IMA Journal of Numerical Analysis (2020) **40**, 1838–1874
doi:10.1093/imanum/drz022
Advance Access publication on 17 May 2019

A mass conserving mixed stress formulation for the Stokes equations

JAY GOPALAKRISHNAN
Fariborz Maseeh Department of Mathematics and Statistