I$, a.e. in Ω, follows from Mazur's Lemma together with (3.7). Indeed, by Mazur's lemma, for a given $\epsilon > 0$, there exists $N = N(\epsilon)$, and $\{\lambda_i(\epsilon)\}$ with $\sum_1^N \lambda_i = 1$, such that

$$\| \sum_1^N \lambda_i \det \nabla x_i - \det \nabla \bar{x} \|_{0,\frac{2p}{n}} < \epsilon.$$

Since $\det \nabla x_k \geq \phi_I$, we conclude that $\det \nabla \bar{x} \geq \phi_I$ almost everywhere and hence, $\bar{x} \in \mathcal{A}$. □

The relation $\phi = \phi_I (\det \nabla \bar{x})^{-1}$ allows us to recover the volume fraction variable $\phi \in L^\infty(\Omega)$, which corresponds to the energy minimizer $\bar{x}$.

Remark 3.3. *In the case that $\mathcal{H}$ is non-convex, the previous theorem does not apply. In such a case, the total energy can be modified by the addition of a regularization term. The new gel energy becomes*

$$\mathcal{E}_r(x,\phi) = \int_\Omega \mathcal{W}_E(F)\,\mathrm{d}X + \int_{\Omega_x} \mathcal{W}_{FH}(\phi) + c_{FH} + \epsilon\,||\nabla_x \phi||^2\,\mathrm{d}x, \tag{3.8}$$

with $\epsilon > 0$ small. Although we do not consider this extension in the remainder of this work, we point out that existence of minimizer for a regularized energy of the form (3.8)*, although with a modified $\mathcal{W}_E$ and $\mathcal{W}_{FH}$, was studied in* [**29**].

4. Linearizations of the Euler-Lagrange equations

With the existence of minimizers proven for the convex case, we turn our attention to the Euler-Lagrange equations. The equations are nonlinear, but we provide a linearization of the equations in this section. This is motivated by the fact that, for the biomedical applications we aim to consider in this exposition, the linear regime may be meaningful and applicable. Moreover, qualitative effects of changes in the environment can be studied through linearized perturbations. The resulting linear boundary value problem will be further studied in Sections 5 and 6.

We consider the energy defined by (2.10), complemented by the energy potentials specified by (2.6), (2.7) and (2.8), and augmented by a body force g and a boundary stress s_0 on $\partial\Omega_1$. Upon taking variations, we obtain the Euler-Lagrange equilibrium equations over the reference domain Ω for the first Piola-Kirchhoff stress tensor $\mathcal{S}$, cf. (2.3). The strong form of the problem formally reads as follows. Find a deformation $x : \Omega \to \Omega_x$ and the associated stress tensor $\mathcal{S}$ such that

$$\mathcal{S} = \nu(\nabla x)\,\nabla x - \kappa(\nabla x)\,\nabla x^{-T} \quad \text{in } \Omega, \tag{4.1a}$$

$$\operatorname{div} \mathcal{S} = g, \quad \text{in } \Omega, \tag{4.1b}$$

$$x = x_0 \text{ on } \partial\Omega_0, \quad \mathcal{S}\cdot n = s_0 \text{ on } \partial\Omega_1, \tag{4.1c}$$

where n denotes the outward oriented normal of the boundary. The coefficients ν and κ are functions of a matrix variable and take the form

$$\begin{aligned} \nu(F) &= \mu_E \phi_I ||F||^{2(p-1)} \\ \kappa(F) &= \mu_E \phi_I ||I||^{2(p-1)} (\det F)^{-\beta} - (\det F)\left(\mathcal{W}_{FH}(\phi) + c_{FH} - \phi \mathcal{W}'_{FH}(\phi)\right). \end{aligned} \tag{4.2}$$

Here, we have reintroduced $\phi = \phi_I (\det F)^{-1}$ for the sake of notational brevity. Recall that $\mu_E > 0, \beta$, and $p \geq 1$ are elastic parameters, the former being the elastic shear modulus. As before, ϕ_I is the volume fraction in the reference configuration. The derivative of the Flory-Huggins potential $\mathcal{W}'_{FH}$ is the derivative with regard to the variable ϕ, cf. (2.8).

For frequent later reference, we label the residual stress (at $F = I$):

$$r(\phi) = \mathcal{W}_{FH}(\phi) + c_{FH} - \phi\mathcal{W}'_{FH}(\phi). \tag{4.3}$$

This quantity plays a key role in the numerical methods to follow. The residual stress is the stress in the reference state defined by the reference domain. For purely elastic materials, the residual stress is often assumed to vanish. In our model, the notion of vanishing residual stress becomes a restriction on the reference volume fraction ϕ_I, since the residual stress will vanish if and only if ϕ_I satisfies $r(\phi_I) = 0$. The existence of such ϕ_I is guaranteed by the lifting of the Flory-Huggins potential by c_{FH}. If ϕ_I is a global minimizer of the Flory-Huggins potential, $\mathcal{W}_{FH}(\phi_I) = -c_{FH}$ and $\mathcal{W}'_{FH}(\phi_I) = 0$, which thereby results in $r(\phi_I) = 0$. However, this constraint is too limiting for our purposes, as the volume fraction in the reference configuration may be arbitrary. We therefore also aim to consider ϕ_I such that $r(\phi_I) \neq 0$. In fact, this is a key point for the following.

4.1. The linearized boundary value problem. The Flory-Huggins energy potential $\mathcal{W}_{FH}$, and hence the constitutive equations of (4.1), depends on the physical parameters $\{a, b, c\}$. In the form given by (2.9), these parameters are explicitly dependent on the temperature, but may also be sensitive to other environmental parameters. In order to obtain a qualitative understanding of the effect of changes in the environment, perturbations of a generic environment parameter T may be considered.

Let the Piola-Kirchhoff stress tensor $\mathcal{S} = \mathcal{S}(F, T)$ be as defined by (4.1a). Assuming that $F = \nabla(x_0 + u)$ and $T = T_0 + \Delta T$, give the first order approximation:

$$\mathcal{S}(F, T) \approx \mathcal{S}(F_0, T_0) + \frac{\partial \mathcal{S}}{\partial F}(F_0, T_0)[\nabla u] + \frac{\partial \mathcal{S}}{\partial T}(F_0, T_0)[\Delta T], \tag{4.4}$$

where $F_0 = \nabla x_0$. The approximation is valid under the assumption of small perturbations in the deformation gradient and the environment parameter. On the right hand side, the first term corresponds to the force induced by the residual stress at (F_0, T_0), the second term is the elasticity tensor linearized about the deformation gradient F_0, and the third term corresponds to the force induced by the change of environment. If $\Delta T = 0$ and $F_0 = I$, (4.4) reduces to the linear elasticity approximation. A set of elementary, but somewhat lengthy calculations give the following expression of the Gateaux derivative of $\mathcal{S}$:

$$\begin{aligned} \frac{\partial \mathcal{S}}{\partial F}(F_0)[G] = {} & \nu(F_0)G + \kappa(F_0)F_0^{-T}G^T F_0^{-T} \\ & + \lambda_0(F_0)(F_0 : G)F_0 + \lambda_1(F_0)(F_0^{-T} : G)F_0^{-T}, \end{aligned} \tag{4.5}$$

where ν and κ are defined by (4.2). The constants λ_0 and λ_1 are defined by

$$\begin{aligned} \lambda_0(F) &= 2(p-1)\mu_E\phi_I||F||^{2(p-2)}, \\ \lambda_1(F) &= \mu_E\phi_I\beta||I||^{2(p-1)}(\det F)^{-\beta} + (\det F)(r(\phi) + \phi^2\mathcal{W}''_{FH}(\phi)), \end{aligned} \tag{4.6}$$

where, once again, we write $\phi = \phi_I(\det F)^{-1}$ and r is defined by (4.3).

We take a closer look at the residual stress at a given deformation gradient F_0. Since the constitutive relation is isotropic, $\mathcal{S}(F_0\cdot) = 0$ implies that $F_0 = f_0 I$ for some scalar f_0. In other words, for the residual stress at F_0 to vanish, F_0 must be a certain pure expansion or contraction. Furthermore, we note that for any such $f_0 I$, the forcing terms of (4.4) correspond to pure pressures. In the case $f_0 = 1$, $\mathcal{S}(I, \cdot) = r(\phi_I)I$, where $r(\phi_I)$ is defined by (4.3). The linear boundary value problem resulting from taking $f_0 = 1$ is summarized below in Problem 1.

PROBLEM 1. *Let Ω be an open and bounded domain in $\mathbb{R}^n$ with Lipschitz boundary $\partial\Omega = \partial\Omega_0 \cup \partial\Omega_1$, $\partial\Omega_0 \cap \partial\Omega_1 = \emptyset$. For a given $\phi_I : \Omega \to (0, 1)$ and given forces f*

and g, find the displacement $u : \Omega \to \mathbb{V}$ and the stress tensor $\sigma : \Omega \to \mathbb{M}$ satisfying the boundary conditions $u|_{\partial\Omega_0} = u_0$ and $\sigma \cdot n|_{\partial\Omega_1} = s_0$ and such that

$$\sigma - C_r[\nabla u] = r(\phi_I)I + f, \tag{4.7a}$$

$$\operatorname{div} \sigma = g. \tag{4.7b}$$

Here, r is defined by (4.3) and C_r is the residual-dependent stiffness tensor:

$$C_r[\nabla u] = \frac{\partial \mathcal{S}}{\partial F}(I)[\nabla u] = \mu(\phi_I)\,\nabla u + (\mu(\phi_I) - r(\phi_I))\,\nabla u^T + \lambda(\phi_I)(\operatorname{div} u)\,I, \tag{4.8}$$

and the generalized Lamé coefficients μ and λ are given by

$$\begin{aligned} \mu(\phi) &= \mu_E \phi ||I||^{2(p-1)} \\ \lambda(\phi) &= 2(p-1)\mu_E \phi ||I||^{2(p-2)} + \beta \mu_E \phi ||I||^{2(p-1)} + r(\phi) + \phi^2 \mathcal{W}''_{FH}(\phi). \end{aligned} \tag{4.9}$$

If ϕ_I is such that $r = r(\phi_I) = 0$, the stiffness operator C_r reduces to the standard linear elasticity tensor. Although the Lamé coefficients μ and λ additionally depend on the reference volume fraction ϕ_I and the Flory-Huggins potential $\mathcal{W}_{FH}$. In this case, the stress tensor is symmetric; that is, $\sigma : \Omega \to \mathbb{S}$. On the other hand, if ϕ_I is such that $r(\phi_I) \neq 0$, then the reference volume fraction is not an equilibrium volume fraction, and the gel is inclined to deform by either swelling or collapsing. We consider two possible approaches to this case and preface their description by noting that both approaches allow us to arbitrarily prescribe a reference volume fraction ϕ_I.

The first approach is simply to consider the system of equations (4.7) directly. However, the additional source introduced by the residual stress and, more importantly, the skew-symmetry part of the displacement gradient must be resolved. This approach is studied carefully in the main part of Section 5.

The second approach is restricted to the case where the material parameters are assumed to be homogenous. In this approach, we can consider linearizations about homogenous, non-identity deformations $F_0 = f_0 I$, where $f_0 \neq 1$. These deformations are pure expansions or contractions with f_0 chosen such that (4.1a) vanishes; that is, $\mathcal{S}(f_0 I, \cdot) = 0$. The stipulation is that f_0 is such that

$$\nu(f_0 I) f_0 - \kappa(f_0 I) f_0^{-1} = 0. \tag{4.10}$$

The linearized equations, resulting from considering (4.4) and (4.5) with $F_0 = f_0 I$, become (4.7) and (4.8) with $r = 0$ and the Lame coefficients

$$\mu(\phi, f_0) = \nu(f_0 I), \quad \lambda(\phi, f_0) = f_0^2 \lambda_0(f_0 I) + f_0^{-2} \lambda_1(f_0 I), \tag{4.11}$$

where λ_0, λ_1 are given by (4.6). In addition, we observe that the linearization about a non-identity state may require some additional care when applying boundary conditions, but postpone further discussion of this until Section 6.

Having set up the equations to study, we continue by discussing weak formulations and discretizations of Problem 1 and the linear stability of such in terms of the gel parameters, paying special attention to the role of residual stress r. We shall demonstrate the use of both the approaches described above in the subsequent sections, but will explicitly state when we consider (4.11) in place of (4.9).

5. Weak formulations and linear stability

The boundary value problem defined by Problem 1 with $r = 0$ is, in effect, the standard linear elasticity problem. Finite element and mixed finite element methods for the approximation of these equations are very well-understood and a multitude of approaches are available. In this work, we consider mixed finite elements, a choice which is motivated by two primary considerations. First, it is well-known that the pure displacement formulation, which results from eliminating the stress

tensor σ and looking for a displacement $u \in H^1(\Omega; \mathbb{V})$, is unreliable if $\lambda \to \infty$ or if the clamped boundary $\partial\Omega_0$ constitute only a small part of the total boundary [**5**]. In our problem, since $\lambda(\phi_I) \to \infty$ as $\phi_I \to 1$ for λ defined by (4.9), we seek weak formulations that are robust in the sense that they afford uniform convergence in the coefficient λ. Mixed finite element methods to offer such a feature. Second, in the biomedical confinement problems we address in this paper, the quantity of main interest is the stress $\sigma \cdot n$. Therefore, it may be natural to approximate this variable directly. A recent survey of mixed finite element methods for linear elasticity was presented by Falk [**13**].

In this section, we shall examine two types of mixed finite element methods for the gel equations as defined by Problem 1: the *displacement-pressure* and *stress-displacement-rotation* formulations. Both are robust with respect to the coefficient λ and typically give second (or higher) order stress approximations for smooth solutions. Both will be used for simulations in Section 6. In the displacement-pressure method, the displacement u and a pressure p are approximated directly. This formulation gives a version of the Stokes equations with an additional stabilization term. The discrete stability requirements on the associated pair of finite element spaces typically mandate higher order approximation spaces for the displacement. This, in turn, induces higher order approximations for the stresses. The second method, the stress-displacement-rotation formulation, approximates the stress, the displacement, and the rotation (the skew-symmetric part of the displacement gradient). Here the stress is approximated as the primal variable and therefore typically has a higher order accuracy than the displacement.

Both types of mixed finite element methods have been thoroughly studied from the linear elasticity view-point, and we refer to [**5, 13**] for a thorough treatment. We emphasize, however, that the case where $r \neq 0$ requires additional care. The main results in this section address the stress-displacement-rotation formulation in that setting.

5.1. Weak formulations for vanishing residual stress. We now state the two mixed formulations applied to the boundary value problem defined by Problem 1 with $r = 0$ and discuss the linear stability requirements placed on the gel parameters. The case $r \neq 0$ will be considered in the subsequent sections. For the sake of clarity, we assume that the boundary conditions are homogenous: $u_0 = 0$ on $\partial\Omega_0$, $s_0 = 0$ on $\partial\Omega_1$. The notation is as introduced in Section 3.

We begin with a discussion of the displacement-pressure formulation. Introducing a pressure $p = \lambda \operatorname{div} u$, one easily obtains the following weak displacement-pressure formulation of (4.7). Given $f \in L^2(\Omega; \mathbb{S})$ and $g \in L^2(\Omega; \mathbb{V})$, find $u \in H^1_0(\Omega; \mathbb{V})$ and $p \in L^2(\Omega)$ such that

$$
\begin{aligned}
\langle 2\mu\varepsilon(u), \varepsilon(v)\rangle + \langle p, \operatorname{div} v\rangle &= -\langle f, \nabla v\rangle - \langle g, v\rangle \quad \forall\, v \in H^1_0(\Omega; \mathbb{V}), \\
\langle \operatorname{div} u, q\rangle - \langle \lambda^{-1} p, q\rangle &= 0 \quad \forall\, q \in L^2(\Omega),
\end{aligned}
\tag{5.1}
$$

where ε is the symmetrized gradient and $\mu = \mu(\phi_I)$ and $\lambda = \lambda(\phi_I)$ are as defined in (4.9). The robustness in the limit $\lambda \to \infty$ is due to the fact that (5.1) in the are the Stokes equations in the limit $\lambda = \infty$. Furthermore, if $\mu > 0$ and $2\mu + n\lambda > 0$, there exist stable solutions u and p to (5.1).

The dependencies in the Lamé coefficients on the parameters of the problem have implications for the stability of the method. Notice that $\phi_I^2 \mathcal{W}''_{FH}(\phi_I) \to \infty$ corresponds to $\lambda \to \infty$. The robustness in λ may therefore be important if ϕ_I is close to 1 (or 0); that is, if the gel is almost dry (or almost fluid) in the reference configuration. It is clear that $\mu(\phi_I) > 0$ when defined by (4.9). In addition, $\lambda > 0$ if the Flory-Huggins energy potential $\mathcal{W}_{FH}$ is convex. If $\mathcal{W}_{FH}$ is sufficiently non-convex, λ may take negative values, and consequently the gel may display auxetic

behavior. It also follows that the condition $2\mu + n\lambda > 0$ can only break if $\mathcal{W}_{FH}$ is sufficiently non-convex. If one instead considers the Lamé parameters obtained by a linearization about a non-identity state (cf. (4.11), similar considerations apply. However, we see from (4.11) and (4.6), that λ might take negative values, even if $\mathcal{W}_{FH}$ is convex.

The alternative approach, the stress-displacement-rotation formulation, approximates the stress tensor σ and the displacement u separately while weakly enforcing the symmetry of the stress tensor. The rotation γ enters as a Lagrange multiplier corresponding to the latter symmetry constraint, and it is easily seen that $\gamma = \operatorname{skw} \nabla u$, where

$$2 \operatorname{skw} \tau = \tau - \tau^T.$$

The corresponding weak formulation of (4.7) with $r = 0$ reads as follows. Given $f \in L^2(\Omega; \mathbb{S})$ and $g \in L^2(\Omega; \mathbb{V})$, find $\sigma \in H_0(\operatorname{div}, \Omega; \mathbb{M})$, $u \in L^2(\Omega; \mathbb{V})$ and $\gamma \in L^2(\Omega; \mathbb{K})$ satisfying

$$\begin{aligned} \langle A_0 \sigma, \tau \rangle + \langle \operatorname{div} \tau, u \rangle + \langle \tau, \gamma \rangle &= \langle A_0 f, \tau \rangle \qquad \forall\, \tau \in H_0(\operatorname{div}, \Omega; \mathbb{M}), \\ \langle \operatorname{div} \sigma, v \rangle + \langle \sigma, \eta \rangle &= \langle g, v \rangle \quad \forall\, v \in L^2(\Omega; \mathbb{V}), \eta \in L^2(\Omega; \mathbb{K}). \end{aligned} \tag{5.2}$$

Here A_0 is the compliance tensor, given as the inverse of the stiffness tensor C_0:

$$A_0 \tau = \frac{1}{2\mu} \left(\tau - \frac{\lambda}{2\mu + n\lambda} \operatorname{tr} \tau \right). \tag{5.3}$$

with $\mu = \mu(\phi_I)$ and $\lambda = \lambda(\phi_I)$ as before. Similar considerations for the existence of stability of solutions with regard to the gel material parameters apply, as for the displacement-pressure formulation. In particular, A will not be uniformly bounded as $2\mu + n\lambda \to 0$. We remark that the symmetry of the stress tensor could also be enforced strongly, in the sense that the stress tensor function space could be restricted to the symmetric matrix fields [**3**]. However, as we shall see in the following, the weak symmetry approach leads itself more easily to the nearly symmetric case.

5.2. Weak formulations incorporating residual stress. The above systems of equations (5.1) and (5.2) illustrate that the weak formulations for the small-strain gel equations are entirely analogous to the weak formulations for the standard linear elasticity equations in the case where the residual stress vanishes. However, if the residual stress does not vanish, the skew-symmetric part of the displacement gradient and an additional source will enter the equations cf. (4.8). In particular, the stress tensor will *not* be symmetric. The displacement-pressure formulation (5.1) can easily be extended to this case. One may introduce a pressure $p = \lambda \operatorname{div} u$ as before and, instead of the symmetrized gradient, consider the displacement gradient and its transpose separately. However, the stress-displacement-rotation formulation requires more attention. In this section, we shall demonstrate how the stress-displacement-rotation formulation (5.2) can be extended to deal with a nearly symmetric stress tensor. To the authors' knowledge, this is an original approach. Accordingly, the derivation is presented in some detail.

The derivation of the stress-displacement-rotation formulation of (5.2) relies on both the symmetry of the stress tensor and the inversion of the stress-strain relation. Thus, traditionally, the symmetry of the stress tensor is a key factor. In the case $r \neq 0$, the stress tensor is not symmetric. However, the form of the stiffness tensor C and the premise that f takes symmetric values give a restriction on the skew component of the stress tensor. More specifically, the skew component of the stress tensor is proportional to the skew symmetric part of the displacement gradient:

$$\operatorname{skw}(\sigma) = r \operatorname{skw}(\nabla u) = r\gamma. \tag{5.4}$$

We note that if $r = 0$, this reduces to the classical symmetry constraint for the stress tensor.

The inversion of the stress-strain relation follows the standard procedure. From (4.8), we find that

$$\operatorname{tr}\varepsilon(u) = \frac{1}{\zeta_r}\left(\operatorname{tr}\sigma - nr - \operatorname{tr} f\right), \quad \zeta_r = 2\mu + n\lambda - r. \tag{5.5}$$

This relation allows us to eliminate of $\operatorname{tr}\varepsilon$ in terms of $\operatorname{tr}\sigma$. Assuming that $2\mu - r \neq 0$, we introduce the residual-dependent compliance tensor A_r:

$$A_r\sigma = \frac{1}{2\mu - r}\left(\sigma - \frac{\lambda}{\zeta_r}\operatorname{tr}\sigma I\right).$$

Further, let $\gamma = \operatorname{skw}(\nabla u)$ and recall that $\varepsilon(u) = \nabla u - \gamma$. The stress-strain relation follows

$$A_r\sigma = \nabla u - k_r\gamma + A_r\left(rI + f\right), \quad k_r = 1 - \frac{r}{2\mu - r}. \tag{5.6}$$

Multiplying (5.6), (4.7b), and (5.4) by fields τ, v and η respectively, and integrating the ∇u term by parts, give the following *nearly symmetric* weak formulation:

$$\begin{aligned} &\langle A_r\sigma, \tau\rangle + \langle \operatorname{div}\tau, u\rangle + \langle k_r\tau, \gamma\rangle = \langle A_r(rI + f), \tau\rangle \quad \forall\, \tau \in H_0(\operatorname{div}, \Omega; \mathbb{M}), \\ &\langle \operatorname{div}\sigma, v\rangle + \langle k_r\sigma, \eta\rangle - \langle rk_r\gamma, \eta\rangle = \langle g, v\rangle \quad \forall\, v \in L^2(\Omega; \mathbb{V}), \eta \in L^2(\Omega; \mathbb{K}). \end{aligned} \tag{5.7}$$

The relation (5.4) has been multiplied by the factor k_r for sake of the symmetry of the system of equations. Observe that $A_r(rI) = r\zeta_r^{-1}I$, where ζ_r is given by (5.5). The system of equations (5.7) can be written in the following, strong form:

$$\begin{pmatrix} A_r & -\nabla & k_rI \\ \operatorname{div} & 0 & 0 \\ k_r\operatorname{skw} & 0 & -rk_r \end{pmatrix} \begin{pmatrix} \sigma \\ u \\ \gamma \end{pmatrix} = \begin{pmatrix} A_r(rI + f) \\ g \\ 0 \end{pmatrix}. \tag{5.8}$$

Finally, we remark that if $r = 0$, then $k_r = 1$ and $rk_r = 0$. Thus, comparing (5.7) to (5.2), we see that the nearly symmetric formulation reduces to the weakly symmetric formulation in that case. We also note that $k_r \in (0, 2)$ when $r \in (-\infty, \mu)$.

5.3. Stability of nearly symmetric stress-displacement formulation. We now inspect the linear stability of the nearly symmetric formulation (5.7), paying special attention to r. The case $r = 0$ was discussed in the previous. For $r \neq 0$, it is easy to see from (5.8) that the system of equations take the form of a saddle point problem with an additional perturbation term. Therefore, this formulation also naturally lends itself to analysis using the Brezzi theory of saddle point problems and its extensions to perturbed saddle point problems [**7**]. However, we note that, in contrast to the standard perturbed case, the parameter r affects virtually all terms of (5.8).

The range of values for r for which the system of equations (5.7) has a stable solution is discussed in Lemma 5.1 and subsequent Theorem 5.2. The lemma gives an upper bound on the residual r in terms of the parameters μ and λ, such that the saddle-point part of (5.7) satisfy the classical Brezzi conditions. However, the $\langle rk_r\gamma, \eta\rangle$ term of (5.7) will act as a de-stabilization for $r < 0$. As this case simply corresponds to a negative residual stress, it is indeed relevant for the current exposition. Thus, we are also interested in a negative, lower bound. Theorem 5.2 makes these notions more precise. The case $f, g = 0$, and the assumption of homogenous elastic and mixing parameters, are considered for the sake of simplicity.

LEMMA 5.1. *For $\mu > 0$ and $\lambda \in \mathbb{R}$, assume that $r < \min(\mu, 2\mu + n\lambda)$. Define*

$$Z_r = \{\tau \in H(\operatorname{div}, \Omega; \mathbb{M}),\ \langle \operatorname{div}\tau, v\rangle + \langle \kappa_r\tau, \eta\rangle = 0 \quad \forall\, v \in L^2(\Omega; \mathbb{V}), \eta \in L^2(\Omega; \mathbb{K})\}.$$

Then, A_r is continuous and coercive on Z_r. In particular, there exist positive constants α_r, a_r such that

$$|\langle A_r\sigma,\tau\rangle| \le a_r||\sigma||_0||\tau||_0 \quad \forall\,\sigma,\tau\in H(\mathrm{div},\Omega;\mathbb{M}), \tag{5.9}$$

$$\langle A_r\tau,\tau\rangle \ge \alpha_r||\tau||_0^2 = \alpha_r||\tau||_{\mathrm{div}}^2 \quad \forall\,\tau\in Z_r. \tag{5.10}$$

Further, for any $v\in L^2(\Omega;\mathbb{V}), \eta\in L^2(\Omega;\mathbb{K})$, there exist a $\tau\in H_0(\mathrm{div},\Omega;\mathbb{M})$ and a $\beta>0$, independent of r, such that

$$||\tau||_{\mathrm{div}}\left(||v||_0+||k_r\eta||_0\right) \le \beta\left(\langle \mathrm{div}\,\tau, v\rangle + \langle k_r\tau,\eta\rangle\right). \tag{5.11}$$

Proof. The proof of this lemma is just a minor modification of the proof for the case with $r=0$, for instance the one given in [**5**, p.308 ff]. We simply remark that it is evident that for $\tau\in Z_r$, $\mathrm{div}\,\tau=0$. The bounds follow as usual, as the assumption $r<\mu$ in particular implies that $2\mu-r>0$ and $k_r>0$. □

Theorem 5.2. *Let $\sigma\in H(\mathrm{div},\Omega;\mathbb{M})$, $u\in L^2(\Omega;\mathbb{V})$ and γ in $L^2(\Omega;\mathbb{K})$ solve (5.7) with $f,g=0$. Assume that μ,λ and r are such that Lemma 5.1 holds. Additionally, if $r<0$, assume that*

$$\alpha_r - \frac{|r|}{k_r}2\beta^2a_r^2 > 0. \tag{5.12}$$

Then, there exists a constant $c>0$ such that

$$||\sigma||_{\mathrm{div}}^2 \le c||rI||_0^2, \quad ||u||_0+k_r||\gamma||_0 \le c(||\sigma||_0+||rI||_0). \tag{5.13}$$

Proof. For σ,u,γ solving (5.7),

$$\langle A_r\sigma,\sigma\rangle + \langle rk_r\gamma,\gamma\rangle = \langle A_r(rI),\sigma\rangle \tag{5.14}$$

Under the assumption that r is such that Lemma 5.1 holds, (5.11) in combination with (5.9) give:

$$||u||_0+|\kappa_r|||\gamma||_0 \le \beta a_r\left(||\sigma||_0+||rI||_0\right). \tag{5.15}$$

Hence, if $r\ge 0$, (5.14) immediately gives the stability bound (5.13). So, consider the case $r<0$. Using (5.14), (5.15) and (5.10), we obtain

$$\left(\alpha_r - \frac{|r|}{k_r}2\beta^2a_r^2\right)||\sigma||_0^2 \le \langle A_r(rI),\sigma\rangle + \frac{|r|}{k_r}2\beta^2a_r^2||rI||_0^2,$$

giving rise to the condition (5.12). □

In essence, the inequality of (5.12) quantifies a value $r_0<0$, such that (5.13) holds for $r\in(r_0,\mu)$. With reference to (4.3), recall that $r=r(\phi_I)$ is induced by a given reference volume fraction ϕ_I and that μ and λ are given by (4.9). We note that the estimates of Lemma 5.1 and Theorem 5.2 degenerate as $r\to\mu$, $2\mu-r+n\lambda\to 0$ or in the limiting case of (5.12). Hence, the behaviour in, and beyond, these limiting cases is not clear. This aspect will be investigated numerically in the next section. In particular, loss of stability is indeed detected for r sufficiently negative.

The approximation properties of a discretization of (5.7) now follows from the standard theory [**7**]. Let $\Sigma_h\subset H_0(\mathrm{div},\Omega;\mathbb{M})$, $V_h\subset L^2(\Omega;\mathbb{V})$ and $Q_h\subset L^2(\Omega;\mathbb{K})$ be finite dimensional spaces associated with a simplicial tessellation $\mathcal{T}_h$ of Ω, h denoting the meshsize. Assume that $\Sigma_h\times V_h\times Q_h$ defines a stable discretization of (5.2), that is, assume that (5.10) and (5.11), viewed over the discrete spaces, are satisfied with positive constants independent of h. Further, assume that the material parameters are such that Theorem 5.2 holds. Then, if $\sigma_h\in\Sigma_h, u_h\in V_h$

and $\gamma_h \in Q_h$ solve (5.7) for all $\tau \in \Sigma_h, v \in V_h$ and $\eta \in Q_h$, there exists $C > 0$, independent of h such that

$$||\sigma - \sigma_h||_{\mathrm{div}} + ||u - u_h||_0 + ||\gamma - \gamma_h||_0 \\ \leq C \left(\inf_{\tau \in \Sigma_h} ||\sigma - \tau||_{\mathrm{div}} + \inf_{v \in V_h} ||u - v||_0 + \inf_{\eta \in Q_h} ||\gamma - \eta||_0 \right).$$

5.4. Choice of discretization spaces. We conclude this section by some considerations regarding specific discretizations of (5.1) and (5.7). As our main emphasis is robust and accurate stress approximation, we briefly discuss some choices towards the achievement of this aim. We denote the space of continuous piecewise polynomials of degree (less than or equal to) k defined relative to a simplicial tessellation $\mathcal{T}_h$ by $\mathcal{P}_k^c(\mathcal{T}_h)$ and the space of discontinuous piecewise polynomials by $\mathcal{P}_k(\mathcal{T}_h)$. The space of piecewise polynomial vector fields of degree k with continuous normal components over edges (or faces) is labelled $\mathrm{BDM}_k(\mathcal{T}_h)$, (cf. [**6, 20**]), with $\mathrm{BDM}_k(\mathcal{T}_h; \mathbb{V}) = \mathrm{BDM}_k(\mathcal{T}_h)^n$.

First, for discretizations $U_h \times P_h \subset H^1(\Omega; \mathbb{V}) \times L^2(\Omega)$ of the displacement-pressure formulation (5.1), we use the generalized Taylor-Hood elements: for $k = 1, 2, \ldots$ continuous, piecewise $(k+1)$'th degree polynomial fields for the displacement and continuous, piecewise k'th degree fields for the pressure. The lowest order elements are thus

$$U_h \times P_h = \mathcal{P}_2^c(\mathcal{T}_h; \mathbb{V}) \times \mathcal{P}_1(\mathcal{T}_h). \tag{5.16}$$

For smooth solutions, $(k+1)$'th order convergence is expected in the H^1-norm of the displacement and hence for a post-calculated stress tensor, as well.

Second, for the nearly symmetric stress-displacement-rotation formulation (5.7), we utilize the lowest order element spaces of the family suggested by Farhloul and Fortin [**14**] and Falk [**13**]; specifically, for $k = 1, 2, \ldots$, piecewise k'th degree polynomial tensor fields with row-wise normal component continuity for the stresses, discontinuous $(k-1)$'th degree for the displacements and continuous k'th degree for the rotations. The lowest order is

$$\Sigma_h \times V_h \times Q_h = \mathrm{BDM}_1(\mathcal{T}_h; \mathbb{V}) \times \mathcal{P}_0(\mathcal{T}_h; \mathbb{V}) \times \mathcal{P}_1^c(\mathcal{T}_h; \mathbb{K}). \tag{5.17}$$

These elements are preferred to the similar spaces considered by Arnold et al. [**2**] because the higher order stress interpolation error is conserved at little additional cost. As for the displacement-pressure discretization, the discretization defined by (5.17) gives $(k+1)$'th order convergence for the stress tensor for smooth solutions.

Thus, in theory, both formulations with the previously described discretizations give the same order of convergence for the stresses. One advantage of the displacement-pressure approach is that it is typically less expensive than the stress-displacement-rotation approach. This aspect may be particularly relevant in three dimensions. However, the disadvantage to using this method is that the stress tensor must be post-calculated. This is evidently not an issue with the stress-displacement-rotation approach. An additional advantage of the stress-displacement-rotation approach is that it gives stress tensors with normal component continuity and hence continuous stresses over inter-element boundaries. The latter consequence is particularly relevant if there are internal interfaces in the domain where the interfacial stress is a quantity of interest.

6. Simulations

In this section, we present some numerical simulations. First, we examine the stability range of the nearly symmetric stress-displacement-rotation formulation numerically in order to compare the observations with the analytical estimates.

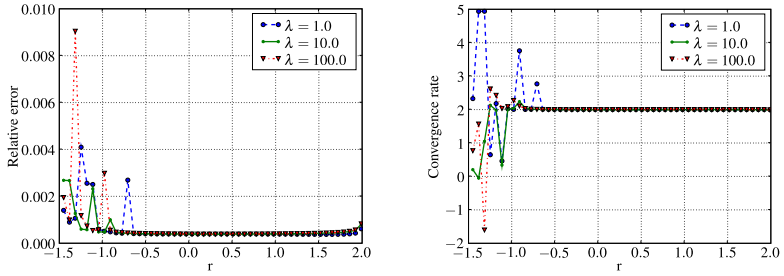


FIGURE 1. Approximation properties of the nearly-symmetric formulation in terms of the residual stress for $\lambda = 1, 10, 100$. Left: Stress approximation errors. Relative error E_r of the stress approximation versus residual stress r. $E_r = ||\sigma^r - \sigma_h^r||_0/||\sigma^r||_0$. $h = 1.0/40$. Right: Convergence rates. Convergence rate c_r of the stress approximations versus residual stress r. $c_r = ||\sigma_{h/2}^r||_0/||\sigma_h^r||$. $h = 1.0/20$

Next, we turn to a more physically realistic application. The gel model and numerical methods presented are applied to study the shear stresses arising from the confinement and environmental effects experienced by an artificial bone implant. All simulations have been performed using the DOLFIN library of the FEniCS project [**15, 17**].

6.1. Numerical stability of nearly symmetric stress formulation. The residual stress r clearly has an impact on the approximation properties of discretizations of the nearly symmetric stress formulation. In order to numerically investigate this effect, we consider a range of values for r and compare the error of approximation and the rate of convergence for a given discretization of (5.7). We remark that such a study can be performed independently of the specific expressions for the Lamé parameters given by (4.9). However, using a range of values for μ and λ relevant for the gel model seems appropriate. In lieu of (4.9) and (5.12), we let $\mu = 1.0$, consider a set of positive λ, and $r \in (-1.5\,\mu, 2\mu)$.

The following domain and exact smooth solutions are considered in order to examine the error and the convergence rate of the approximations. We let $\Omega = (0,1)^2 \subset \mathbb{R}^2$ and consider the boundary $\partial\Omega = \partial\Omega_0 \cup \partial\Omega_1$, where

$$\partial\Omega_0 = \{X \in \partial\Omega, X_0 \in \{0,1\}\}, \quad \partial\Omega_1 = \partial\Omega \backslash \partial\Omega_0$$

Define the following smooth displacement:

$$u = u(x_0, x_1) = (\sin(\pi x_0), \sin(\pi x_0)). \tag{6.1}$$

For a given r, we define the associated stress tensor σ^r and rotation γ^r in accordance with (4.8) and (5.4), respectively. We also let $g^r = \operatorname{div}\sigma^r$, $s_0^r = \sigma^r \cdot n$ on $\partial\Omega_1$. By construction, σ^r, u, γ solve (5.7) with the boundary conditions $u|_{\partial\Omega_0} = 0$ and $\sigma^r \cdot n|_{\partial\Omega_1} = s_0^r$. Furthermore, we let $\mathcal{T}_h$ be a uniform, regular triangulation of Ω, and we consider the discretization defined by (5.17). Recall that if the residual stress r, viewed in connection with the Lamé coefficients μ and λ, is such that (5.12) holds, we expect second order convergence for the stress approximations in the L^2 norm.

The relative error of the stress approximation for a given mesh size $h = 1.0/40$ and the rate of convergence for the stress in the L^2-norm between $h = 1.0/20$ and $h/2 = 1.0/40$ are plotted in Figure 1 for $\lambda \in \{1.0, 10.0, 100.0\}$. We observe that for

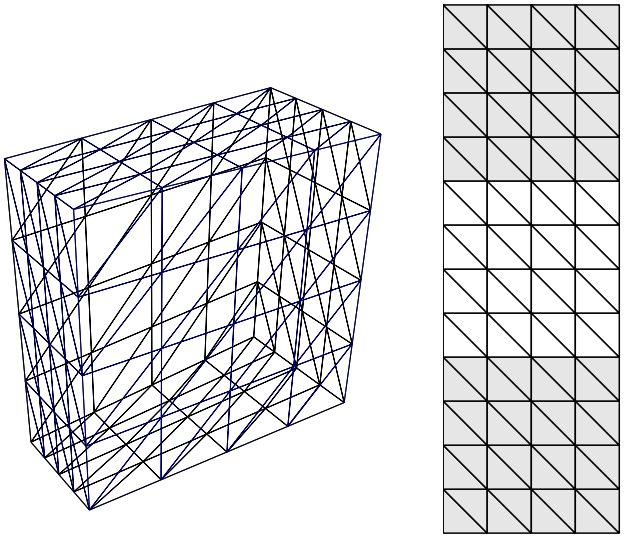
$\lambda = 1.0$, the relative error of the approximation seems uniform for $-0.5 < r < 1.0$, and the convergence rate is of second order. For $r < -0.5$, the relative error grows in a non-regular fashion, and the convergence rates vary dramatically. Further investigations show that the convergence rates are also highly dependent on the mesh size. The observed behaviour agrees with the theoretical results. The method is not guaranteed stable for r sufficiently negative. Indeed, the above observations give evidence of instabilities for $r < -0.5$. We note that no instability is observed for r close to 1.0, indicating that the requirement $r > \mu$, motivated by $k_r > 0$, could be relaxed. For $r \to 2.0$, we note only a slight increase in the relative error and no observed variation in the convergence rate. This indicates that the stress approximation is only slightly affected by the small coercivity constant. The cases $\lambda = 10.0$ and $\lambda = 100.0$ are entirely analogous to the case $\lambda = 1.0$. This indicates that increasing λ does not have a deleterious impact on the performance of the method.

6.2. Swelling-induced stress in artificial bone implants. We now turn to consider a specific biomedical application: an artificial bone implant used for high tibial osteotomy. In this procedure, the knee is re-aligned to shift the body weight from a damaged area of the knee to the side with healthy cartilage. This is performed by removing a wedge from the shinbone and then chemically gluing the bone implant into the open space in order to realign the knee. One natural question becomes whether and how the additional moisture of the body affects the implant. Another relevant question is how the confinement of the gel affects the stresses. In particular, both the implant itself and the glue attaching the implant to the bone may fail when exposed to high shear forces.

In the following, we aim to apply the appropriately linearized gel equations and the numerical methods of the previous sections to examine the shear stresses acting on the implant. First, we study the stresses arising from the confinement of an implanted polymer gel. Second, using the approximation given in (4.4), we investigate how temperature changes in the body may affect the implant. Finally, an interface problem between the bone and the implant is considered. Although we consider a specific device, the simulations illustrate the more general situation of swelling in a confined area under environmental and mechanical forces.

Polymers used in biomedical devices which provide mechanical reinforcement must have low interaction with water and be rigid so as to prohibit large swelling. The following homogenous parameter values are representative of polymers used in artificial bone implants [**9, 21**]. In (2.9) and (2.7), we let $V_m = 0.1\text{nm}^3$, $N_1 = 1000$, $N_2 = 1$, $\mu_E = 1.0$ GPa, $\beta = 2$, $\Delta w = (150\,\text{K})K_B$, and consider an initial temperature $T = 310$ K. The choice of Δw is such that $\chi(T) = 0.5$ at $T = 300$ K, in agreement with [**9**]. The associated Flory-Huggins energy density has a single minima at $\phi_I^N \approx 0.66$ and hence $r(\phi_I^N) = 0$. A physically realistic value for the volume fraction of the polymer, however, is typically close to 1. With this in mind, we consider a set of reference volume fraction values ϕ_I in the range $[0.6, 1.0)$.

Confinement. If the reference volume fraction is higher than the volume fraction corresponding to a stress-free reference state, the gel will respond with an initial swelling. In order to isolate this effect, we consider the linearized equations about a non-identity state $F_0 = f_0 I$ such that $\mathcal{S}(F_0) = 0$ cf. (4.10). For $\phi_I = 0.995$, $f_0 \approx 1.0038$. The equations to be considered are thus those of Problem 1 with $f, g = 0$, but with μ and λ given by (4.11). The non-identity linearization has ramifications for the boundary conditions. Since the total deformation can be given as $x(X) = u(X) + f_0 X$, in order to represent a bone implant entirely confined at some



(a) (b)

FIGURE 2. The simulation domains. (a): A coarse tessellation of the domain occupied by the gel. (b): A coarse triangulation of the cross-sectional domain occupied by the gel and the bone.

$\partial\Omega_0$, we let

$$u(X) = u_0(X) = (1 - f_0)X, \quad X \in \partial\Omega_0, \tag{6.2}$$

thus ensuring no deformation of the reference body at the given boundary. However, for $f_0 \neq 1$, the inclination to swell versus the confinement of the implant yields a non-zero u_0. In particular, shear stress forces are induced. At $\partial\Omega_1$, we let $\sigma \cdot n = s_0 = 0$.

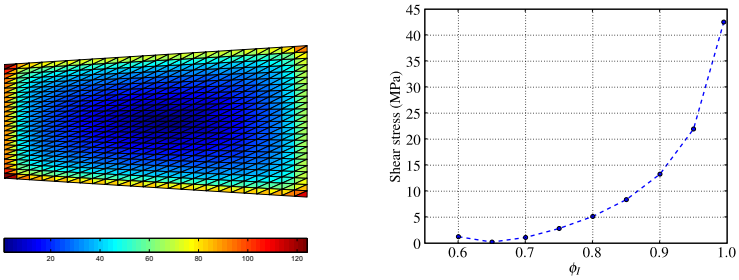
Motivated by typical implant shapes, we consider a hexahedral reference domain for the bone implant defined by an isosceles trapezoidal base. More precisely, let $\Omega^0 = (0, 20) \times (0, 10) \times (0, 20)$ (mm^3) and in general, $\Omega^\theta = \{(X_0, X_1, X_2)\} \subseteq \Omega^0$ such that

$$X_1 \in ((2 - X_0)\tan(\theta), 1 - (2 - X_0)\tan(\theta)).$$

for $0 \leq \theta \leq \theta_{\max} = \arctan(1/4)$. Thus, $\Omega^{\theta_{\max}}$ is defined by a isosceles triangular base. In the following, we consider $\theta = \arctan(1/16)$. This domain is illustrated in Figure 2(a). Further, the implant is assumed to be confined at the top and bottom boundaries, where it adjoins the bone and stress-free at the remaining boundary:

$$\partial\Omega_0 = \{X \in \partial\Omega, X_2 \in \{0, 2\}\}, \quad \partial\Omega_1 = \partial\Omega \backslash \partial\Omega_0. \tag{6.3}$$

The resulting boundary value problem was simulated using the displacement-pressure formulation (5.1) with the Taylor–Hood elements (5.16) on a regular tessellation of Ω consisting of approximately 80 000 tetrahedral cells. The stress tensor was post-calculated from the displacement and the pressure approximations using (4.7a). The quantity of interest is the shear stresses at the top and bottom boundaries of the implant. Due to symmetry, we only consider the top boundary Γ. The magnitude of the tangential component of the stress at the barycenter of each facet at the top boundary is shown in Figure 3(a) for $\phi_I = 0.995$. The reported values are in the range $(0, 125)$ MPa, but should be interpreted qualitatively. We



(a) (b)

FIGURE 3. Shear stresses, in MPa, resulting from the confinement of a gel in a non-equilibrium state. (a): The magnitude of the shear stress vector measured at the barycenters of each facet of the top boundary Γ for $\phi_I = 0.995$. (b): The average shear stress over the top boundary S_Γ versus initial volume fraction ϕ_I.

see that the stresses are, relatively speaking, low in the interior of the top facet, but that a boundary layer forms, which gives very high stress values at the corners. In fact, these pointwise stress values does not seem to be bounded in the sense that the maximal value depends on the mesh size. Considering the unreliability of this pointwise stress magnitude, we instead measure the average shear stresses over the top boundary:

$$S_\Gamma^2 = \frac{1}{\int_\Gamma 1 \,\mathrm{ds}} \left(\int_\Gamma \sigma_{20}^2 + \sigma_{21}^2 \,\mathrm{ds} \right). \tag{6.4}$$

The shear stresses S_Γ corresponding to ϕ_I in the range $(0.60, 1.0)$ is plotted in Figure 3(b). We see that the stresses are very close to zero for $\phi_I = 0.65$, as this is close to the natural volume fraction of the energy. The average shear stresses grow as ϕ_I approaches 1.0: At $\phi_I = 0.995$, the average shear stress is approximately 42.1 MPa. We conclude that the confinement versus the swelling gives significant shear stress values for the implant.

Temperature. We now turn to inspect how an increase in temperature affects the implant. In order to obtain a qualitative understanding of the effects of changes in temperature, we consider a perturbation about the initial temperature, as discussed in connection with (4.4). The temperature dependence of the given model occurs only through the coefficients of the Flory-Huggins potential, and thus, cf. (4.2), the resulting force f reads:

$$f = f(\Delta T) = \frac{\partial}{\partial T} \left(\mathcal{W}_{FH}(\phi, T) + c_{FH}(T) - \phi \mathcal{W}'_{FH}(\phi, T) \right) (\Delta T) \, (\det F) F^{-T}.$$

We let $T = 310$ K as before and consider $\Delta T = 1.0$. The parameter values, domain, and discretization is as for the previous experiment. The resulting pressure force is of the order 10^{-1} MPa for $\phi_I = 0.995$ and thus several orders of magnitude smaller than the shear stresses resulting from the confinement. In order to isolate the effects of the change in temperature, we modify the boundary conditions of (6.2), and let $u_0 = 0$ on $\partial\Omega_0$. This corresponds to letting the implant swell to the equilibrium position and penalizing the additional swelling resulting from the changes in temperature only.

The resulting average shear stresses at the top boundary of the implant, following the same procedure as in the previous, are illustrated in Figure 4. We point at

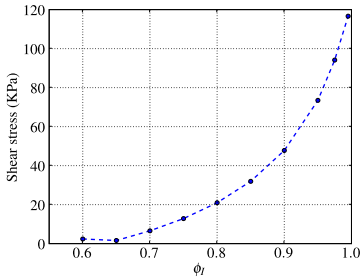


FIGURE 4. Shear stresses, in KPa, resulting from temperature changes. Average shear stress over the top boundary S_Γ versus initial volume fraction ϕ_I.

the difference in units between Figures 3(b) and 4. The average shear stress values are in the range $(1.61, 117)$ KPa and thus significantly lower than those associated with the full confinement.

Gel–bone interaction. Finally, we aim to consider the interface problem between the gel implant and the bone to which it is attached. In the first experiment of this section, the implant was assumed to be entirely confined at the top and bottom boundary. This assumption does not take the elastic properties and behaviour of the bone into account. In particular, in reality, it seems meaningful to assume that the bone will absorb some of the swelling of the gel. Here, we aim to examine the shear stresses of the gel–bone interface when the bone is assumed to be linearly elastic. In particular, we consider a two-dimensional cross-section, in the longitudinal direction, of the gel implant and the bone.

Let $\Omega = (0, 20) \times (0, 60)\,\mathrm{mm}^2$. This reference domain is illustrated in Figure 2(b). The lower and upper interfaces between the gel and the bone are placed at the line segments defined by $X_1 = 20.0$ and $X_1 = 40.0$, respectively. Thus, the middle third of the domain is occupied by the gel, while the upper and lower thirds are occupied by the bone. We assume that the bone is confined at the top and bottom boundaries; that is, at the line segments defined by $X_1 = 0$ and $X_1 = 60$. Both the bone and the gel are assumed to be stress-free at the left and right boundaries.

The gel parameter values and the range of reference volume fractions to be considered are as before. However, instead of linearizing about a non-identity, but stress-free state, we now consider the small-strain equations resulting from the linearization about the identity deformation. Thus, the Lamé parameters of the gel, $\mu(\phi_I)$ and $\lambda(\phi_I)$, are as defined by (4.9). The residual stress $r(\phi_I)$ will act as a forcing term, cf. (4.7a). We remark that for above parameter values and the given range of reference volume fractions, $r(\phi_I) \in (-0.17, 0.0036)$. Furthermore, we give the bone the Lamé parameters $\mu_\mathrm{b} = 4.0$ GPa and $\lambda_\mathrm{b} = 20.0$ GPa. Under these assumptions, both the bone and the gel are governed by (4.7). The Lamé parameters μ and λ, however, are now heterogenous.

As in the previous, any non-equilibrium gel reference volume fraction ϕ_I will produce a swelling or collapse of the domain. Our main interest lies with the shear stresses at the upper and lower material interfaces. As these are interfaces in the interior of the domain, the stress-displacement-rotation formulation (5.7) and the discretization defined by (5.17) are used for this simulation. We denote the upper gel-bone interface by Γ_{gb}. Keeping with the previous, we consider the average shear

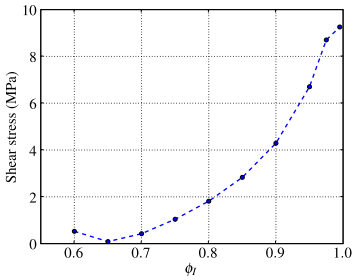


FIGURE 5. Average shear stress, in MPa, over the upper gel–bone boundary $S_{\Gamma_{\text{gb}}}$ versus initial volume fraction ϕ_I.

stress over this interface:

$$S^2_{\Gamma_{\text{gb}}} = \frac{1}{\int_{\Gamma_{\text{gb}}} 1 \, \text{ds}} \left(\int_{\Gamma_{\text{gb}}} \sigma^2_{01} \, \text{ds} \right).$$

The triangulation used for the simulation contains approximately 25 000 triangles. The simulated average shear stress values are plotted versus the reference volume fraction in Figure 5. For $\phi_I = 0.65$, $S_{\Gamma_{\text{gb}}} = 8.79 \times 10^{-2}$ MPa. This relatively small value is expected as this volume fraction is close to the reference volume fraction with zero residual stress. The shear stresses increase as the volume fraction increases: at $\phi_I = 0.995$, $S_{\Gamma_{\text{gb}}} = 9.26$ MPa. We note that the shear stresses are smaller than the ones resulting from the total confinement, but still significant.

Acknowledgments. The authors wish to thank Suping Lyu, Medtronic Inc for supplying appropriate material parameters. The first author also wishes to thank Ragnar Winther for fruitful discussions and comments on the manuscript. The first author also acknowledges the support of the Centre of Mathematics for Applications, University of Oslo.

CENTER FOR BIOMEDICAL COMPUTING, SIMULA RESEARCH LABORATORY, P.O.BOX 134, 1325 LYSAKER, NORWAY. meg@simula.no

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA, 127 VINCENT HALL, 206 CHURCH STREET S. E., MINNEAPOLIS, MN 55455 mcc@math.umn.edu.

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA, 127 VINCENT HALL, 206 CHURCH STREET S. E., MINNEAPOLIS, MN 55455. mice0012@math.umn.edu

Bibliography

[1] E. S. Almeida and R. L. Spilker. Mixed and penalty finite element models for the nonlinear behaviour of biphasic soft tissues in finite deformation: Part I – alternate formulations. *Computer methods in Biomechanics and Biomedical engineering*, 1:25–46, 1997.

[2] D. N. Arnold, R. Falk, and R. Winther. Mixed finite element methods for linear elasticity with weakly imposed symmetry. *Math. Computation*, 76:1699–1723, 2007.

[3] D. N. Arnold and R. Winther. Mixed finite elements for elasticity. *Numer. Math.*, 92:401–419, 2002.

[4] J. M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 63:337–403, 1977.

[5] D. Braess. *Finite elements. Theory, fast solvers, and applications in solid mechanics*. Cambridge University Press, 2001.

[6] F. Brezzi, J. Douglas, Jr., and L. D. Marini. Two families of mixed finite elements for second order elliptic problems. *Numer. Math.*, 47:217–235, 1985.

[7] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Springer, 1991.

[8] M. C. Calderer, B. Chabaud, S. Lyu, and H. Zhang. Modeling approaches to the dynamics of hydrogel swelling. *Journal of Computational and Theoretical Nanoscience*, 2008. Accepted.

[9] M. C. Calderer and H. Zhang. Incipient dynamics of swelling of gels. *SIAM J. Appl. Math.*, 68:1641–1664, 2008.

[10] P. Ciarlet. *Mathematical Elasticity, Volume I: Three-Dimensional Elasticity*. North-Holland, Amsterdam, 1987.

[11] J. Dolbow, E. Fried, and H. Ji. Chemically induced swelling of hydrogels. *Journal of Mechanics and Physics of Solids*, 52:51–84, 2004.

[12] J. Dolbow, E. Fried, and H. Ji. A numerical strategy for investigating the kinetic response of stimulus-responsive hydrogels. *Computer methods in applied mechanics and engineering*, 194:4447–4480, 2005.

[13] R. S. Falk. Finite element methods for linear elasticity. In *Mixed Finite Elements, Compatibility Conditions and Applications*. Springer, 2008.

[14] M. Farhloul and M. Fortin. Dual hybrid methods for the elasticity and the Stokes problem: a unified approach. *Numer. Math.*, 76:417–440, 1997.

[15] T. FEniCS project. FEniCS. http//www.fenics.org/.

[16] P. Flory. *Principles of Polymer Chemistry*. Cornell University Press, 1953.

[17] J. Hoffman, J. Jansson, A. Logg, G. N. Wells, et al. DOLFIN. http//www.fenics.org/dolfin/.

[18] W. Hong, Z. Liu, and Z. Sou. Inhomogenous swelling of a gel in equilibrium with a solvent and mechanical load. *Submitted*, 2008.

[19] W. Hong, X. H. Zhao, J. Zhou, and Z. Suo. A theory of coupled diffusion and large deformation in polymeric gels. *Journal of the Mechanics and Physics of Solids*, 56:1779–1793, 2008.

[20] J. C. Nédélec. A new family of mixed finite elements in $\mathbf{R}^3$. *Numer. Math.*, 50:57–81, 1986.

[21] Y. Osada and E. K. Kajiwari. *Gels Handbook*. Academic Press, 2001.

[22] D. J. Segalman and W. R. Witkowski. Two-dimensional finite element analysis of a polymer gel drug delivery system. *Materials Science and Engineering C*, 2:243–249, 1995.

[23] H. Simpson and S. Spector. On composite matrices and strong ellipticity for isotropic elastic materials. *Arch Rat Mech Anal*, 84:55–68, 1983.

[24] T. Tanaka. Collapse of gels and the critical endpoint. *Phys. Rev. Lett.*, 40:820–823, 1978.

[25] C. Truesdell. *Rational Thermodynamics*. Springer, 1984.

[26] T. Yamaue and M. Doi. Swelling dynamics of constrained thin-plate gels under an external force. *Phys. Rev. E*, 70:011401, 2004.
[27] T. Yamaue and M. Doi. Theory of one-dimensional swelling dynamics of polymer gels under mechanical constraint. *Phys. Rev. E*, 69:041402, 2004.
[28] T. Yamaue and M. Doi. The stress diffusion coupling in the swelling dynamics of cylindrical gels. *J. Chem. Phys.*, 122:084703, 2005.
[29] H. Zhang. Static and dynamical problems of hydrogel swelling: Modeling and analysis. Ph.D. thesis, University of Minnesota, 2007.

Paper III

EFFICIENT ASSEMBLY OF H(DIV) AND H(CURL) CONFORMING FINITE ELEMENTS

MARIE E. ROGNES, ROBERT C. KIRBY AND ANDERS LOGG

Abstract. In this note, we discuss how to efficiently evaluate and assemble general finite element variational forms on $H(\mathrm{div})$ and $H(\mathrm{curl})$. In particular, we extend a previously presented representation theorem for affinely mapped elements to Piola-mapped elements, discuss a simple numbering strategy that removes the need to contend with directions of facet normals and tangents, and present an automated and easy-to-use implementation that allows a user to specify finite element variational forms on $H(\mathrm{div})$ and $H(\mathrm{curl})$ in close to mathematical notation.

1. Introduction

The Sobolev spaces $H(\mathrm{div})$ and $H(\mathrm{curl})$ play an important role in many applications of mixed finite element methods to partial differential equations. Examples include second order elliptic partial differential equations, Maxwell's equations for electromagnetism and the linear elasticity equations. Mixed finite element methods may provide advantages over standard H^1 finite element discretizations in terms of added robustness, stability, and flexibility. However, implementing $H(\mathrm{div})$ and $H(\mathrm{curl})$ methods requires additional code complexity for constructing basis functions and evaluating variational forms, which helps to explain their relative scarcity in practice.

The FEniCS project [**13, 24**] comprises a collection of free software components for the automated solution of differential equations. One of these components is the FEniCS form compiler (FFC) [**18, 19, 23**]. FFC allows finite element spaces and multilinear forms over simplicial meshes to be specified in a form language close to the mathematical abstraction and notation. The form compiler generates low-level (C++) code for efficient form evaluation and assembly based on an efficient tensor contraction. Moreover, the FErari project [**17, 20, 21, 22**] has developed specialized techniques for further optimizing this code based on underlying discrete structure.

FFC relies on the FInite element Automatic Tabulator (FIAT) [**14, 15, 16**] for the tabulation of finite element basis functions. FIAT provides methods for efficient tabulation of finite element basis functions and their derivatives at any

particular point. In particular, FIAT provides $H(\mathrm{div})$ elements such as the families of Raviart–Thomas [**31**], Brezzi–Douglas–Marini [**7**], and Brezzi–Douglas–Fortin–Marini [**6**], as well as $H(\mathrm{curl})$ elements of the Nedelec types [**27, 28**].

Previous iterations of FFC have enabled easy use of H^1 and L^2 conforming finite element spaces, including discontinuous Galerkin formulations, but support for $H(\mathrm{div})$ and $H(\mathrm{curl})$ spaces has been absent. In this paper, we extend the previous work [**18, 19, 29**] to allow simple and efficient compilation of variational forms on $H(\mathrm{div})$ and $H(\mathrm{curl})$, including mixed formulations on combinations of H^1, $H(\mathrm{div})$, $H(\mathrm{curl})$, and L^2.

Implementations of $H(\mathrm{div})$ and $H(\mathrm{curl})$ finite element spaces, in particular of arbitrary degree, are not prevalent. There is, to our knowledge, no implementations that utilize the compiled approach to combine the efficiency of low-level optimized code with a fully automated high-level interface. Some finite element packages, such as FEAP [**12**], do not provide $H(\mathrm{div})$ or $H(\mathrm{curl})$ type elements at all. Others, such as FreeFEM [**30**], typically only provide low-order elements such as the lowest-order Raviart–Thomas elements. Some libraries like deal.II[2] [**5**] or FEMSTER [**9**] do provide arbitrary degree elements of Raviart–Thomas and Nedelec type, but do not automate the evaluation of variational forms. NGSolve [**33**] provides arbitrary order $H(\mathrm{div})$ and $H(\mathrm{curl})$ elements along with automated assembly, but only for a predefined set of bilinear forms.

The outline of this paper is as follows. We begin by reviewing basic aspects of the function spaces $H(\mathrm{div})$ and $H(\mathrm{curl})$ in Section 2, and provide examples of variational forms defined on these spaces. We continue, in Section 3, by summarizing the $H(\mathrm{div})$ and $H(\mathrm{curl})$ conforming finite elements implemented by FIAT. In Section 4, we recap the multilinear form framework of FFC and present an extension of the representation theorem from [**19**]. Subsequently, in Section 5, we provide some notes on the assembly of $H(\mathrm{div})$ and $H(\mathrm{curl})$ elements. Particular emphasis is placed on aspects not easily found in the standard literature, such as choice of orientation of geometric entities. In Section 6, we return to the examples introduced in Section 2 and illustrate the ease and terseness with which even complicated mixed finite element formulations may be expressed in the FFC form language. Convergence rates in agreement with theoretically predicted results are presented to substantiate the veracity of the implementation. Finally, we make some concluding remarks and indicate directions for further work in Section 7.

2. $H(\mathrm{div})$ AND $H(\mathrm{curl})$

In this section, we summarize some basic facts about the Sobolev spaces $H(\mathrm{div})$ and $H(\mathrm{curl})$ and discuss conforming finite element spaces associated with these. Our primary focus is on properties relating to inter-element continuity and change of variables. The reader is referred to the monographs [**8**] and [**26**] for a more thorough analysis of $H(\mathrm{div})$ and $H(\mathrm{curl})$ respectively.

2.1. Definitions. For an open domain $\Omega \subset \mathbb{R}^n$, we let $L^2(\Omega; \mathbb{R}^n)$ denote the space of square-integrable vector fields on Ω with associated norm $||\cdot||_0$ and inner-product $\langle \cdot, \cdot \rangle$, and abbreviate $L^2(\Omega) = L^2(\Omega; \mathbb{R}^1)$. We define the following standard differential operators on smooth fields v: $D^\alpha v = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_m}^{\alpha_m} v$ for a multiindex α of length m, $\operatorname{div} v = \sum_{i=1}^n \partial_{x_i} v_i$, $\operatorname{curl} v = (\partial_{x_2} v_3 - \partial_{x_3} v_2, \partial_{x_3} v_1 - \partial_{x_1} v_3, \partial_{x_1} v_2 - \partial_{x_2} v_1)$ and $\operatorname{rot} v = \partial_{x_1} v_2 - \partial_{x_2} v_1$. We may then define the spaces $H^m(\Omega)$, $H(\operatorname{div}, \Omega)$, and

[2] deal.II considers quadrilateral meshes only, where the bases may be constructed from tensor products.

$H(\mathrm{curl},\Omega)$ by

$$H^m(\Omega) = \{v \in L^2(\Omega) :\ D^\alpha v \in L^2(\Omega),\ |\alpha| \le m\}, \quad m = 1, 2, \ldots,$$
$$H(\mathrm{div},\Omega) = \{v \in L^2(\Omega;\mathbb{R}^n) :\ \mathrm{div}\, v \in L^2(\Omega)\},$$
$$H(\mathrm{curl},\Omega) = \begin{cases} \{v \in L^2(\Omega;\mathbb{R}^2) :\ \mathrm{rot}\, v \in L^2(\Omega)\}, & \Omega \subset \mathbb{R}^2, \\ \{v \in L^2(\Omega;\mathbb{R}^3) :\ \mathrm{curl}\, v \in L^2(\Omega;\mathbb{R}^3)\}, & \Omega \subset \mathbb{R}^3, \end{cases}$$

with derivatives taken in the distributional sense. The reference to the domain Ω will be omitted when appropriate and the associated norms will be denoted $||\cdot||_m$, $||\cdot||_{\mathrm{div}}$, and $||\cdot||_{\mathrm{curl}}$. Furthermore, we let $\mathbb{M}$ denote the space of matrices and $H(\mathrm{div},\Omega;\mathbb{M})$ denote the space of square-integrable matrix fields with square-integrable row-wise divergence.

For the sake of compact notation, we shall also adopt the exterior calculus notation of [**2**] and let $\Lambda^k(\Omega)$ denote the space of smooth differential k-forms on Ω, and let $L^2\Lambda^k(\Omega)$ denote the space of square integrable differential k-forms on Ω. We further let d denote the exterior derivative with adjoint δ, and define $H\Lambda^k(\Omega) = \{v \in L^2\Lambda^k(\Omega), \mathrm{d}\, v \in L^2\Lambda^k(\Omega)\}$. Further, $\mathcal{P}_r\Lambda^k$ is the space of polynomial k-forms of up to and including degree r, and $\mathcal{P}_r^-\Lambda^k$ denotes the reduced space as defined in [**2**, Section 3.3].

2.2. Examples. The function spaces $H(\mathrm{div})$ and $H(\mathrm{curl})$ are the natural function spaces for an extensive range of partial differential equations, in particular in mixed formulations. We sketch some examples in the following, both for motivational purposes and for later reference. The examples considered here are mixed formulations of the Hodge Laplace equations, the standard eigenvalue problem for Maxwell's equations and a mixed formulation for linear elasticity with weakly imposed symmetry. We return to these examples in Section 6.

EXAMPLE 2.1 (Mixed formulation of Poisson's equation). *The most immediate example involving the space $H(\mathrm{div})$ is a mixed formulation of Poisson's equation: $-\Delta u = f$ in $\Omega \subset \mathbb{R}^n$. By introducing the flux $\sigma = -\nabla u$ and assuming Dirichlet boundary conditions for u, we obtain the following mixed variational problem: Find $\sigma \in H(\mathrm{div},\Omega)$ and $u \in L^2(\Omega)$ satisfying*

$$\langle \tau, \sigma\rangle - \langle \mathrm{div}\, \tau, u\rangle + \langle v, \mathrm{div}\, \sigma\rangle = \langle v, f\rangle, \tag{2.1}$$

for all $\tau \in H(\mathrm{div},\Omega)$ and $v \in L^2(\Omega)$.

EXAMPLE 2.2 (The Hodge Laplacian). *In more generality, we may consider weak formulations of the Hodge Laplacian equation $(\mathrm{d}\,\delta + \delta\,\mathrm{d})u = f$ on a domain $\Omega \subset \mathbb{R}^n$, see* [**2**, Section 7]. *For simplicity of presentation, we assume that Ω is contractible such that the space of harmonic forms on Ω vanishes. The formulation in Example 2.1 is the equivalent of seeking $u \in H\Lambda^n$ and $\sigma = \delta\, u \in H\Lambda^{n-1}$ for $n = 2, 3$ with natural boundary conditions (the appropriate trace being zero). To see this, we test $\sigma = \delta\, u$ against $\tau \in H\Lambda^{n-1}$ and test $(\mathrm{d}\,\delta + \delta\,\mathrm{d})u = f$ against $v \in H\Lambda^n$ to obtain*

$$\langle \tau, \sigma\rangle - \langle \tau, \delta\, u\rangle + \langle v, \mathrm{d}\, \sigma\rangle = \langle v, f\rangle,$$

noting that $du = 0$ for $u \in H\Lambda^n$. Integrating by parts, we obtain

$$\langle \tau, \sigma\rangle - \langle \mathrm{d}\, \tau, u\rangle + \langle v, \mathrm{d}\, \sigma\rangle = \langle v, f\rangle. \tag{2.2}$$

We may restate (2.2) in the form (2.1) by making the identifications $\delta\, u = -\nabla\, u$, $\mathrm{d}\, \tau = \mathrm{div}\, \tau$, and $\mathrm{d}\, \sigma = \mathrm{div}\, \sigma$. If $\Omega \subset \mathbb{R}^3$, we may also consider the following mixed formulations of the Hodge Laplace equation:

(i) *Find $\sigma \in H\Lambda^1 = H(\mathrm{curl})$ and $u \in H\Lambda^2 = H(\mathrm{div})$ such that*

$$\langle \tau, \sigma\rangle - \langle \mathrm{curl}\, \tau, u\rangle + \langle v, \mathrm{curl}\, \sigma\rangle + \langle \mathrm{div}\, v, \mathrm{div}\, u\rangle = \langle v, f\rangle \tag{2.3}$$

for all $\tau \in H\Lambda^1$, $v \in H\Lambda^2$.

(ii) *Find* $\sigma \in H\Lambda^0 = H^1$ *and* $u \in H\Lambda^1 = H(\mathrm{curl})$ *such that*

$$\langle \tau, \sigma \rangle - \langle \nabla \tau, u \rangle + \langle v, \nabla \sigma \rangle + \langle \mathrm{curl}\, v, \mathrm{curl}\, u \rangle = \langle v, f \rangle \tag{2.4}$$

for all $\tau \in H\Lambda^0, v \in H\Lambda^1$.

Example 2.3 (Cavity resonator). *The time-harmonic Maxwell equations in a cavity with perfectly conducting boundary induces the following eigenvalue problem: Find resonances* $\omega \in \mathbb{R}$ *and eigenfunctions* $E \in H_0(\mathrm{curl}, \Omega)$, *satisfying*

$$\langle \mathrm{curl}\, F, \mathrm{curl}\, E \rangle = \omega^2 \langle F, E \rangle \quad \text{for all } F \in H_0(\mathrm{curl}, \Omega), \tag{2.5}$$

where $H_0(\mathrm{curl}, \Omega) = \{v \in H(\mathrm{curl}, \Omega) |\ v \times n|_{\partial\Omega} = 0\}$. *Note that the formulation (2.5) disregards the original divergence-free constraint for the electric field* E *and thus includes the entire kernel of the* curl *operator, corresponding to* $\omega = 0$ *and electric fields of the form* $E = \nabla \psi$.

Example 2.4 (Elasticity with weakly imposed symmetry). *Navier's equations for linear elasticity can be reformulated using the stress tensor* σ, *the displacement* u *and an additional Lagrange multiplier* γ *corresponding to the symmetry of the stress constraint. The weak equations for* $\Omega \subset \mathbb{R}^2$, *with the natural*[3] *boundary condition* $u|_{\partial\Omega} = 0$, *take the following form: Given* $f \in L^2(\Omega; \mathbb{R}^n)$, *find* $\sigma \in H(\mathrm{div}, \Omega; \mathbb{M})$, $u \in L^2(\Omega; \mathbb{R}^n)$, *and* $\gamma \in L^2(\Omega)$ *such that*

$$\langle \tau, A\sigma \rangle + \langle \mathrm{div}\, \tau, u \rangle + \langle v, \mathrm{div}\, \sigma \rangle + \langle \mathrm{skw}\, \tau, \gamma \rangle + \langle \eta, \mathrm{skw}\, \sigma \rangle = \langle v, f \rangle \tag{2.6}$$

for all $\tau \in H(\mathrm{div}, \Omega; \mathbb{M})$, $v \in L^2(\Omega; \mathbb{R}^n)$, *and* $\eta \in L^2(\Omega)$. *Here,* A *is the compliance tensor, and* $\mathrm{skw}\, \tau$ *is the scalar representation of the skew-symmetric component of* τ, *more precisely,* $2\,\mathrm{skw}\, \tau = \tau_{21} - \tau_{12}$. *This formulation has the advantage of being robust with regard to nearly incompressible materials and provides an alternative foundation for complex materials with non-local stress-strain relations. For more details, we refer the reader to* [**3**].

2.3. Continuity-preserving mappings for $H(\mathrm{div})$ and $H(\mathrm{curl})$. At this point, we turn our attention to a few results on continuity-preserving mappings for $H(\mathrm{div})$ and $H(\mathrm{curl})$. The results are classical and we refer to [**8, 26**] for a more thorough treatment.

First, it follows from Stokes' theorem that in order for piecewise $H(\mathrm{div})$ vector fields to be in $H(\mathrm{div})$ globally, the traces of the normal components over patch interfaces must be continuous and analogously tangential continuity is required for piecewise $H(\mathrm{curl})$ fields. More precisely, we have the following. Let $\mathcal{T}_h = \{K\}$ be a partition of Ω into subdomains. Define the space Σ_h of piecewise $H(\mathrm{div})$ functions relative to this partition $\mathcal{T}_h$:

$$\Sigma_h = \{\phi \in L^2(\Omega; \mathbb{R}^n) :\ \phi|_K \in H(\mathrm{div}, K) \ \text{ for all } K \in \mathcal{T}_h\}. \tag{2.7}$$

Then $\phi \in \Sigma_h$ is in $H(\mathrm{div}, \Omega)$ if and only if the normal traces of ϕ are continuous across all element interfaces. Analogously, if $\phi|_K \in H(\mathrm{curl}, K)$ for all $K \in \mathcal{T}_h$, then $\phi \in H(\mathrm{curl}, \Omega)$ if and only if the tangential traces are continuous across all element interfaces.

Second, we turn to consider a non-degenerate mapping $F : \Omega_0 \to F(\Omega_0) = \Omega$ with Jacobian $DF(X)$, $X \in \Omega_0 \subset \mathbb{R}^n$. For $\Phi \in H^m(\Omega_0)$, the mapping $\mathcal{F}$ defined by

$$\mathcal{F}(\Phi) = \Phi \circ F^{-1}, \tag{2.8}$$

[3]Note that the natural boundary condition in this mixed formulation is a Dirichlet condition, whereas for standard H^1 formulations the natural boundary condition would be a Neumann condition.

is an isomorphism from $H^m(\Omega_0)$ to $H^m(\Omega)$. This, however, is not the case for $H(\mathrm{div})$ or $H(\mathrm{curl})$, since $\mathcal{F}$ does not in general preserve continuity of normal nor tangential traces. Instead, one must consider the contravariant and covariant Piola mappings which preserve normal and tangential continuity respectively.

DEFINITION 2.5 (The contravariant and covariant Piola mappings). *Let $\Omega_0 \subset \mathbb{R}^n$, let F be a non-degenerate mapping from Ω_0 onto $F(\Omega_0) = \Omega$ with $J = DF(X)$, and let $\Phi \in L^2(\Omega_0; \mathbb{R}^n)$.*
The contravariant Piola mapping $\mathcal{F}^{\mathrm{div}}$ is defined by

$$\mathcal{F}^{\mathrm{div}}(\Phi) = \frac{1}{\det J} J\Phi \circ F^{-1}. \tag{2.9}$$

The covariant Piola mapping $\mathcal{F}^{\mathrm{curl}}$ is defined by

$$\mathcal{F}^{\mathrm{curl}}(\Phi) = J^{-T}\Phi \circ F^{-1}. \tag{2.10}$$

REMARK 2.6. *We remark that the contravariant Piola mapping is usually defined with an absolute value, $\mathcal{F}^{\mathrm{div}}(\Phi) = \frac{1}{|\det J|} J\Phi \circ F^{-1}$. However, omitting the absolute value, as in* (2.9)*, can simplify the assembly of $H(\mathrm{div})$ elements, as will be expounded in Section 5.*

The contravariant Piola mapping is an isomorphism of $H(\mathrm{div}, \Omega_0)$ onto $H(\mathrm{div}, \Omega)$ and the covariant Piola mapping is an isomorphism of $H(\mathrm{curl}, \Omega_0)$ onto $H(\mathrm{curl}, \Omega)$. In particular, the contravariant Piola mapping preserves normal traces and the covariant Piola mapping preserves tangential traces. We illustrate this below in the case of simplicial meshes in two and three space dimensions (triangles and tetrahedra).

EXAMPLE 2.7 (Piola mapping on triangles in $\mathbb{R}^2$). *Let K_0 be a triangle with vertices X^i and edges E^i for $i = 1, 2, 3$. We define the unit tangents by $T^i = E^i/||E^i||$. We further define the unit normals by $N^i = RT^i$ where*

$$R = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{2.11}$$

is the clockwise rotation matrix.

Now, assume that K_0 is affinely mapped to a (non-degenerate) simplex K with vertices x^i. The affine mapping $F_K : K_0 \to K$ takes the form $x = F_K(X) = JX + b$ and satisfies $x^i = F_K(X^i)$ for $i = 1, 2, 3$. It follows that edges are mapped by

$$e = x^i - x^j = J(X^i - X^j) = JE.$$

Similarly, normals are mapped by

$$||e||n = Re = RJE = (\det J)J^{-T}RE = (\det J)J^{-T}||E||N,$$

where we have used that $\frac{1}{\det J} RJR^T = J^{-T}$ and thus $RJ = (\det J)J^{-T}R$ for $J \in \mathbb{R}^{2\times 2}$.

The relation between the mappings of tangents and normals (or edges and rotated edges) may be summarized in the following commuting diagrams.

$$\begin{array}{ccc} T & \xrightarrow{J||E||/||e||} & t \\ {\scriptstyle R}\downarrow & & \downarrow{\scriptstyle R} \\ N & \xrightarrow{(\det J)J^{-T}||E||/||e||} & n \end{array} \qquad \begin{array}{ccc} E & \xrightarrow{J} & e \\ {\scriptstyle R}\downarrow & & \downarrow{\scriptstyle R} \\ ||E||N & \xrightarrow{(\det J)J^{-T}} & ||e||n \end{array} \tag{2.12}$$

With this in mind, we may study the effect of the Piola transforms on normal and tangential traces. Let $\Phi \in C^\infty(K_0; \mathbb{R}^n)$ and let $\phi = \mathcal{F}^{\mathrm{div}}(\Phi)$. Then

$$||e||\,\phi(x) \cdot n = ||e||\left((\det J)^{-1}J\Phi(X)\right)^T\left((\det J)J^{-T}||E||/||e||N\right) = ||E||\,\Phi(X) \cdot N.$$

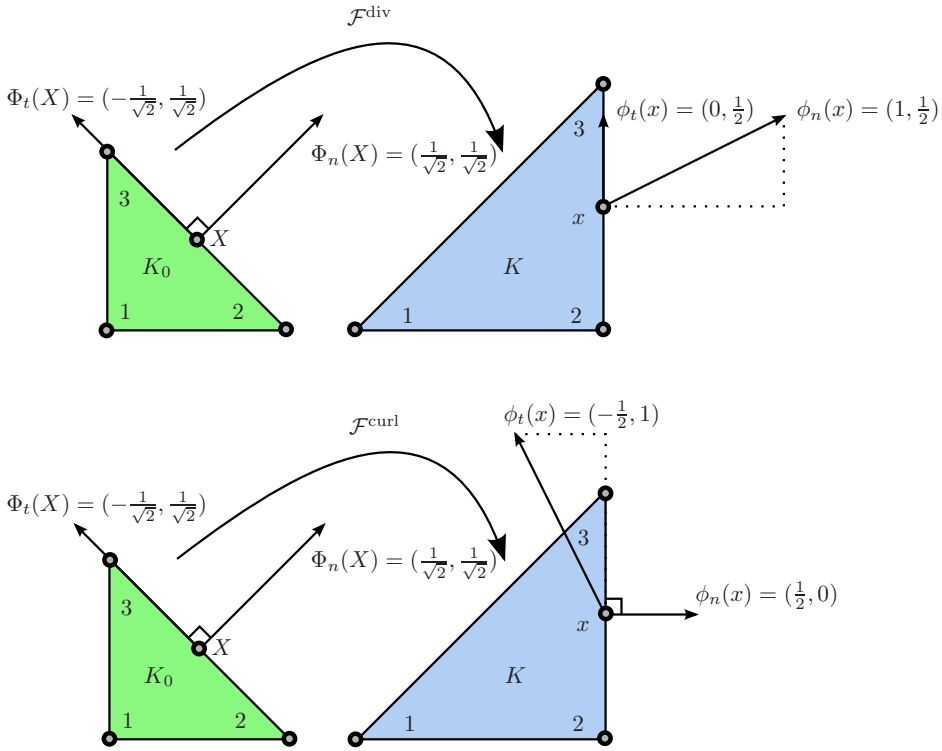


FIGURE 1. Mapping two vector fields Φ_n and Φ_t between two triangles using the contravariant and covariant Piola mappings. The contravariant Piola mapping (above) preserves normal traces of vector fields, and the covariant Piola mapping (below) preserves tangential traces of vector fields. This means in particular that the contravariant Piola mapping maps tangents to tangents (which have a zero normal component), and that the covariant Piola mapping maps normals to normals (which have a zero tangential component). Note that this is somewhat counter-intuitive; the contravariant $H(\mathrm{div})$ Piola mapping always maps tangential fields to tangential fields but does not in general map normal fields to normal fields. However, in both cases the normal component (being zero and one respectively) is preserved.

Thus, the contravariant Piola mapping preserves normal traces for vector fields under affine mappings, up to edge lengths. In general, the same result holds for smooth, non-degenerate mappings F_K if the Jacobian $DF_K(X)$ is invertible for all $X \in K_0$.

Similarly, let $\phi = \mathcal{F}^{\mathrm{curl}}(\Phi)$. Then

$$||e||\,\phi(x) \cdot t = ||e||\left(J^{-T}\Phi(X)\right)^T (J||E||/||e||) = ||E||\,\Phi(X) \cdot T. \tag{2.13}$$

Thus, the covariant Piola preserves tangential traces for vector fields, again up to edge lengths. Observe that the same result holds for tetrahedra without any modifications The effect of the contravariant and covariant Piola mappings on normal and tangential traces is illustrated in Figure 1, where $||E|| = ||e||$ for simplicity.

EXAMPLE 2.8 (Contravariant Piola mapping on tetrahedra in $\mathbb{R}^3$). *Now, let K_0 be a tetrahedron. As explained above, the covariant Piola mapping preserves tangential traces. To study the effect of the contravariant Piola mapping on normal traces, we define the face normals of K by $N = \frac{E^i \times E^j}{||E^i \times E^j||}$. Then*

$$||e^i \times e^j||n = JE^i \times JE^j = \det JJ^{-T}(E^i \times E^j) = ||E^i \times E^j|| \det JJ^{-T}N,$$

since $(Ju) \times (Jv) = \det JJ^{-T}(u \times v)$. Let $\Phi \in C^\infty(K_0; \mathbb{R}^n)$ and let $\phi = \mathcal{F}^{\mathrm{div}}(\Phi)$. Then, it follows that

$$||e^i \times e^j|| \, \phi(x) \cdot n = ||E^i \times E^j|| \, \Phi(X) \cdot N.$$

Thus, the contravariant Piola mapping preserves normal traces, up to the area of faces.

We finally remark that if $J \in \mathbb{R}^{2\times 2}$ defines a conformal, orientation-preserving map, the contravariant and covariant Piola mappings coincide. In $\mathbb{R}^3$, J must also be orthogonal for this to occur.

3. $H(\mathrm{div})$ AND $H(\mathrm{curl})$ CONFORMING FINITE ELEMENTS

To construct $H(\mathrm{div})$ and $H(\mathrm{curl})$ conforming finite element spaces, that is, discrete spaces V_h satisfying $V_h \subset H(\mathrm{div})$ or $V_h \subset H(\mathrm{curl})$, one may patch together local function spaces (finite elements) and make an appropriate matching of degrees of freedom over shared element facets[4]. In particular, one requires that degrees of freedom corresponding to normal traces match for $H(\mathrm{div})$ conforming discretizations and that tangential traces match for $H(\mathrm{curl})$ conforming discretizations. Several finite element spaces with degrees of freedom chosen to facilitate this exist, such as those in [**6, 7, 27, 28, 31**]. We summarize in Table 1 those $H(\mathrm{div})$ and $H(\mathrm{curl})$ conforming finite elements that are supported by FIAT and hence by FFC. In general, FFC can wield any finite element space that may be generated from a local basis through either of the afore described mappings. In Table 2, we also summarize the basic approximation properties of these elements for later comparison with numerical results in Section 6.

For the reasons above, it is common to define the degrees of freedom for each of the elements in Table 1 as moments of either normal or tangential traces over element facets. However, one may alternatively consider point values of traces at suitable points on element facets (in addition to any internal degrees of freedom). Thus, the degrees of freedom for the lowest order Raviart–Thomas space on a triangle may be chosen as the normal components at the edge midpoints, and for the lowest order Brezzi–Douglas–Marini space, we may consider the normal components at two points on each edge (positioned symmetrically on each edge and not touching the vertices). This, along with the appropriate scaling by edge length, is how the degrees of freedom are implemented in FIAT.

4. REPRESENTATION OF $H(\mathrm{div})$ AND $H(\mathrm{curl})$ VARIATIONAL FORMS

In this section, we discuss how multilinear forms on $H(\mathrm{div})$ or $H(\mathrm{curl})$ may be represented as a particular tensor contraction, allowing for pre-computation of integrals on a reference element and thus efficient assembly of linear systems. We

[4]We refer to any geometric entity of positive codimension in the mesh (such as an edge of a triangle or an edge or face of a tetrahedron) as a *facet*.

Simplex	$H(\mathrm{div})$			$H(\mathrm{curl})$		
$K \subset \mathbb{R}^2$	BDM_r	$\mathcal{P}_r\Lambda^1(K)$	[**7**]			
	RT_{r-1}	$\mathcal{P}_r^-\Lambda^1(K)$	[**31**]	NED_{r-1}	—	
	BDFM_r	—	[**6**]			
$K \subset \mathbb{R}^3$	$\mathrm{BDM}_r/\mathrm{NED}_r^2$	$\mathcal{P}_r\Lambda^2(K)$	[**28**]			
	$\mathrm{RT}_{r-1}/\mathrm{NED}_{r-1}^1$	$\mathcal{P}_r^-\Lambda^2(K)$	[**27**]	NED_{r-1}^1	$\mathcal{P}_r^-\Lambda^1(K)$	[**27**]
	BDFM_r	—				

TABLE 1. $H(\mathrm{div})$ and $H(\mathrm{curl})$ conforming finite elements on triangles and tetrahedra supported by FIAT and FFC for $r \geq 1$. When applicable, the elements are listed with their exterior calculus notation, along with their original references. Note that for $K \subset \mathbb{R}^3$, the Raviart–Thomas and Brezzi–Douglas–Marini elements are also known as the first and second kind $H(\mathrm{div})$ Nedelec (face) elements respectively.

Finite element	Interpolation estimates	
$\mathcal{P}_r\Lambda^k(\Omega)$	$\|\|v - \Pi_h v\|\|_0 \leq Ch^{m+1}\|\|v\|\|_m,$	$\|\|v - \Pi_h v\|\|_{\mathrm{div,curl}} \leq Ch^m\|\|v\|\|_m$
$\mathcal{P}_r^-\Lambda^k(\Omega)$	$\|\|v - \Pi_h v\|\|_0 \leq Ch^m\|\|v\|\|_m,$	$\|\|v - \Pi_h v\|\|_{\mathrm{div,curl}} \leq Ch^m\|\|v\|\|_m$
BDFM_r	$\|\|v - \Pi_h v\|\|_0 \leq Ch^m\|\|v\|\|_m,$	$\|\|v - \Pi_h v\|\|_{\mathrm{div}} \leq Ch^m\|\|v\|\|_m$

TABLE 2. Approximation properties of the spaces from Table 1. $C > 0$, $r \geq 1$, $1 \leq m \leq r$. It is assumed that $v \in H^m(\Omega; \mathbb{R}^n)$ and Π_h denotes the canonical interpolation operator onto the element space in question.

follow the notation from [**18, 24**] and extend the representation theorem from [**19**] for multilinear forms on H^1 and L^2 to $H(\mathrm{div})$ and $H(\mathrm{curl})$. The main new component is that we must use the appropriate Piola mapping to map basis functions from the reference element.

4.1. Multilinear forms and their representation. Let $\Omega \subset \mathbb{R}^n$ and let $\{V_h^j\}_{j=1}^\rho$ be a set of finite dimensional spaces associated with a tessellation $\mathcal{T} = \{K\}$ of Ω. We consider the following canonical linear variational problem: Find $u_h \in V_h^2$ such that

$$a(v, u_h) = L(v) \quad \forall v \in V_h^1, \tag{4.1}$$

where a and L are bilinear and linear forms on $V_h^1 \times V_h^2$ and V_h^1 respectively. Discretizing (4.1), one obtains a linear system $AU = b$ for the degrees of freedom U of the discrete solution u_h.

In general, we shall be concerned with the discretization of a general multilinear form of arity ρ,

$$a : V_h^1 \times V_h^2 \times \cdots \times V_h^\rho \to \mathbb{R}. \tag{4.2}$$

Typically, the arity is $\rho = 1$ (linear forms) or $\rho = 2$ (bilinear forms) but forms of higher arity also appear (see [**19**]). For illustration purposes, we consider the discretization of the mixed Poisson problem (2.1) in the following example.

EXAMPLE 4.1 (Discrete mixed Poisson). *Let Σ_h and W_h be discrete spaces approximating $H(\mathrm{div}, \Omega)$ and $L^2(\Omega)$ respectively. We may then write (2.1) in the*

canonical form (4.1) by defining

$$a((\tau_h, v_h), (\sigma_h, u_h)) = \langle \tau_h, \sigma_h \rangle - \langle \operatorname{div} \tau_h, u_h \rangle + \langle v_h, \operatorname{div} \sigma_h \rangle, \tag{4.3a}$$

$$L((\tau_h, v_h)) = \langle v_h, f \rangle, \tag{4.3b}$$

for $(\tau_h, v_h) \in V_h^1 = \Sigma_h \times W_h$ *and* $(\sigma_h, u_h) \in V_h^2 = V_h^1$.

To discretize the multilinear form (4.2), we let $\{\phi_k^j\}_{k=1}^{N_j}$ denote a basis for V_h^j for $j = 1, 2, \ldots, \rho$ and define the global tensor

$$A_i = a(\phi_{i_1}^1, \phi_{i_2}^2, \ldots, \phi_{i_\rho}^\rho), \tag{4.4}$$

where $i = (i_1, i_2, \ldots, i_\rho)$ is a multiindex. If the multilinear form is defined as an integral over $\Omega = \cup_{K \in \mathcal{T}_h} K$, the tensor A may be computed by assembling the contributions from all elements,

$$A_i = a(\phi_{i_1}^1, \phi_{i_2}^2, \ldots, \phi_{i_\rho}^\rho) = \sum_{K \in \mathcal{T}_h} a^K(\phi_{i_1}^1, \phi_{i_2}^2, \ldots, \phi_{i_\rho}^\rho). \tag{4.5}$$

where a^K denotes the contribution from element K. We further let $\{\phi_k^{K,j}\}_{k=1}^{n_j}$ denote the local finite element basis for V_h^j on K and define the *element tensor* A^K by

$$A_i^K = a^K(\phi_{i_1}^{K,1}, \phi_{i_2}^{K,2}, \ldots, \phi_{i_\rho}^{K,\rho}). \tag{4.6}$$

The assembly of the global tensor A thus reduces to the computation of the element tensor A^K on each element K and the insertion of the entries of A^K into the global tensor A.

In [**18**], it was shown that if the local basis on each element K may be obtained as the image of a basis on a *reference element* K_0 by the standard (affine) isomorphism $\mathcal{F}_K : H^1(K_0) \to H^1(K)$, then the element tensor A^K may be represented as a tensor contraction of a *reference tensor* A^0, only depending on the form a and the reference basis, and a *geometry tensor* G_K, depending on the geometry of the particular element K,

$$A_i^K = A_{i,\alpha}^0 G_K^\alpha, \tag{4.7}$$

with summation over the multiindex α. It was further demonstrated in [**21**], that this representation may significantly reduce the operation count for computing the element tensor compared to standard evaluation schemes based on quadrature. Below, we extend the representation (4.7) to hold not only for bases that may be affinely mapped from a reference element, but also for finite element spaces that must be transformed by a Piola mapping.

4.2. A representation theorem. We now state the general representation theorem for multilinear forms on H^1, $H(\text{curl})$, $H(\text{div})$ (and L^2). Instead of working out the details of the proof here, we refer the reader to the proof presented in [**19**] for H^1, and illustrate the main points for $H(\text{div})$ and $H(\text{curl})$ by a series of examples.

THEOREM 4.2. *Let* $K_0 \subset \mathbb{R}^n$ *be a reference element and let* $F_K : K_0 \to K = F_K(K_0)$ *be a non-degenerate, affine mapping with Jacobian* J_K. *For* $j = 1, 2, \ldots, \rho$, *let* $\{\phi_k^{K,j}\}_k$ *denote a basis on* K *generated from a reference basis* $\{\Phi_k^j\}_k$ *on* K_0, *that is,* $\phi_k^{K,j} = \mathcal{F}_K^j(\Phi_k^j)$ *where* $\mathcal{F}_K^j$ *is either of the mappings defined by* (2.8), (2.9) *or* (2.10).

Then there exists a reference tensor A_i^0, *independent of* K, *and a geometry tensor* G_K *such that* $A^K = A^0 : G_K$, *that is,*

$$A_i^K = \sum_{\alpha \in \mathcal{A}} A_{i\alpha}^0 G_K^\alpha \quad \forall i \in \mathcal{I}, \tag{4.8}$$

for a set of primary indices $\mathcal{I}$ *and* secondary indices $\mathcal{A}$. *In fact, the reference tensor* A^0 *takes the following canonical form,*

$$A^0_{i\alpha} = \sum \int_{K_0} \prod_j D_X^{(\cdot)} \Phi^j_{(\cdot)}[(\cdot)] \,\mathrm{d}X, \tag{4.9}$$

that is, it is the sum of integrals of products of basis function components and their derivatives on the reference element K_0, *and the geometry tensor* G_K *is the outer product of the coefficients* $c_{(\cdot)}$ *of any weight functions with a tensor that depends only on the Jacobian* J_K,

$$G^\alpha_K = \prod c_{(\cdot)} \frac{|\det J_K|}{(\det J_K)^\gamma} \sum \prod \frac{\partial X_{(\cdot)}}{\partial x_{(\cdot)}} \prod \frac{\partial x_{(\cdot)}}{\partial X_{(\cdot)}}. \tag{4.10}$$

4.3. Examples. To this end, we start by considering the vector-valued $L^2(\Omega)$ inner product, defining a bilinear form:

$$a(v, u) = \int_\Omega v \cdot u \,\mathrm{d}x. \tag{4.11}$$

In the following, we let x denote coordinates on K and let X denote coordinates on the reference element K_0. F_K is an affine mapping from K_0 to K, that is, $x = F_K(X) = J_K X + x_K$. We further let ϕ^K denote a field on K obtained as the image of a field Φ on the reference element K_0, $\phi^K = \mathcal{F}_K^{(\cdot)}(\Phi)$. We aim to illustrate the differences and similarities of the representations of the mass matrix for different choices of mappings $\mathcal{F}_K$, in particular affine, contravariant Piola, and covariant Piola.

Example 4.3 (The mass matrix with affinely mapped basis). *Let* $\mathcal{F}_K$ *be the affine mapping,* $\mathcal{F}_K(\Phi) = \Phi \circ F_K^{-1}$. *Then, the element matrix* A^K *for (4.11) is given by*

$$A^K_i = \int_K \phi^{K,1}_{i_1}(x) \cdot \phi^{K,2}_{i_2}(x) \,\mathrm{d}x = |\det J_K| \int_{K_0} \Phi^1_{i_1}[\beta](X)\, \Phi^2_{i_2}[\beta](X) \,\mathrm{d}X, \tag{4.12}$$

where we use $\Phi[\beta]$ *to denote component* β *of the vector-valued function* Φ *and implicit summation over the index* β. *We may thus represent the element matrix as the tensor contraction (4.7) with reference and geometry tensors given by*

$$\begin{aligned} A^0_i &= \int_{K_0} \Phi^1_{i_1}[\beta](X)\, \Phi^2_{i_2}[\beta](X) \,\mathrm{d}X, \\ G^K &= |\det J_K|. \end{aligned}$$

We proceed to examine the representation of the mass matrix when the basis functions are transformed with the contravariant and the covariant Piola transforms.

Example 4.4 (The mass matrix with contravariantly mapped basis). *Let* $\mathcal{F}^{\mathrm{div}}_K$ *be the contravariant Piola mapping,*

$$\mathcal{F}^{\mathrm{div}}_K(\Phi) = \frac{1}{\det J_K} J_K \Phi \circ F_K^{-1}.$$

Then, the element matrix A^K *for* (4.11) *is given by*

$$\begin{aligned} A^K_i &= \int_K \phi^{K,1}_{i_1}(x) \cdot \phi^{K,2}_{i_2}(x) \,\mathrm{d}x \\ &= \frac{|\det J_K|}{(\det J_K)^2} \frac{\partial x_\beta}{\partial X_{\alpha_1}} \frac{\partial x_\beta}{\partial X_{\alpha_2}} \int_{K_0} \Phi^1_{i_1}[\alpha_1](X)\, \Phi^2_{i_2}[\alpha_2](X) \,\mathrm{d}X. \end{aligned}$$

We may thus represent the element matrix as the tensor contraction (4.7) *with reference and geometry tensors given by*

$$A^0_{\alpha,i} = \int_{K_0} \Phi^1_{i_1}[\alpha_1](X)\, \Phi^2_{i_2}[\alpha_2](X)\, \mathrm{d}X,$$
$$G^K_\alpha = \frac{|\det J_K|}{(\det J_K)^2} \frac{\partial x_\beta}{\partial X_{\alpha_1}} \frac{\partial x_\beta}{\partial X_{\alpha_2}}.$$

EXAMPLE 4.5 (The mass matrix with covariantly mapped basis). *Let* $\mathcal{F}^{\mathrm{curl}}_K$ *be the covariant Piola mapping,*

$$\mathcal{F}^{\mathrm{curl}}_K(\Phi) = J_K^{-T} \Phi \circ F_K^{-1}.$$

Then, the element tensor (matrix) A^K *for* (4.11) *is given by*

$$\begin{aligned} A^K_i &= \int_K \phi^{K,1}_{i_1}(x) \cdot \phi^{K,2}_{i_2}(x)\, \mathrm{d}x \\ &= |\det J_K| \frac{\partial X_{\alpha_1}}{\partial x_\beta} \frac{\partial X_{\alpha_2}}{\mathrm{d}x_\beta} \int_{K_0} \Phi^1_{i_1}[\alpha_1](X)\, \Phi^2_{i_2}[\alpha_2](X)\, \mathrm{d}X. \end{aligned}$$

We may thus represent the element matrix as the tensor contraction (4.7) with reference and geometry tensors given by

$$A^0_{\alpha,i} = \int_{K_0} \Phi^1_{i_1}[\alpha_1](X)\, \Phi^2_{i_2}[\alpha_2](X)\, \mathrm{d}X,$$
$$G^K_\alpha = |\det J_K| \frac{\partial X_{\alpha_1}}{\partial x_\beta} \frac{\partial X_{\alpha_2}}{\partial x_\beta}.$$

We observe that the representation of the mass matrix differs for affine, contravariant Piola and covariant Piola. In particular, the geometry tensor is different for each mapping, and the reference tensor has rank two for the affine mapping, but rank four for the Piola mappings. We also note that the reference tensor for the mass matrix in the case of the covariant Piola mapping transforms in the same way as the reference tensor for the stiffness matrix in the case of an affine mapping (see [**18**]).

We conclude by demonstrating how the divergence term from (4.3) is transformed with the contravariant Piola (being the relevant mapping for $H(\mathrm{div})$).

EXAMPLE 4.6 (Divergence term). *Let* $\mathcal{F}_K$ *be the affine mapping, let* $\mathcal{F}^{\mathrm{div}}_K$ *be the contravariant Piola mapping, and consider the bilinear form*

$$a(v, \sigma) = \int_K v\ \mathrm{div}\, \sigma\, \mathrm{d}x. \tag{4.13}$$

for $(v, \sigma) \in V^1 \times V^2$. *Then, if* $\phi^{K,1} = \mathcal{F}_K(\Phi^1)$ *and* $\phi^{K,2} = \mathcal{F}^{\mathrm{div}}_K(\Phi^2)$, *the element matrix* A^K *for (4.13) is given by*

$$A^K_i = \int_K \phi^{K,1}_{i_1}\ \mathrm{div}\, \phi^{K,2}_{i_2}\, \mathrm{d}x = \frac{|\det J_K|}{\det J_K} \frac{\partial x_\beta}{\partial X_{\alpha_1}} \frac{\partial X_{\alpha_2}}{\partial x_\beta} \int_{K_0} \Phi^1_{i_1}\ \frac{\partial \Phi^2_{i_2}[\alpha_1]}{\partial X_{\alpha_2}}\, \mathrm{d}X.$$

Noting that $\frac{\partial x_\beta}{\partial X_{\alpha_1}} \frac{\partial X_{\alpha_2}}{\partial x_\beta} = \delta_{\alpha_1 \alpha_2}$, *we may simplify to obtain*

$$A^K_i = \frac{|\det J_K|}{\det J_K} \int_{K_0} \Phi^1_{i_1}\ \frac{\partial \Phi^2_{i_2}[\alpha_1]}{\partial X_{\alpha_1}}\, \mathrm{d}X = \pm \int_{K_0} \Phi^1_{i_1}\ \mathrm{div}\, \Phi^2_{i_2}\, \mathrm{d}X.$$

We may thus represent the element matrix as the tensor contraction (4.7) with reference and geometry tensors given by

$$A^0_i = \int_{K_0} \Phi^1_{i_1}\ \mathrm{div}\, \Phi^2_{i_2}\, \mathrm{d}X,$$
$$G^K_\alpha = \pm 1.$$

The simplification in the final example is a result of the isomorphism, induced by the contravariant Piola transform, between $H(\mathrm{div}, K_0)$ and $H(\mathrm{div}, K)$. The form compiler FFC takes special care of such and similar simplifications.

5. Assembling $H(\mathrm{div})$ and $H(\mathrm{curl})$ elements

To guarantee global continuity with Piola-mapped elements, special care has to be taken with regard to the numbering and orientation of geometric entities, in particular the interplay between local and global orientation. This is well-known, but is rarely discussed in the standard references[5]. We here discuss some of these issues and give a strategy for dealing with directions of normals and tangents that simplifies assembly over $H(\mathrm{div})$ and $H(\mathrm{curl})$. In fact, we demonstrate that one may completely remove the need for contending with directions by using an appropriate numbering scheme for the simplicial mesh.

5.1. Numbering scheme. The numbering and orientation of geometric entities in FFC follows the UFC specification [**1**]. In short, the numbering scheme works as follows. A global index is assigned to each vertex of the tessellation $\mathcal{T}_h$ (consisting of triangles or tetrahedra). If an edge adjoins two vertices v_i and v_j, we define the direction of the edge as going from vertex v_i to vertex v_j if $i < j$. This gives a unique orientation of each edge. The same convention is used locally to define the directions of the local edges on each element. Thus, if an edge adjoins the first and second vertices of a tetrahedron, then the direction is from the first to the second vertex. A similar numbering strategy is employed for faces. The key is now to require that the vertices of each element are always ordered based on the their global indices.

For illustration, consider first the two-dimensional case. Let K_0 be the UFC reference triangle, that is, the triangle defined by the vertices $\{(0,0),(1,0),(0,1)\}$. Assume that $K = F_K(K_0)$ and $K' = F_{K'}(K_0)$ are two physical triangles sharing an edge e with normal n. If e adjoins vertices v_i and v_j and is directed from v_i to v_j, it follows from the numbering scheme that $i < j$. Since the vertices of both K and K' are ordered based on their global indices, and the local direction (as seen from K or K') of an edge is based on the local indices of the vertices adjoining that edge, this means that the local direction of the edge e will agree with the global direction, both for K and K'. Furthermore, if we define edge normals as clock-wise rotated tangents, K and K' will agree on the direction of the normal of the common edge. The reader is encouraged to consult Figure 2 for an illustration.

The same argument holds for the direction of edges in three dimensions. Moreover, if face normals are consistently defined in terms of the edges, it is straightforward to ensure a common direction. In particular, FIAT uses the two first edges of each face to define the direction of the face normal. As a consequence, two adjacent tetrahedra sharing a common face will always agree on the direction of the normal of that face.

We emphasize that the numbering scheme above does not result in a consistent orientation of the boundary of each element. It does however ensure that two adjacent elements sharing a common edge or face will always agree on the orientation of that edge or face. In addition to facilitating the treatment of tangential and normal traces, a unique orientation of edges and faces simplifies assembly of higher order Lagrange elements. A similar numbering scheme is proposed in the monograph [**26**] for tetrahedra in connection with $H(\mathrm{curl})$ finite elements.

[5]Some details may be found in [**26, 32**],

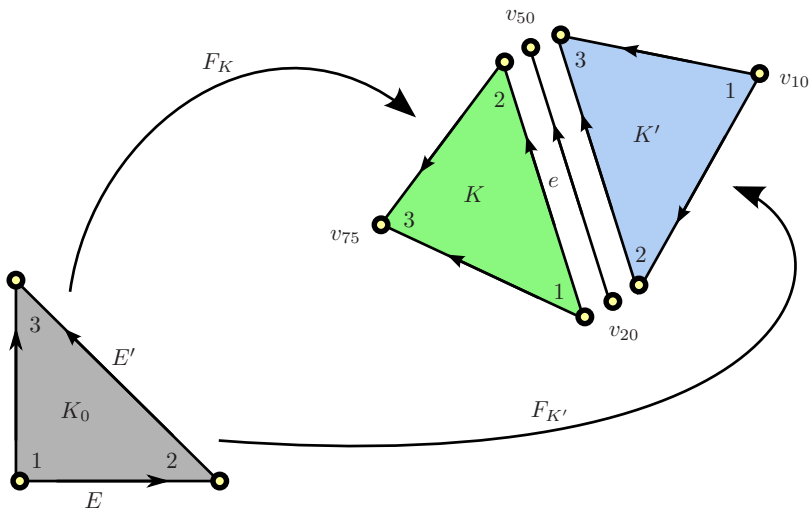


FIGURE 2. Two adjacent triangles will always agree on the direction of a common edge tangent or normal.

5.2. Mapping nodal basis functions. Next, we show how this numbering scheme and the FIAT choice of degrees of freedom give the necessary $H(\mathrm{div})$ or $H(\mathrm{curl})$ continuity. Assume that we have defined a set of nodal basis functions on K_0, that is, $\{\Phi_i\}_{i=1}^n$ such that

$$\ell_i(\Phi_j) = \delta_{ij}, \quad i, j = 1, 2, \ldots, n,$$

for a set of degrees of freedom $\{\ell_i\}_{i=1}^n$. These basis functions are mapped to two physical elements K and K' by an appropriate transformation $\mathcal{F}$ (contravariant or covariant Piola), giving a set of functions on K and K' respectively. We demonstrate below that as a consequence of the above numbering scheme, these functions will indeed be the restrictions to K and K' of an appropriate global nodal basis.

Consider $H(\mathrm{curl})$ and a global degree of freedom ℓ defined as the tangential component at a point x on a global edge e with tangent t, weighted by the length of the edge e,

$$\ell(v) = \|e\|\, v(x) \cdot t = v(x) \cdot e.$$

Let $\mathcal{F}^{\mathrm{curl}}$ be the covariant Piola mapping as before and let ϕ^K and $\phi^{K'}$ be two basis functions on K and K' obtained as the mappings of two nodal basis functions say Φ and Φ', on K_0,

$$\phi^K = \mathcal{F}_K^{\mathrm{curl}}(\Phi) \quad \text{and} \quad \phi^{K'} = \mathcal{F}_{K'}^{\mathrm{curl}}(\Phi').$$

Assume further that Φ is the nodal basis function corresponding to evaluation of the tangential component at the point $X \in K_0$ along the edge E, and that Φ' is the nodal basis function corresponding to evaluation of the tangential component at the point $X' \in K_0$ along the edge E'. Then, if $x = F_K(X) = F_{K'}(X')$, the covariant Piola mapping ensures that

$$\phi^K(x) \cdot e = \Phi(X) \cdot E = 1 \quad \text{and} \quad \phi^{K'}(x') \cdot e' = \Phi'(X') \cdot E' = 1.$$

Thus, since $e = e'$, it follows that

$$\ell(\phi^K) = \ell(\phi^{K'}).$$

Continuity for $H(\mathrm{div})$ may be demonstrated similarly.

In general, FFC allows elements for which the nodal basis on the reference element K_0 is mapped exactly to the nodal basis for each element K under some mapping $\mathcal{F}$, whether this be affine change of coordinates or one of the Piola transformations. While this enables a considerable range of elements, as considered in this paper, it leaves out many other elements of interest. As an example, the Hermite triangle or tetrahedron [**10**], does not transform equivalently. The Hermite triangle has degrees of freedom which are point values at the vertices and the barycenter, and the partial derivatives at each vertex. Mapping the basis function associated with a vertex point value affinely yields the correct basis function for K, but not for the derivative basis functions. A simple calculation shows that a function with unit x-derivative and vanishing y-derivative at a point generally maps to a function for which this is not the case. In fact, the function value basis functions transform affinely, but the pairs of derivative basis functions at each vertex must be transformed together, that is, a linear combination of their image yields the correct basis functions.

Examples of other elements requiring more general types of mappings include the scalar-valued Argyris and Morley elements as well as the Arnold-Winther symmetric elasticity element [**4**] and the Mardal-Tai-Winther element for Darcy-Stokes flow [**25**]. Recently, a special-purpose mapping for the Argyris element has been developed by Dominguez and Sayas [**11**], and we are generalizing this work as an extension of the FIAT project as outlined below.

If $\{\Phi_i\}$ is the reference finite element basis and $\{\phi_i^K\}_i$ is the physical finite element basis, then equivalent elements satisfy $\phi_i^K = \mathcal{F}_K(\Phi_i)$ for each i. If the elements are not equivalent under $\mathcal{F}_K$, then $\{\phi_i^K\}_i$ and $\{\mathcal{F}_K(\Phi_i)\}_i$ form two different bases for the polynomial space. Consequently, there exists a matrix M^K such that $\phi_i^K = \sum_j M_{ij}^K \mathcal{F}_K(\Phi_j)$. In the future, we hope to extend FIAT to construct this matrix M and FFC to make use of it in constructing variational forms, further extending the range of elements available to users.

5.3. A note about directions. An alternative orientation of shared facets gives rise to a special case of such transformations. It is customary to direct edges in a fashion that gives a consistent orientation of the boundary of each triangle. However, this would necessarily mean that two adjacent triangles disagree on the direction of their common edge. In this setting, normals would naturally be directed outward from each triangle, which again would imply that two adjacent triangles disagree on the direction of the normal on a common edge. It can be demonstrated that it is then more appropriate to define the contravariant Piola mapping in the following slightly modified form,

$$\mathcal{F}^{\mathrm{div}}(\Phi) = \frac{1}{|\det J_K|} J_K \Phi \circ F_K^{-1},$$

that is, the determinant of the Jacobian appears without a sign.

To ensure global continuity, one would then need to introduce appropriate sign changes for the mapped basis functions. For two corresponding basis functions ϕ^K and $\phi^{K'}$ as above, one would change the sign of $\phi^{K'}$ or ϕ^K such that both basis functions correspond to the same global degree of freedom. Thus, one may consider obtaining the basis functions on the physical element by first mapping the nodal basis functions from the reference element, and then correcting those basis functions with a change of sign,

$$\tilde{\phi}^K = \mathcal{F}(\Phi),$$
$$\phi^K = \pm \tilde{\phi}^K.$$

TABLE 3. FFC code for the mixed Poisson equation.

```
r = 3
S = FiniteElement("BDM", "triangle", r)
V = FiniteElement("DG", "triangle", r - 1)
element = S + V

(tau, v) = TestFunctions(element)
(sigma, u) = TrialFunctions(element)

a = (dot(tau, sigma) - dot(div(tau), u) + dot(v, div(sigma))*dx
L = dot(v, f)*dx
```

This would correspond to a diagonal M^K transformation where the entries are all ± 1.

Since a multilinear form is linear in each of its arguments, this approach corresponds to first computing a tentative element tensor $\tilde{A}^K$ and then obtaining A^K from $\tilde{A}^K$ by a series of rank one transforms. However, this procedure is unnecessary if the contravariant Piola mapping is defined as in (2.9) and the numbering scheme described in Section 5.1 is employed.

6. EXAMPLES

In order to demonstrate the veracity of the implementation, and the ease with which the $H(\mathrm{div})$ and $H(\mathrm{div})$ conforming elements can be employed, we now present a set of numerical examples and include the FFC code used to define the variational forms. In particular, we return to the examples introduced in Section 2 which include formulations of the Hodge Laplace equations, the cavity resonator eigenvalue problem and the weak symmetry formulation for linear elasticity.

6.1. The Hodge Laplacian. Consider the weak formulations of the Hodge Laplace equation introduced in Examples 2.1 and 2.2. For $\Omega \subset \mathbb{R}^2$ and differential 1- and 2-forms, we have the mixed Poisson equation (2.1). Stable choices of conforming finite element spaces $\Sigma_h \times V_h \subset H(\mathrm{div}) \times L^2$ include $V_h = \mathrm{DG}_{r-1}$ in combination with $\Sigma_h \in \{\mathrm{RT}_{r-1}, \mathrm{BDFM}_r, \mathrm{BDM}_r\}$ for $r = 1, 2, \ldots$. The FFC code corresponding to the latter choice of elements is given in Table 3. Further, for $\Omega \subset \mathbb{R}^3$, we give the FFC code for the formulation of (2.3) with the element spaces $\mathrm{NED}^1_{r-1} \times \mathrm{RT}_{r-1} \subset H(\mathrm{curl}) \times H(\mathrm{div})$ in Table 4.

For testing purposes, we consider a regular tessellation of the unit square/cube, $\Omega = [0,1]^n$, $n = 2, 3$, and a given smooth source for the two formulations. In particular for (2.1), we solve for

$$u(x_1, x_2) = C \sin(\pi x_1) \sin(\pi x_2), \tag{6.1}$$

with C a suitable scaling factor, and for (2.3), we let

$$u(x_1, x_2, x_3) = \begin{pmatrix} x_1^2(x_1-1)^2 \sin(\pi x_2) \sin(\pi x_3) \\ x_2^2(x_2-1)^2 \sin(\pi x_1) \sin(\pi x_3) \\ x_3^2(x_3-1)^2 \sin(\pi x_1) \sin(\pi x_2) \end{pmatrix}. \tag{6.2}$$

Note that u given by (6.2) is divergence-free and such that $u \times n = 0$ on the exterior boundary, and thus satisfies the implicit natural boundary conditions of (2.3).

A comparison of the exact and the approximate solutions for a set of uniformly refined meshes gives convergence rates in perfect agreement with the theoretical values indicated by Table 2, up to a precision limit. Logarithmic plots of the L^2

TABLE 4. FFC code for the curl-div formulation of the Hodge Laplace equation.

```
r = 2
CURL = FiniteElement("Nedelec", "tetrahedron", r - 1)
DIV = FiniteElement("RT", "tetrahedron", r - 1)
element = CURL + DIV

(tau, v) = TestFunctions(element)
(sigma, u) = TrialFunctions(element)

a = (dot(tau, sigma) - dot(curl(tau), u) + dot(v, curl(sigma)) \
    + dot(div(v), div(u)))*dx
L = dot(v, f)*dx
```

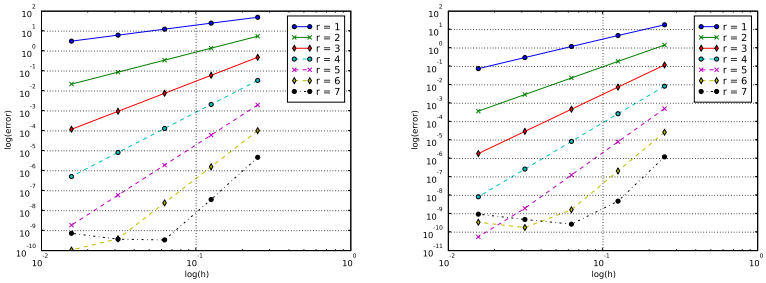


FIGURE 3. Convergence rates for the discretized mixed Poisson equation (2.1) using $\mathrm{RT}_{r-1} \times \mathrm{DG}_{r-1}$ (left) and $\mathrm{BDM}_r \times DG_{r-1}$ (right), r = 1, 2, ..., 7. Logarithmic plots of the L^2 error of the flux approximation: $||\sigma - \sigma_h||_0$ versus mesh size. The convergence rates in the left plot are $\mathcal{O}(h^r)$ and the convergence rates in the right plot are $\mathcal{O}(h^{r+1})$, cf. Table 2. The error does not converge below $\sim 10^{-10}$ in our experiments as a result of limited precision in the evaluation of integrals and/or linear solvers. The exact source of the limited precision has not been investigated in detail.

error of the flux using $\Sigma_h \in \{\mathrm{RT}_{r-1}, \mathrm{BDM}_r\}$ versus the mesh size for $r = 1, 2, \ldots, 7$, can be inspected in Figure 3 for the mixed Poisson problem (with $C = 100$).

For the curl-div formulation of the Hodge Laplace equation (2.3), we have included convergence rates for u and σ in Table 5. Note that the convergence rates for the combinations $\mathrm{NED}_r \times \mathrm{RT}_r$ and $\mathrm{NED}_r \times \mathrm{BDM}_r$, $r = 1, 2$ are of the same order, except for the $||\cdot||_{\mathrm{div}}$ error of u, though the former combination is computationally more expensive.

6.2. The cavity resonator. The analytical non-zero eigenvalues of the Maxwell eigenvalue problem (2.5) with $\Omega = [0, \pi]^n$, $n = 2, 3$, are given by

$$\omega^2 = m_1^2 + m_2^2 + \ldots + m_n^2, \quad m_i \in \{0\} \cup \mathbb{N} \tag{6.3}$$

where at least $n - 1$ of the terms m_i must be nonzero. It is well-known [**26**] that discretizations of this eigenvalue problem using H^1 conforming finite elements produce spurious and highly mesh-dependent eigenvalues ω^2. The edge elements of

Element	$\|\|\sigma - \sigma_h\|\|_0$	$\|\|\sigma - \sigma_h\|\|_{\text{curl}}$	$\|\|u - u_h\|\|_0$	$\|\|u - u_h\|\|_{\text{div}}$
$\text{NED}_0 \times \text{RT}_0$	0.99	0.98	0.99	0.98
$\text{NED}_1 \times \text{BDM}_1$	1.96	2.00	1.95	0.96
$\text{NED}_1 \times \text{RT}_1$	1.97	1.97	1.98	1.98
$\text{NED}_2 \times \text{BDM}_2$	3.00	2.99	2.97	1.97
$\text{NED}_2 \times \text{RT}_2$	2.98	2.96	2.97	2.97

TABLE 5. Averaged convergence rates for the discretized curl-div formulation of the Hodge Laplace equation (2.3) using $\text{NED}_{r-1} \times \text{RT}_{r-1}$, $r = 1, 2, 3$ and $\text{NED}_r \times \text{BDM}_r$, $r = 1, 2$. Number of degrees of freedom in the range $80\,000 - 300\,000$.

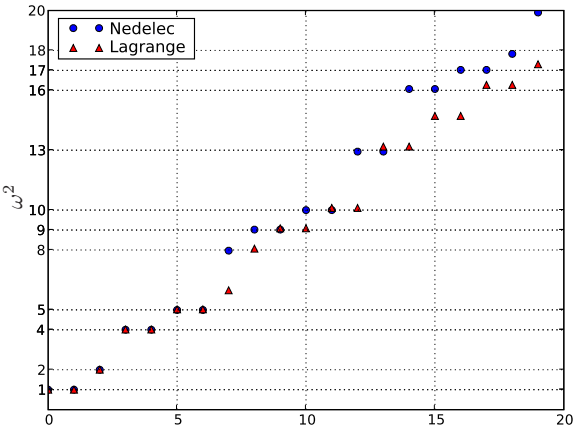


FIGURE 4. The first 20 eigenvalues of the cavity resonator problem computed using first order Nedelec elements (NED_0) and Lagrange elements (P_1) on a coarse (16×16) criss-cross mesh. The exact analytical values are indicated by the horizontal grid lines.

the Nedelec type however give convergent approximations of the eigenvalues. This phenomenon is illustrated in Figure 4. There, the first 20 non-zero eigenvalues $\omega_{h,N}^2$ produced by the Nedelec edge elements on a regular criss-cross triangulation are given in comparison with the corresponding Lagrange eigenvalue approximations $\omega_{h,L}^2$. Note the treacherous spurious Lagrange approximations such as $\omega_{h,L}^2 \approx 6, 15$.

6.3. Elasticity with weakly imposed symmetry. As a final example, we consider a mixed finite element formulation of the equations of linear elasticity with the symmetry of the stress tensor imposed weakly as given in Example 2.4. In the homogeneous, isotropic case, the inner product induced by the compliance tensor A reduces to

$$\langle \tau, A\sigma \rangle = \nu \langle \tau, \sigma \rangle - \zeta \langle \operatorname{tr} \tau, \operatorname{tr} \sigma \rangle$$

for ν, ζ material parameters. A stable family of finite element spaces for the discretization of (2.6) is given by [**3**]: $\text{BDM}_r^2 \times \text{DG}_{r-1}^2 \times \text{DG}_{r-1} \subset H(\text{div}, \Omega; \mathbb{M}) \times$

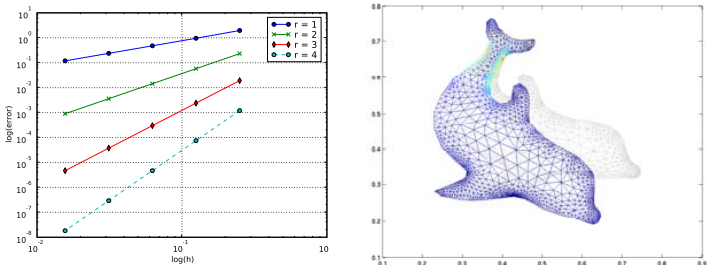


FIGURE 5. Left: Convergence rates for elastic stress approximations of (2.6). Logarithmic plot of $H(\mathrm{div})$ error of the approximated stress σ versus mesh size. The convergence rates are $\mathcal{O}(h^r)$, $r = 1, 2, 3, 4$. Right: Elastic dolphin hanging by the tail under a gravitational force.

TABLE 6. FFC code for linear elasticity with weak symmetry.

```
def A(sigma, tau, nu, zeta):
    return (nu*dot(sigma, tau) - zeta*trace(sigma)*trace(tau))*dx

def b(tau, w, eta):
    return (div(tau[0])*w[0] + div(tau[1])*w[1] + skew(tau)*eta)*dx

nu =  0.5
zeta =  0.2475
r = 2

S = FiniteElement("BDM", "triangle", r)
V = VectorElement("Discontinuous Lagrange", "triangle", r-1)
Q = FiniteElement("Discontinuous Lagrange", "triangle", r-1)
MX = MixedElement([S, S, V, Q])

(tau0, tau1, v, eta) = TestFunctions(MX)
(sigma0, sigma1, u, gamma) = TrialFunctions(MX)
sigma = [sigma0, sigma1]
tau = [tau0, tau1]

a = A(sigma, tau, nu, zeta) + b(tau, u, gamma) + b(sigma, v, eta)
L = dot(v, f)*dx
```

$L^2(\Omega; \mathbb{R}^n) \times L^2(\Omega)$, $r = 1, 2, \ldots$. The 17 lines of FFC code sufficient to define this discretization are included in Table 6.

Again to demonstrate convergence, we consider a regular triangulation of the unit square and solve for the smooth solution

$$u(x_0, x_1) = \begin{pmatrix} -x_1 \sin(\pi x_0) \\ 0.5\pi x_1^2 \cos(\pi x_0) \end{pmatrix}. \tag{6.4}$$

The theoretically predicted convergence rate of the discretization introduced above is of the order $\mathcal{O}(h^r)$ for all computed quantities. The numerical experiments corroborate this prediction. In particular, the convergence of the stress approximation in the $H(\mathrm{div})$ norm can be examined in Fig. 5.

7. Conclusions

The relative scarcity of $H(\mathrm{div})$ and $H(\mathrm{curl})$ mixed finite element formulations in practical use, may be attributed to their higher theoretical and implementational threshold. Indeed, more care is required to implement their finite element basis functions than the standard Lagrange bases, and assembly poses additional difficulties.

However, as demonstrated in this work, the implementation of mixed finite element formulations over $H(\mathrm{div})$ and $H(\mathrm{curl})$ may be automated and thus be used with the same ease as standard formulations over H^1. In particular, the additional challenges in the assembly can be viewed as not essentially different from those encountered when assembling higher-order Lagrange elements. The tools (FFC, FIAT, DOLFIN) used to compute the results presented here are freely available as part of the FEniCS project [**13**] and it is our hope that this may contribute to further the use of mixed formulations in applications.

Centre of Mathematics for Applications, P.O.Box 1053, 0316 Oslo, Norway. meg@cma.uio.no

Texas Tech University, Department of Mathematics and Statistics, P.O. Box 1042, Lubbock, TX, 79409–1042, USA robert.c.kirby@ttu.edu
This work is supported by the United States Department of Energy Office of Science under grant number DE–FG02–07ER25821.

Center for Biomedical Computing, Simula Research Laboratory / Department of Informatics, University of Oslo, P.O. Box 134, 1325 Lysaker, Norway. logg@simula.no
This work is supported by a Center of Excellence grant from the Research Council of Norway to the Center for Biomedical Computing at Simula Research Laboratory. This work is also supported by an Outstanding Young Investigator grant from the Research Council of Norway, NFR 180450.

Bibliography

[1] M. Alnæs, A. Logg, K.-A. Mardal, O. Skavhaug, and H. P. Langtangen. *UFC Specification and User Manual 1.1*, 2008. URL: `http://www.fenics.org/ufc/`.

[2] D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus, homological techniques, and applications. *Acta numerica*, 2006.

[3] D. N. Arnold, R. S. Falk, and R. Winther. Mixed finite element methods for elasticity with weakly imposed symmetry. *Math. Computation*, 2007.

[4] D. N. Arnold and R. Winther. Mixed finite elements for elasticity. *Numer. Math.*, 92(3):401–419, 2002.

[5] W. Bangerth, R. Hartmann, and G. Kanschat. `deal.II` Differential Equations Analysis Library, 2006. URL: `http://www.dealii.org/`.

[6] F. Brezzi, J. Douglas, Jr., M. Fortin, and L. D. Marini. Efficient rectangular mixed finite elements in two and three space variables. *RAIRO Modél. Math. Anal. Numér.*, 21(4):581–604, 1987.

[7] F. Brezzi, J. Douglas, Jr., and L. D. Marini. Two families of mixed finite elements for second order elliptic problems. *Numer. Math.*, 47(2):217–235, 1985.

[8] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.

[9] P. Castillo, R. Rieben, and D. White. Femster: an object-oriented class library of discrete differential forms. *to appear in ACM Trans. Math. Software*, 2005.

[10] P. G. Ciarlet. *Numerical Analysis of the Finite Element Method*. Les Presses de l'Universite de Montreal, 1976.

[11] V. Domínguez and F.-J. Sayas. Algorithm 884: A simple Matlab implementation of the Argyris element. *ACM Transactions on Mathematical Software*, 35(2):16, July 2008. Article 16, 11 pages.

[12] FEAP. FEAP A Finite Element Analysis Program. URL: `http://www.ce.berkeley.edu/~rlt/feap/`.

[13] J. Hoffman, J. Jansson, C. Johnson, M. G. Knepley, R. C. Kirby, A. Logg, L. R. Scott, and G. N. Wells. *FEniCS*, 2006. `http://www.fenics.org/`.

[14] R. C. Kirby. Algorithm 839: FIAT, a new paradigm for computing finite element basis functions. *ACM Transactions on Mathematical Software*, 30(4):502–516, Dec. 2004.

[15] R. C. Kirby. *FIAT*, 2006. URL: `http://www.fenics.org/fiat/`.

[16] R. C. Kirby. Optimizing FIAT with Level 3 BLAS. *ACM Transactions on Mathematical Software*, 32(2):223–235, June 2006.

[17] R. C. Kirby, M. G. Knepley, A. Logg, and L. R. Scott. Optimizing the evaluation of finite element matrices. *SIAM J. Sci. Comput.*, 27(3):741–758, 2005.

[18] R. C. Kirby and A. Logg. A compiler for variational forms. *ACM Transactions on Mathematical Software*, 32(9):417–444, 2006.

[19] R. C. Kirby and A. Logg. Efficient compilation of a class of variational forms. *ACM Transactions on Mathematical Software*, 33(10), 2007.

[20] R. C. Kirby and A. Logg. Benchmarking domain-specific compiler optimizations for variational forms. *ACM Transactions on Mathematical Software*, 35(2):10, July 2008. Article 10, 18 pages.

[21] R. C. Kirby, A. Logg, L. R. Scott, and A. R. Terrel. Topological optimization of the evaluation of finite element matrices. *SIAM J. Sci. Comput.*, 28(1):224–240, 2006.

[22] R. C. Kirby and L. R. Scott. Geometric optimization of the evaluation of finite element matrices. *SIAM J. Sci. Comput.*, 29(2):827–841, 2007.
[23] A. Logg. *FFC*, 2006. `http://www.fenics.org/ffc/`.
[24] A. Logg. Automating the finite element method. *Arch. Comput. Methods Eng.*, 14(11):93–138, 2007.
[25] K. A. Mardal, X.-C. Tai, and R. Winther. A robust finite element method for Darcy-Stokes flow. *SIAM J. Numer. Anal.*, 40(5):1605–1631 (electronic), 2002.
[26] P. Monk. *Finite Element Methods for Maxwell's Equations*. Oxford Science Publications, 2003.
[27] J.-C. Nédélec. Mixed finite elements in $\mathbf{R}^3$. *Numer. Math.*, 35(3):315–341, 1980.
[28] J.-C. Nédélec. New mixed finite elements in $\mathbf{R}^3$. *Numer. Math.*, 50(1):57–81, 1986.
[29] K. Oelgaard, A. Logg, and G. N. Wells. Automated code generation for discontinuous galerkin methods. *To appear in SIAM J. Sci. Comput.*, 2008.
[30] O. Pironneau, F. Hecht, A. L. Hyaric, and K. Ohtsuka. FreeFEM, 2006. URL: `http://www.freefem.org/`.
[31] P.-A. Raviart and J. M. Thomas. Primal hybrid finite element methods for 2nd order elliptic equations. *Math. Comp.*, 31(138):391–413, 1977.
[32] A. Schneebeli. An $H(\mathrm{curl};\omega)$ conforming FEM: Nedelec's element of the first type. Technical report, 2003.
[33] J. Schöberl. `NGSolve, 2008. URL: http://www.hpfem.jku.at/ngsolve/index.html/.`

Paper IV

STABILITY OF LAGRANGE FINITE ELEMENTS FOR THE MIXED LAPLACIAN

MARIE E. ROGNES AND DOUGLAS N. ARNOLD

Abstract. The stability properties of continuous piecewise polynomial approximations $\mathcal{P}_r^c$ (Lagrange elements) for the vector variable, and discontinuous piecewise polynomial approximations of one order lower $\mathcal{P}_{r-1} \supseteq \operatorname{div} \mathcal{P}_r^c$ for the scalar field, of the mixed Laplacian are investigated numerically for $r = 1, 2, 3$. For $r = 2, 3$, the element pair $\mathcal{P}_r^c \times \operatorname{div} \mathcal{P}_r^c$ is stable for all mesh families tested. In particular, it is stable on diagonal mesh families, in contrast to its behaviour for the Stokes equations. For $r = 1$, stability holds for some meshes, but not for others. Moreover, in $H(\operatorname{div}) \times L^2$, the natural setting for this problem, convergence was observed precisely for the methods that were observed to be stable. However, it seems that optimal order L^2 estimates for the vector variable, known to hold for $r \geq 4$, are not attainable for $r < 4$.

1. Introduction

In this note, we consider approximations of the mixed Laplace equations with Dirichlet boundary conditions: Given a source g, find the velocity u and the pressure p such that

$$u - \nabla p = 0, \ \operatorname{div} u = g \ \text{ in } \Omega, \qquad p = 0 \ \text{ on } \partial\Omega$$

for a domain $\Omega \subset \mathbb{R}^2$ with boundary $\partial\Omega$. The equations offer the classical weak formulation: Find a square integrable vector field with square integrable divergence $u \in H(\operatorname{div}, \Omega)$ and a square integrable function $p \in L^2(\Omega)$ such that

$$\int_\Omega u \cdot v + q \operatorname{div} u + p \operatorname{div} v = \int_\Omega g\, q \tag{1.1}$$

for all $v \in H(\operatorname{div}, \Omega)$ and $q \in L^2(\Omega)$. The above formulation can be discretized using a pair of finite dimensional spaces $V_h \times Q_h$, yielding discrete approximations $u_h \in V_h$ and $p_h \in Q_h$ satisfying (1.1) for all $v \in V_h$ and $q \in Q_h$.

As is well-known, the spaces V_h and Q_h must satisfy certain stability, or compatibility, conditions for the discretization to be well-behaved [**7**]. More precisely,

there must exist positive constants α and β such that for any h,

$$0 < \alpha \leq \alpha_h = \inf_{u \in Z_h} \sup_{v \in Z_h} \frac{\langle u, v \rangle}{||u||_{\mathrm{div}} ||v||_{\mathrm{div}}}, \tag{1.2a}$$

$$0 < \beta < \beta_h^{\mathrm{div}} = \inf_{q \in Q_h} \sup_{v \in V_h} \frac{\langle \mathrm{div}\, v, q \rangle}{||v||_{\mathrm{div}} ||q||_0}. \tag{1.2b}$$

Here, $||\cdot||_{\mathrm{div}}$ and $||\cdot||_0$ denote the norms on $H(\mathrm{div}, \Omega)$ and $L^2(\Omega)$ respectively, $\langle \cdot, \cdot \rangle$ is the $L^2(\Omega)$ inner product and

$$Z_h = \{ v \in V_h \mid \langle \mathrm{div}\, v, q \rangle = 0 \quad \forall\, q \in Q_h \}. \tag{1.3}$$

The two conditions will be referred to as the Brezzi coercivity and the Brezzi inf-sup condition for the mixed Laplacian respectively. The classical conforming discretizations of (1.1) rely on the finite element families of Raviart and Thomas [**18**] or Brezzi, Douglas and Marini [**8**] for the space $V_h \subset H(\mathrm{div})$ in order to satisfy these conditions.

In this note, we shall consider the Lagrange vector element spaces, that is, continuous piecewise polynomial vector fields defined relative to a triangulation $\mathcal{T}_h$, for the space V_h. This is motivated by the following reasons. First, these spaces are fairly inexpensive, simple to implement and post-process, and in frequent use for other purposes. Second, such pairs would allow continuous approximations of the velocity variable, or when viewed in connection with linear elasticity, lay the ground for continuous approximations of the stress tensor. Moreover, in the recent years, there has been an interest in mixed finite element discretizations that are both stable for (1.1) and for the Stokes equations:

$$\int_\Omega \nabla u : \nabla v + q\ \mathrm{div}\, u + p\ \mathrm{div}\, v = \int_\Omega f\, v, \tag{1.4}$$

for all $v \in H^1(\Omega; \mathbb{V})$ such that $\int_\Omega v = 0$ and all $q \in L^2(\Omega)$. The search for conforming such discretizations is complicated by the fact that the existing, stable discretizations of (1.1) are such that $V_h \not\subset H^1(\Omega; \mathbb{V})$. On the other hand, the existing stable discretizations of (1.4) are typically unstable for (1.1) [**14**]. The existence of stable discretizations $\{V_h \times Q_h\}_h$ of (1.1) such that $V_h \subset H^1(\Omega; \mathbb{V})$ becomes a natural separate question. Unfortunately, there are no known such finite element discretizations that are stable for any admissible triangulation family $\{\mathcal{T}_h\}$. In this note, we aim to numerically examine cases where a reduced stability property may be identified. In this sense, the investigations here are in the spirit of the work of Chapelle and Bathe [**10**], and Qin [**17**].

For a family of conforming discretizations $\{V_h \times Q_h\}_h$ of (1.1) such that $\mathrm{div}\, V_h \subseteq Q_h$ for each h, the condition (1.2a) is trivial. The stability conditions thus reduce to the condition (1.2b), namely the question of bounded Brezzi inf-sup constant β_h^{div}. On the other hand, recall that for the Stokes formulation (1.4), the corresponding Brezzi coercivity condition is trivial by Korn's inequality. Hence, for any family of conforming discretizations, the stability conditions for Stokes reduce to that of a uniform bound for the Brezzi inf-sup constant β_h^1. Here,

$$\beta_h^1 = \inf_{q \in Q_h} \sup_{v \in V_h} \frac{\langle \mathrm{div}\, v, q \rangle}{||v||_1 ||q||_0}, \tag{1.5}$$

when $V_h \subset H^1(\Omega; \mathbb{V})$, $Q_h \subset L^2(\Omega)$, and $||\cdot||_1$ denotes the norm on $H^1(\Omega)$. Further, such a bound immediately gives (1.2b) since $\beta_h^1 \leq \beta_h^{\mathrm{div}}$ by definition. Hence, if $\mathrm{div}\, V_h \subseteq Q_h$, stability for Stokes gives stability for the mixed Laplacian.

The conditions $\mathrm{div}\, V_h \subseteq Q_h$ and $V_h \subset H^1(\Omega; \mathbb{V})$ are clearly satisfied by the element pairs consisting of continuous piecewise polynomial vector fields of degree less than or equal to r, and discontinuous piecewise polynomials of degree $r - 1$

for $r = 1, 2, \ldots$. This family could be viewed as an attractive family of elements for both the Stokes equations and the mixed Laplacian. However, the Brezzi inf-sup constant(s) will not be bounded for all r. For $r \geq 4$, Scott and Vogelius demonstrated that these finite element spaces will be stable for the Stokes equations on triangulations that have no nearly singular vertices, that is, triangulations that are not singular in the appropriate sense [**19**]. The lower order cases, $1 \leq r \leq 3$, were studied carefully by Qin, concluding that the elements are not stable in general [**17**]. However, they are stable for some specific families of triangulations, and can be stabilized by removal of spurious pressure modes on some other classes of triangulations. (The space N_h of spurious pressure modes is defined in (2.8) below.) In general, the stability of finite element spaces for the Stokes equations has been exhaustively investigated. In addition to the previous references, surveys are presented in [**4, 9**]. However, to our knowledge, a careful study of the lower order cases has not been conducted for the mixed Laplacian.

As the stability for the mixed Laplacian is a weaker requirement when $\operatorname{div} V_h \subseteq Q_h$, there may be a greater class of triangulations for which the elements form a stable discretization. In fact, this is known to be true. One example is provided by the pairing of continuous piecewise linear vector fields V_h and the subspace of discontinuous piecewise constants Q_h such that $Q_h = \operatorname{div} V_h$ on crisscross triangulations of the unit square. Qin proved that there does not exist a $\beta > 0$ such that $\beta_h^1 > \beta$ for any h [**17**, Lemma 7.3.1]. On the other hand, Boffi et al. proved that such a bound exists for β_h^{div} [**3**]. We shall present numerical evidence suggesting that there is a range of triangulations for which the above holds. The main results are summarized below.

Spurious modes: For $r = 2, 3$ and for all triangulations tested, the dimension of the space N_h of spurious modes is equal to the number of interior singular vertices σ. However, for $r = 1$ and one of the triangulation families studied ("Flipped", which is defined in Figure 1), $\dim N_h$ is strictly greater than σ.

Stability: For all triangulations we have tested, the method seems at least reduced stable (i.e. stable after removal of spurious modes, if any), for $r = 2, 3$. This is in contrast to the situation for the Stokes equations, where for some triangulations, such as the diagonal triangulation, the method is not stable for $r = 2, 3$, while for other triangulations, it is. For $r = 1$, reduced stability holds for some triangulations, but fails for others, including the diagonal triangulation.

Convergence: We also studied convergence of the method on diagonal triangulations. For such meshes, the method was observed to be stable for $r > 1$, but unstable for $r = 1$. Theory predicts optimal convergence of p in $L^2(\Omega)$ and u in $H(\operatorname{div}, \Omega)$ for a stable method and this is in fact what was observed. Such optimal convergence holds for $r = 2, 3, 4$, but not in the apparently unstable case $r = 1$. In the case $r \geq 4$, it is known that u converges at one order higher in $L^2(\Omega)$ than in $H(\operatorname{div}, \Omega)$. No such increase of order was observed for $r < 4$.

The note is organized as follows. We introduce further notation and summarize some key points of the theory of mixed finite element methods in Section 2. Further, we derive some eigenvalue problems associated with the stability conditions and give a characterization of the Brezzi inf-sup constant for the mixed Laplacian β_h^{div} in Section 3. These eigenvalue problems applied to the Stokes equations were also stated by Malkus [**13**], and (in part) by Qin [**17**], and provide a foundation for numerical investigations of the Brezzi stability conditions. We rely on the eigenvalue problem associated with the Brezzi inf-sup constant in Section 4. The other

eigenvalue problems are included for the sake of completeness, ease of comparison and as starting-points for the study of other pairs of element spaces. Section 4 is devoted to the study of continuous piecewise polynomials in two dimensions for the velocity and discontinuous piecewise polynomials for the pressure.

2. Notation and preliminaries

The notion of reduced stability of families of mixed finite element spaces is a key point in this note. In order to make this notion precise, this preliminary section aims to introduce notation and summarize the stability notions for finite element discretizations of abstract saddle point problems.

If V is an inner product space, we denote the dual space by V^*, the inner product on V by $\langle \cdot, \cdot \rangle_V$ and the induced norm by $|| \cdot ||_V$. Let Ω be an open and bounded domain in $\mathbb{R}^d$ with boundary $\partial\Omega$. We let $H^m(\Omega)$, for $m = 0, 1, \ldots$, denote the standard Sobolev spaces of square integrable functions with m weak derivatives and denote their norm by $|| \cdot ||_m$. Accordingly, $H^0(\Omega) = L^2(\Omega)$. The space of polynomials of degree r on Ω is denoted $\mathcal{P}_r(\Omega)$. The space of vectors in $\mathbb{R}^d$ is denoted $\mathbb{V}$, and in general, $X(\Omega; \mathbb{V})$ denotes the space of vector fields on Ω for which each component is in $X(\Omega)$. For brevity however, the space of vector fields in $L^2(\Omega; \mathbb{V})$ with square integrable divergence is written $H(\mathrm{div}, \Omega)$ with norm $|| \cdot ||_{\mathrm{div}}$ and semi-norm $| \cdot |_{\mathrm{div}} = || \operatorname{div} \cdot ||_0$. The subscripts and the reference to the domain Ω will be omitted when considered superfluous.

Let $\mathcal{T}_h$ denote an admissible simplicial tessellation of Ω, h measuring the mesh size of the tessellation. We shall frequently refer to spaces of piecewise polynomials defined relative to such, and label the spaces of continuous, and discontinuous, piecewise polynomials of degree less than or equal to r as follows.

$$\mathcal{P}_r^c = \mathcal{P}_r^c(\mathcal{T}_h) = \{p \in H^1(\Omega) \mid p|_K \in \mathcal{P}_r(K) \quad \forall\, K \in \mathcal{T}_h\} \quad r = 1, 2, \ldots,$$
$$\mathcal{P}_r = \mathcal{P}_r(\mathcal{T}_h) = \{p \in L^2(\Omega) \mid p|_K \in \mathcal{P}_r(K) \quad \forall\, K \in \mathcal{T}_h\} \quad r = 0, 1, \ldots.$$

The classical abstract saddle point problem reads as follows [**7, 9**]. For given Hilbert spaces V and Q and data $(f, g) \in V^* \times Q^*$, find $(u, p) \in V \times Q$ satisfying

$$a(u, v) + b(v, p) + b(u, q) = \langle f, v \rangle + \langle g, q \rangle \quad \forall\, (v, q) \in V \times Q, \tag{2.1}$$

where a and b are assumed to be continuous, bilinear forms on $V \times V$ and $V \times Q$, respectively. We shall assume here and throughout that a is symmetric. Following [**2**], there exists a unique and stable solution (u, p) of (2.1), if there exists a positive constant γ such that

$$0 < \gamma \leq \inf_{0 \neq (u,p)} \sup_{0 \neq (v,q)} \frac{a(u, v) + b(v, p) + b(u, q)}{||(u, p)||_{V \times Q} ||(v, q)||_{V \times Q}}. \tag{2.2}$$

Equivalently [**7**], existence, uniqueness and stability of the solution are guaranteed by the existence of positive constants α, β such that

$$0 < \alpha \leq \inf_{0 \neq u \in Z} \sup_{0 \neq v \in Z} \frac{a(u, v)}{||u||_V ||v||_V}, \tag{2.3a}$$

$$0 < \beta \leq \inf_{0 \neq q \in Q} \sup_{0 \neq v \in V} \frac{b(v, q)}{||v||_V ||q||_Q}, \tag{2.3b}$$

where $Z = \{v \in V \mid b(v, q) = 0 \quad \forall\, q \in Q\}$. We shall refer to γ as the continuous Babuška inf-sup constant, and α and β as the continuous Brezzi coercivity and Brezzi inf-sup constants, respectively.

Given finite dimensional spaces $V_h \subset V$ and $Q_h \subset Q$, defined relative to a tessellation $\mathcal{T}_h$ of Ω, the Galerkin discretization of (2.1) takes the following form.

Find $(u_h, p_h) \in V_h \times Q_h$ satisfying

$$a(u_h, v) + b(v, p_h) + b(u_h, q) = \langle f, v\rangle + \langle g, q\rangle \quad \forall\, (v, q) \in V_h \times Q_h. \tag{2.4}$$

The discrete analogy of (2.2) now reads:

$$0 \le \gamma_h = \inf_{0\neq(u,p)\in V_h\times Q_h} \sup_{0\neq(v,q)\in V_h\times Q_h} \frac{a(u,v) + b(v,p) + b(u,q)}{||(u,p)||_{V\times Q}||(v,q)||_{V\times Q}}, \tag{2.5}$$

and respectively for (2.3):

$$0 \le \alpha_h = \inf_{0\neq u\in Z_h} \sup_{0\neq v\in Z_h} \frac{a(u,v)}{||u||_V||v||_V}, \tag{2.6a}$$

$$0 \le \beta_h = \inf_{0\neq q\in Q_h} \sup_{0\neq v\in V_h} \frac{b(v,q)}{||v||_V||q||_Q}, \tag{2.6b}$$

where

$$Z_h = \{v \in V_h \mid b(v, q) = 0 \quad \forall\, q \in Q_h\}. \tag{2.7}$$

The values γ_h, α_h and β_h will be referred to as the Babuška inf-sup, Brezzi coercivity and Brezzi inf-sup constants, respectively. For given $V_h \times Q_h$, there exists a unique solution of (2.4) if α_h and β_h (or equivalently γ_h) are positive. Further, if α_h and β_h are uniformly bounded from below for a family of discretization spaces $\{V_h \times Q_h\}_h$, parameterized over h, one obtains the quasi-optimal approximation estimate [**7**]:

$$||u - u_h||_V + ||p - p_h||_Q \le C \left(\inf_{v\in V} ||u - v||_{V_h} + \inf_{q\in Q} ||p - q||_{Q_h} \right).$$

The uniform boundedness condition motivates the notion of stability for pairs of finite element spaces.

Definition 2.1 (Stable discretization). *A family of finite element discretizations $\{V_h \times Q_h\}_h$ is* stable in $V \times Q$ *if the Brezzi coercivity and inf-sup constants α_h and β_h (or equivalently the Babuška inf-sup constant γ_h) are bounded from below by a positive constant independent of h.*

We shall also, in agreement with the standard terminology, say that $\{V_h \times Q_h\}_h$ satisfies the Brezzi coercivity, respectively inf-sup, condition if α_h, respectively β_h, is uniformly bounded from below.

There are families of discretizations that are not stable in the sense defined above, but have a reduced stability property. More precisely, for a pair $V_h \times Q_h$ consider the space of spurious modes $N_h \subseteq Q_h$:

$$N_h = \{q \in Q_h \mid b(v, q) = 0 \quad \forall\, v \in V_h\}. \tag{2.8}$$

For a stable discretization, N_h contains only the zero element. Conversely, $\beta_h = 0$ if and only if N_h contains non-zero elements. Then, if N_h is non-trivial, it may be natural to consider the reduced space $N_h^\perp$, the orthogonal complement of N_h in Q_h, in place of Q_h. This motivates the definition of the reduced Brezzi inf-sup constant, relating to the stability of $V_h \times N_h^\perp$:

$$\tilde{\beta}_h = \inf_{0\neq q\in N_h^\perp} \sup_{0\neq v\in V_h} \frac{b(v,q)}{||v||_V||q||_Q}, \tag{2.9}$$

and the following definition of reduced stable. By definition, $\tilde{\beta}_h \neq 0$.

Definition 2.2 (Reduced stable discretization). *A family of discretizations $\{V_h \times Q_h\}_h$ is* reduced stable in $V \times Q$ *if the Brezzi coercivity constant α_h and the reduced Brezzi inf-sup constant $\tilde{\beta}_h$, defined by* (2.6a) *and* (2.9) *respectively, are bounded from below by a positive constant independent of h.*

3. Eigenvalue problems related to the Babuška-Brezzi constants

For a given set of discrete spaces, the Babuška and Brezzi constants defined by (2.5) and (2.6) can be computed by means of eigenvalue problems. The form and properties of the eigenvalue problem associated with the Brezzi inf-sup constant for the Stokes equations was discussed by Qin [**17**]. Since also the Brezzi coercivity constant plays a role for the mixed Laplacian, we begin this section by deriving how the Brezzi coercivity constant can be computed by similar eigenvalue problems. These eigenvalue problems were also stated, and carefully analysed from an algebraic view-point, by Malkus [**13**] in connection with the displacement-pressure formulation of the linear elasticity equations. We continue by observing that the continuous Brezzi inf-sup constant can be naturally associated with the smallest eigenvalue of the Laplacian itself.

3.1. Eigenvalue problems for the discrete Babuška-Brezzi constants. Let $V_h \subset V$ and $Q_h \subset Q$ be given finite dimensional spaces as before. By definition, the Babuška inf-sup constant $\gamma_h = |\lambda_{\min}|$ when $\lambda_{\min}$ is the smallest (in modulus) eigenvalue of the following generalized eigenvalue problem: Find $\lambda \in \mathbb{R}$, $0 \neq (u, p) \in V_h \times Q_h$ satisfying

$$a(u, v) + b(v, p) + b(u, q) = \lambda \left(\langle u, v \rangle_V + \langle p, q \rangle_Q \right) \quad \forall\, (v, q) \in V_h \times Q_h. \tag{3.1}$$

Further, the following lemma identifies an eigenvalue problem associated with the Brezzi inf-sup constant.

Lemma 3.1 (Qin [**17**, Lemma 5.1.1 – 5.1.2]). *Let $\lambda_{\min}$ be the smallest eigenvalue of the following generalized eigenvalue problem: Find $\lambda \in \mathbb{R}$, $0 \neq (u, p) \in V_h \times Q_h$ satisfying*

$$\langle u, v \rangle_V + b(v, p) + b(u, q) = -\lambda \langle p, q \rangle_Q, \quad \forall\, (v, q) \in V_h \times Q_h. \tag{3.2}$$

Then, $\lambda \geq 0$ and for β_h defined by (2.6b), $\beta_h = \sqrt{\lambda_{\min}}$.

It can also be shown that the reduced Brezzi inf-sup constant $\tilde{\beta}_h$ equals the square-root of the smallest non-zero eigenvalue of (3.2) [**17**, Theorem 5.1.1].

The Babuška and Brezzi inf-sup constants are thus easily computed, given bases for the spaces V_h and Q_h. As for (3.1), it is easily seen that the Brezzi coercivity constant $\alpha_h = |\lambda_{\min}|$ where $\lambda_{\min}$ is the smallest (in modulus) eigenvalue of the eigenvalue problem: Find $\lambda \in \mathbb{R}$ and $0 \neq u \in Z_h$ such that

$$a(u, v) = \lambda \langle u, v \rangle_V \quad \forall\, v \in Z_h. \tag{3.3}$$

However, a basis for Z_h is usually not readily available, thus hindering the actual computation of the eigenvalues of (3.3). Instead, the above eigenvalue problem over Z_h can be extended to a generalized eigenvalue problem over $V_h \times Q_h$: Find $\lambda \in \mathbb{R}$ and $0 \neq (u, p) \in V_h \times Q_h$ such that

$$a(u, v) + b(v, p) + b(u, q) = \lambda \langle u, v \rangle_V \quad \forall\, (v, q) \in V_h \times Q_h. \tag{3.4}$$

The following lemma establishes the equivalence between (3.3) and (3.4).

Lemma 3.2. *If (λ, u) is an eigenpair of* (3.3)*, there exists a $p \in Q_h$ such that $(\lambda, (u, p))$ is an eigenpair of* (3.4)*. Conversely, if $(\lambda, (u, p))$ is an eigenpair of* (3.4) *and $u \neq 0$, then $u \in Z_h$ and (λ, u) is an eigenpair of* (3.3)*. For $0 \neq p \in N_h$ and any scalar λ, $(\lambda, (0, p))$ is an eigenpair of* (3.4)*, and these are the only eigenpairs of* (3.4) *with $u = 0$.*

Proof. Let (λ, u) be an eigenpair of (3.3). Define $B_h : V_h \to Q_h$ such that $\langle B_h v, q \rangle_Q = b(v, q)$ for all $q \in Q_h$. Since $B_h : Z_h^\perp \to B_h(V_h)$ is an isomorphism, $p \in B_h(V_h) \subset Q_h$ is well-defined by

$$\langle p, q \rangle_Q = \lambda \langle u, B_h^{-1} q \rangle_V - a(u, B_h^{-1} q) \quad \forall\, q \in B_h(V_h).$$

Then, for any $v \in Z_h^\perp$, p satisfies

$$b(v,p) = \langle B_h v, p\rangle_Q = \lambda\langle u, v\rangle_V - a(u,v).$$

Further, by definition $b(v,p) = 0$ for any $v \in Z_h$. Hence, by the assumption that (λ, u) is an eigenpair of (3.3), $(\lambda, (u,p))$ satisfies (3.4). The converse statement is obvious. Finally, letting $u = 0$ in (3.4), we see that $(\lambda, (0,p)$ satisfies (3.4) if and only if $p \in N_h$, but for any $\lambda \in \mathbb{R}$. □

In the subsequent section, we shall numerically investigate the stability of families of finite element discretizations $\{V_h \times Q_h\}_h$ such that $\operatorname{div} V_h \subseteq Q_h$ for the mixed Laplacian, using the eigenvalue problem (3.2) in terms of standard bases for the spaces V_h and Q_h. For such an investigation of discretizations where $\operatorname{div} V_h \not\subseteq Q_h$, for instance for equal order approximations, the eigenvalue problem (3.4) would be equally important. Note that, as a consequence of the last observation in Lemma 3.2, if N_h is non-trivial, the generalized eigenvalue problem (3.4) is computationally not well-posed since any scalar λ is an eigenvalue.

3.2. A characterization of the Brezzi inf-sup constant for the mixed Laplacian. We now turn from the general setting to consider the $H(\operatorname{div}) \times L^2$ formulation of the mixed Laplacian (1.1). In Lemma 3.3 below, we show that the Brezzi inf-sup constant can be identified with the smallest eigenvalue of the negative Laplacian. Consequently, if a discretization family $\{V_h \times Q_h\}_h$ guarantees eigenvalue convergence for the mixed Laplace eigenvalue problem, and is such that $\operatorname{div} V_h \subseteq Q_h$, the Brezzi inf-sup constant of the discretization will converge to the continuous Brezzi inf-sup constant.

LEMMA 3.3. *Let $V \subseteq H(\operatorname{div}, \Omega)$ and $Q \subseteq L^2(\Omega)$ be such that that $\operatorname{div} V \subseteq Q$. Consider the Brezzi inf-sup eigenvalue problem* (3.2) *applied to* (1.1)*:*

$$\langle u, v\rangle_{\operatorname{div}} + \langle \operatorname{div} v, p\rangle + \langle \operatorname{div} u, q\rangle = -\lambda\langle p, q\rangle \quad \forall\, (v,q) \in V \times Q. \tag{3.5}$$

Consider also the mixed Laplace eigenvalue problem:

$$\langle \hat{u}, v\rangle + \langle \operatorname{div} v, \hat{p}\rangle + \langle \operatorname{div} \hat{u}, q\rangle = -\hat{\lambda}\langle \hat{p}, q\rangle \quad \forall\, (v,q) \in V \times Q. \tag{3.6}$$

Then, $(\lambda, (u,p))$ is an eigenpair of (3.5) *if and only if $(\hat{\lambda}, (\hat{u}, \hat{p}))$ is an eigenpair of* (3.6) *where $\hat{\lambda} = \lambda(1-\lambda)^{-1}$, $\hat{u} = u$ and $\hat{p} = (1-\lambda)p$.*

PROOF. Assume that $(\lambda, (u,p))$ is an eigenpair of (3.5). First, note that $\lambda \neq 1$. Letting $\lambda = 1$, $v = u$ and $q = -\operatorname{div} u$ in (3.5), implies that $u = 0$. Further, $v = 0$ and $q = p$ gives that $p = 0$. Hence, $\lambda = 1$ is only associated with the zero solution, which by definition, cannot form an eigenpair. Also, letting $v = u$, $q = -p$, gives that $\lambda \geq 0$. Next, by the assumption $\operatorname{div} V \subseteq Q$,

$$\langle \operatorname{div} u, \operatorname{div} v\rangle = -\lambda\langle p, \operatorname{div} v\rangle \quad \forall\, v \in V.$$

Hence,

$$\langle u, v\rangle + \langle \operatorname{div} u, q\rangle + (1-\lambda)\langle \operatorname{div} v, p\rangle = -\lambda\langle p, q\rangle.$$

Letting $\hat{u} = u$, $\hat{p} = (1-\lambda)p$ and $\hat{\lambda} = \lambda(1-\lambda)^{-1}$, give that $(\hat{\lambda}, (\hat{u}, \hat{p}))$ solves (3.6). The converse holds by similar arguments. □

The equivalence demonstrated in the lemma above affords a simple characterization of the Brezzi inf-sup constant for the mixed Laplacian. The eigenvalue problem (3.5) with $V = H(\operatorname{div}, \Omega)$ and $Q = L^2(\Omega)$ is the eigenvalue problem associated with the continuous Brezzi inf-sup constant $\beta^{\operatorname{div}}$, cf. Lemma 3.1. Hence, $\beta^{\operatorname{div}}$ is the square-root of the smallest eigenvalue of (3.5). On the other hand, the eigenvalue problem (3.6) is a mixed weak formulation of the standard eigenvalue

problem for the negative Laplacian with Dirichlet boundary conditions, given in strong form below:

$$-\Delta\hat{p} = \hat{\lambda}\,\hat{p} \text{ in } \Omega, \quad \hat{p} = 0 \text{ on } \partial\Omega. \tag{3.7}$$

Thus, if μ is the smallest eigenvalue of (3.7), $\beta^{\mathrm{div}} = \sqrt{\mu(1+\mu)^{-1}}$.

REMARK 3.4. *An alternative eigenvalue problem arises from noting that, under the assumption* $\operatorname{div} V \subseteq Q$, (3.5) *implies that* $\lambda p = -\operatorname{div} u$. *Hence if* $(\lambda, (u,p))$ *solves* (3.5), *then either* $\lambda = 0, u = 0$ *and* $p \perp \operatorname{div} V$, *or* $\lambda \neq 1$ *and for* $\hat{\lambda} = \lambda(1-\lambda)^{-1}$, $(\hat{\lambda}, u)$ *will be an eigenpair of the problem:*

$$\langle \operatorname{div} u, \operatorname{div} v\rangle = \hat{\lambda}\langle u, v\rangle \quad \forall\, v \in V.$$

This eigenproblem was studied in [**5**].

Now, consider a stable discretization family $\{V_h \times Q_h\}_{h\to 0}$ of (1.1) such that $\operatorname{div} V_h \subseteq Q_h$, with Brezzi inf-sup constants $\{\beta_h^{\mathrm{div}}\}_h$. Let μ_h denote the smallest eigenvalue approximation of (3.6) by $V_h \times Q_h$. As a consequence of the previous, if $\mu_h \to \mu$, then $\beta_h^{\mathrm{div}} \to \beta^{\mathrm{div}}$. In other words, if the discretization family is stable, satisfies $\operatorname{div} V_h \subseteq Q_h$, and gives eigenvalue convergence, then the Brezzi inf-sup constant will converge to the continuous Brezzi inf-sup constant. Note however, that the discrete stability conditions are *not* sufficient for the convergence of approximations to the eigenvalue problem (3.6) [**1, 3**].

Mixed finite element discretizations of (1.1) based on the Raviart-Thomas [**18**] and Brezzi-Douglas-Marini [**8**] families of $H(\operatorname{div})$ conforming elements are such that $\hat{\lambda}_h \to \hat{\lambda}$ and hence $\beta_h^{\mathrm{div}} \to \beta^{\mathrm{div}}$. The case where β_h^{div} seems to be uniformly bounded in h, but $\beta_h^{\mathrm{div}} \not\to \beta^{\mathrm{div}}$ is exemplified in the subsequent section. Finally, note that if Ω is the unit square: $\Omega = (0,1)^2$, the smallest eigenvalue of (3.7) is $2\pi^2$ and so

$$\beta^{\mathrm{div}} = \sqrt{\frac{2\pi^2}{1+2\pi^2}} \approx 0.975593. \tag{3.8}$$

4. Lower order Lagrange elements for the mixed Laplacian

From here on, we restrict our attention to finite element discretizations of the mixed Laplacian (1.1) on a polyhedral domain $\Omega \subset \mathbb{R}^2$. The primary aim is to examine the stability, or reduced stability, and convergence properties of such, utilizing Lagrange elements, that is, continuous piecewise polynomials, for the vector variable and discontinuous piecewise polynomials for the scalar variable:

$$V_h \times Q_h = \mathcal{P}_r^c(\mathcal{T}_h; \mathbb{V}) \times \mathcal{P}_{r-1}(\mathcal{T}_h), \tag{4.1}$$

for $r = 1, 2, \ldots$. Although the Brezzi conditions are in general not satisfied for these discretizations, stability or reduced stability may be identified on families of structured triangulations. The pair (4.1) is clearly such that $\operatorname{div} V_h \subseteq Q_h$. Therefore, the stability of the discretization relies on a uniform bound for the Brezzi inf-sup constant only. Further, a uniform lower bound on the Brezzi inf-sup constant for the Stokes equations induces the corresponding bound for the mixed Laplacian. Hence, the results on the reduced stability of this element pair for the Stokes equations can be directly applied to the mixed Laplacian. In the following, new numerical evidence is presented and compared to the known results.

The stability of the $\mathcal{P}_r^c(\mathbb{V}) \times \mathcal{P}_{r-1}$ family of elements, for both the Stokes equations and the mixed Laplacian, depends on the polynomial degree r and the structure of the triangulation family $\{\mathcal{T}_h\}_h$. For triangulations that have interior singular vertices, the space of spurious modes N_h, defined by (2.8) applied to (1.1), will be non-trivial. Here, an interior vertex is labelled singular if the edges meeting at that vertex fall on two straight lines. Let x denote an interior singular vertex

	Diagonal	Flipped	Zigzag	Crisscross	Union Jack
σ	0	0	0	n^2	$\frac{1}{2}n(n-2)$
$\dim N_h$	0	$\begin{cases}(\frac{1}{2}n-1)^2 & r=1\\ 0 & r=2,3\end{cases}$	0	n^2	$\frac{1}{2}n(n-2)$

TABLE 1. The number of interior singular vertices σ and the dimension of the space of spurious modes, $\dim N_h$, for labelled families of $n \times n$ triangulations of the unit square, cf. Figure 1. For the flipped, zigzag and Union Jack meshes, $\dim N_h$ is conjectural.

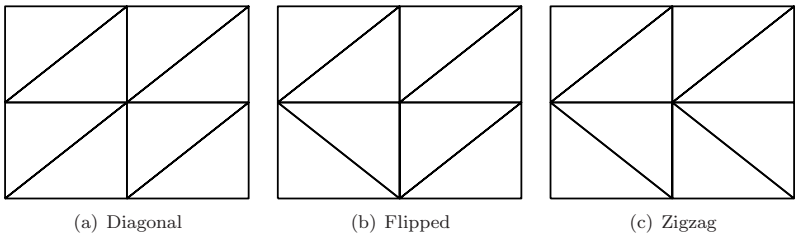
and let ω_x be the star of x. For any $r \geq 1$, there exists a $p \in N_h$ such that p is supported in ω_x [**15, 16**]. Consequently, letting σ denote the number of interior singular vertices of a triangulation, $\dim N_h \geq \sigma$. Scott and Vogelius showed that the converse is true for $r \geq 4$: If there are no interior singular vertices, then $\dim N_h = 0$ and so $\beta_h^{\mathrm{div}} > \beta_h^1 > 0$. Moreover, they proved that for a family of meshes without interior singular vertices, β_h^1 remains bounded above zero as long as the meshes do not tend to singularity as $h \to 0$. For the precise statement and more details, see [**19**] or [**6**, Section 10.6].

As we shall see below, for $r < 4$, the space of spurious modes may be non-trivial even when there are no singular vertices. Further, for the Stokes equations, more restrictive conditions than the above must be placed on the triangulations in order to obtain a uniform bound for the Stokes Brezzi inf-sup constant [**17**]. The stability properties of these lower order discretizations for the mixed Laplacian is the main question of interest in the following.

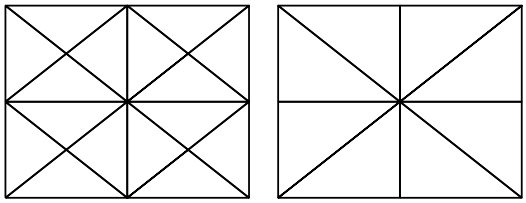
REMARK 4.1. *We shall not consider the pairing of continuous versus discontinuous polynomials of other polynomial degrees than those of* (4.1)*. This choice is easily motivated. First, a dimension count shows that the pairing of continuous piecewise polynomial vector fields with discontinuous piecewise polynomials of the same or higher degree must have a non-trivial space of spurious modes. Second, although the Brezzi inf-sup constant is uniformly bounded for the pairs* $\mathcal{P}_r^c(\mathbb{V}) \times \mathcal{P}_{r-2}$, $r = 2, 3, \ldots$*, the Brezzi coercivity constant for the mixed Laplacian is not uniformly bounded, and thus stability fails.*

4.1. Stability. In the spirit of [**17**, Section 5], we aim to numerically investigate the stability of $\mathcal{P}_r^c(\mathbb{V}) \times \mathcal{P}_{r-1}$ for $r = 1, 2, 3$ on certain families of structured triangulations of the unit square. The triangulation patterns considered are illustrated and labelled in Figure 1. For n even, an $n \times n$ triangulation of each family is constructed by first partitioning the domain into $n \times n$ squares, and subsequently dividing each block of 2×2 squares into triangles by the respective patterns. For instance, an $n \times n$ diagonal triangulation is formed by dividing the unit square into $n \times n$ subsquares, and dividing each subsquare into triangles by the positive diagonal. Throughout, we identify $h = 1/n$ and assume that $n > 2$. Observe that the diagonal, flipped, and zigzag triangulations contain no interior singular vertices, while the crisscross and the Union Jack triangulation contain n^2 and $\frac{1}{2}n(n-2)$ interior singular vertices respectively. This is summarized in the first row of Table 1.

Recall that $\dim N_h \geq \sigma$ for $r \geq 1$ and equality holds for $r \geq 4$. Qin proved that equality holds for $1 \leq r \leq 3$ in the case of the diagonal and the crisscross meshes and numerically observed equality for the flipped mesh for $r = 2$ [**17**]. Our own experiments show that equality holds for the zigzag and Union Jack meshes



(a) Diagonal (b) Flipped (c) Zigzag



(d) Crisscross (e) Union Jack

FIGURE 1. Structured 2×2 triangulations of the unit square.

for $1 \leq r \leq 3$. Equality also holds for the flipped mesh when $r = 2, 3$, but not for $r = 1$. These results are summarized in the second row of Table 1.

We continue by studying the behaviour of the Brezzi inf-sup constants on the above triangulations. The cases $r = 2, 3$ are considered first, but we will return to the case $r = 1$ below. For the Stokes equations, it is known that the diagonal and the crisscross triangulation families exhibit very different behaviour for $r = 2, 3$ [**17**]. Namely, although there are non-trivial spurious modes on the crisscross triangulation family, the reduced Brezzi inf-sup constant is uniformly bounded. In contrast, for the diagonal family, the Brezzi inf-sup constant decays as approximately $\mathcal{O}(h)$. As the discretization is reduced stable for the Stokes equations on crisscross triangulations, it is also reduced stable for the mixed Laplacian. A natural question becomes whether the lack of stability on diagonal triangulations for the Stokes equations is also present for the mixed Laplacian.

In view of Lemma 3.1, we shall make an attempt at answering this question through a set of numerical experiments. For a given r and a given $\mathcal{T}_h$, the smallest, and smallest non-zero, eigenvalue of (3.5) for $V = \mathcal{P}_r^c(\mathcal{T}_h, \mathbb{V})$, $Q = \mathcal{P}_{r-1}(\mathcal{T}_h)$ give the Brezzi inf-sup and reduced Brezzi inf-sup constant. These eigenvalues for the triangulation families considered, computed using LAPACK, SLEPc [**11**] and DOLFIN [**12**], are given for $r = 1, 2, 3$ in Tables 2, 3 and 4 respectively. For the purpose of identifying spurious modes, eigenvalues below a threshold of 10^{-4} have been tabulated to zero[6].

For the diagonal meshes, the numerical experiments indicate that in contrast to Stokes, the mixed Laplacian Brezzi inf-sup constants are bounded from below for both $r = 2, 3$. For the flipped and zigzag meshes, experiments give similar results; Neither exhibits any spurious modes. Moreover, while the Brezzi inf-sup constant decays approximately as $\mathcal{O}(h)$ for the Stokes equations [**17**], it appears to be uniformly bounded for the mixed Laplacian. For the Union Jack family, the same experiment gives $\frac{1}{2}n(n-2)$ spurious modes, but the reduced Brezzi inf-sup constant

[6]Had a smaller threshold been chosen, some of the zero eigenvalues associated to interior singular vertices would have been missed for $r = 2$ on the Union Jack mesh of size $n = 6$.

	β_h^{div}		$\tilde{\beta}_h^{\mathrm{div}}$ $(\dim N_h)$	
n	Diagonal	Zigzag	Flipped	Union Jack
4	0.847171	0.791967	0.945496 (1)	0.976985 (4)
6	0.716677	0.626865	0.945619 (4)	0.976271 (12)
8	0.605576	0.505968	0.947850 (9)	0.975985 (24)
10	0.517707	0.420180	0.946138 (16)	0.975847 (40)
12	0.449060	0.357720	0.944833 (25)	0.975770 (60)
14	0.394963	0.310731	0.943880 (36)	0.975724 (84)
16	0.351684	0.274303	0.943142 (49)	0.975693 (112)

TABLE 2. The mixed Laplacian (reduced) Brezzi inf-sup constant for $\mathcal{P}_1^c(\mathcal{T}_h;\mathbb{V}) \times \mathcal{P}_0(\mathcal{T}_h)$ on labelled structured families of triangulations $\mathcal{T}_h$. The dimension of the space of spurious modes in parenthesis if non-trivial.

	β_h^{div}			$\tilde{\beta}_h^{\mathrm{div}}$ $(\dim N_h)$
n	Diagonal	Zigzag	Flipped	Union Jack
4	0.975627	0.955956	0.943790	0.975628 (4)
6	0.975600	0.952460	0.940480	0.975603 (12)
8	0.975595	0.951384	0.938717	0.975595 (24)
10	0.975594	0.950906	0.937684	0.975594 (40)
12	0.975594	0.950638	0.936992	0.975593 (60)
14	0.975593	0.950458		

TABLE 3. The mixed Laplacian (reduced) Brezzi inf-sup constant for $\mathcal{P}_2^c(\mathcal{T}_h;\mathbb{V}) \times \mathcal{P}_1(\mathcal{T}_h)$ on labelled structured families of triangulations $\mathcal{T}_h$. The dimension of the space of spurious modes in parenthesis if non-trivial.

	β_h^{div}			$\tilde{\beta}_h^{\mathrm{div}}$ $(\dim N_h)$
n	Diagonal	Zigzag	Flipped	Union Jack
4	0.972244	0.975594	0.975594	0.975594 (4)
6	0.967304	0.975593	0.975593	0.975593 (12)
8	0.964845	0.975593	0.975593	0.975593 (24)
10	0.963412			
12	0.962484			

TABLE 4. The mixed Laplacian (reduced) Brezzi inf-sup constant for $\mathcal{P}_3^c(\mathcal{T}_h;\mathbb{V}) \times \mathcal{P}_2(\mathcal{T}_h)$ on labelled structured families of triangulations $\mathcal{T}_h$. The dimension of the space of spurious modes in parenthesis if non-trivial.

again seems to be uniformly bounded. In summary for $r = 2, 3$, the $\mathcal{P}_r^c(\mathbb{V}) \times \mathcal{P}_{r-1}$ elements appear to be at least reduced stable for all the families considered.

With the discussion in Section 3.2 in mind, we also note that the Brezzi inf-sup constant converges to the exact value, given by (3.8), for some, but not all, of these meshes. For $r = 2$, the Brezzi inf-sup constant seems to converge to the exact value on the diagonal meshes, but not on the flipped or the zigzag meshes. The situation is the opposite for $r = 3$. There, the Brezzi inf-sup constant seems to converge to the exact value on the zigzag and flipped meshes, but not for the diagonal meshes.

The situation is different and more diverse in the lowest-order case: $r = 1$. Boffi et al. proved that $\mathcal{P}_1^c(\mathbb{V}) \times \mathcal{P}_0$ is in fact reduced stable for the mixed Laplacian on crisscross meshes [**3**]. It is not reduced stable for Stokes [**17**]. However, the element pair does not seem to be stable on diagonal meshes. The values in the first column of Table 2 indicate that the Brezzi inf-sup constant decays approximately as $\mathcal{O}(h)$. The same is the case for the zigzag meshes. For the Union Jack meshes, the situation is similar to the crisscross case. That is, the number of singular modes match the number of interior singular vertices and the reduced Brezzi inf-sup constant appears to be bounded from below. Finally, the flipped meshes display a surprising behaviour. There seem to be $(\frac{n}{2} - 1)^2$ spurious modes, even though there are no singular vertices. This is the only case where we have observed $\dim N_h > \sigma$. However, the reduced Brezzi inf-sup constant appears to be uniformly bounded.

4.2. Convergence. In the previous, we have considered the stability of the $\mathcal{P}_r^c(\mathbb{V}) \times \mathcal{P}_{r-1}$ elements. Now, we proceed to examine the convergence properties of these elements on the diagonal meshes. Conjecturing that $\mathcal{P}_r^c(\mathbb{V}) \times \mathcal{P}_{r-1}$ is stable on this mesh family for $r \geq 2$, in accordance with the numerical evidence presented above, the standard theory gives the error estimate

$$||u - u_h||_{\text{div}} + ||p - p_h||_0 \leq Ch^r \left(||u||_{r+1} + ||p||_r \right). \tag{4.2}$$

For $r \geq 4$, the L^2 error estimate for the velocity can be improved [**6**, Theorem 10.4.9], thus yielding:

$$||u - u_h||_0 \leq Ch^{r+1} ||u||_{r+1}. \tag{4.3}$$

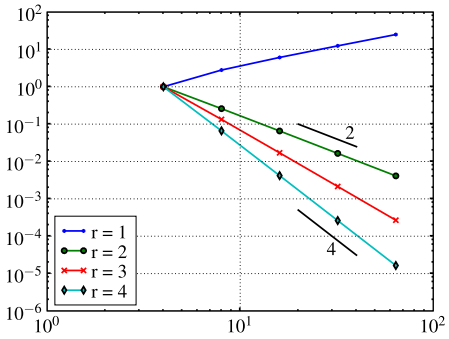
In order to verify (4.2) and to inspect whether (4.3) could be attained for $r = 2, 3$, we consider a standard smooth exact solution to the Laplacian with pure Dirichlet boundary conditions:

$$p(x, y) = \sin(2\pi x) \sin(2\pi y), \quad u = \nabla p, \quad g = \operatorname{div} u. \tag{4.4}$$

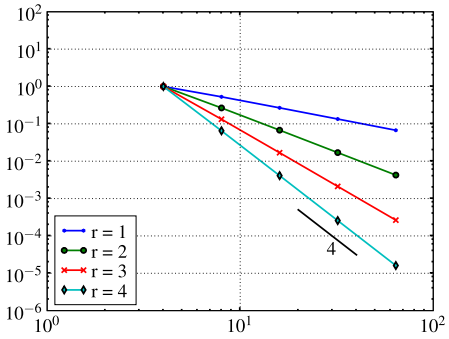
The errors of the $\mathcal{P}_r^c(\mathbb{V}) \times \mathcal{P}_{r-1}$ approximations for $1 \leq r \leq 4$ on diagonal meshes can be examined in Figure 2. To compute the errors, both the source function g and the exact solutions u, p have been represented by sixth order piecewise polynomial interpolants $\Pi_h g$ and $\Pi_h u, \Pi_h p$, whereupon the errors have been calculated exactly (up to numerical precision).

For $r = 1$, we observed the discretization to be unstable on diagonal meshes. As expected in this case, neither the pressure nor the velocity approximation seems to converge in the L^2 norm. This indicates that the estimates for the approximation error, based on the standard estimates and the decaying Brezzi inf-sup constant, cannot be improved. On the other hand, for $r > 1$, we observed the method to be stable. For $r = 2, 3, 4$, the orders of convergence in the $H(\text{div})$ norm of the velocity and the L^2 norm of the pressure approximations are indeed optimal, as predicted by (4.2).

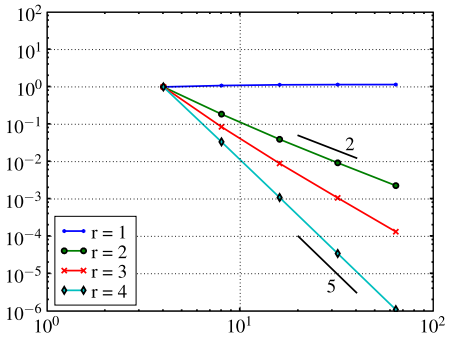
The situation seems different for the convergence of the velocity approximation in the L^2 norm. For $r \geq 4$, a convergence rate of order $r + 1$ is predicted by (4.3). This is also observed for $r = 4$ in Figure 2(c). On the other hand, for $r = 2$, the rate of convergence seems to be of order r, and thus one order suboptimal. The



(a) Normalized pressure errors in $||\cdot||_0$.



(b) Normalized velocity errors in $|\cdot|_{\mathrm{div}}$



(c) Normalized velocity errors in $||\cdot||_0$

FIGURE 2. The errors of $\mathcal{P}_r^c(\mathbb{V}) \times \mathcal{P}_{r-1}$ approximations for $r = 1, 2, 3, 4$ on diagonal meshes versus mesh number n. The errors have been normalized, that is, multiplied by the inverse of the error at the smallest mesh $n = 4$.

same is observed for $r = 3$. We therefore conjecture that the estimate (4.3) does not hold for $r = 2, 3$.

Acknowledgments. The first author acknowledges the support of the Centre of Mathematics for Applications, University of Oslo.

Center for Biomedical Computing, Simula Research Laboratory, P.O.Box 134, 1325 Lysaker, Norway. meg@simula.no

School of Mathematics, 512 Vincent Hall, 206 Church Street S.E., University of Minnesota, Minneapolis, MN 55455. arnold@umn.edu

Bibliography

[1] D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus, homological techniques and applications. *Acta numerica*, 15:1–155, 2006.

[2] I. Babuška. The finite element method with Lagrangian multipliers. *Numer. Math.*, 20:179–192, 1973.

[3] D. Boffi, F. Brezzi, and L. Gastaldi. On the problem of spurious eigenvalues in the approximation of linear elliptic problems in mixed form. *Math. Computation*, 69:121–140, 1999.

[4] D. Boffi, F. Brezzi, and L. Gastaldi. Finite elements for the Stokes problem. In *Mixed Finite Elements, Compatibility Conditions, and Applications*. Springer, 2007.

[5] D. Boffi, R. G. Duran, and L. Gastaldi. A remark on spurious eigenvalues in a square. *Appl. Math. Lett.*, 12:107–114, 1999.

[6] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 1996.

[7] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *R.A.I.R.O. Anal. Numer.*, 2:129–151, 1974.

[8] F. Brezzi, J. Douglas, and L. D. Marini. Two families of mixed elements for second order elliptic problems. *Numer. Math.*, 47:217–235, 1985.

[9] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Springer, 1991.

[10] D. Chapelle and K. J. Bathe. The inf-sup test. *Computers and Structures*, 47:537–545, 1993.

[11] V. Hernandez, J. E. Roman, and V. Vidal. SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems. *ACM Transactions on Mathematical Software*, 31:351–362, 2005.

[12] J. Hoffman, J. Jansson, A. Logg, G. N. Wells, et al. DOLFIN.

[13] D. S. Malkus. Eigenproblems associated with the discrete LBB condition for incompressible finite elements. *Int. J. Engng. Sci.*, 19:1299–1310, 1981.

[14] K.-A. Mardal, X.-C. Tai, and R. Winther. A robust finite element method for Darcy–Stokes flow. *SIAM Jour. Num. Anal.*, 40:1605–1631, 2002.

[15] J. Morgan and L. R. Scott. A nodal basis for C^1 piecewise polynomials of degree $n \geq 5$. *Math. Computation*, 29:737–740, 1975.

[16] J. Morgan and L. R. Scott. The dimension of the space of C^1 piecewise polynomials. 1996.

[17] J. Qin. *On the Convergence of Some Simple Finite Elements for Incompressible Flows*. PhD thesis, Penn State, 1994.

[18] P. A. Raviart and J. M. Thomas. A mixed finite element method for second order elliptic problems. *Mathematical Aspects of the Finite Element Meth. Lecture Notes in Math., Springer*, 606:292–315, 1977.

[19] L. R. Scott and Vogelius. Norm estimates for a maximal right inverse of the divergence operator. *R.A.I.R.O. Anal. Numer.*, 19:111–143, 1985.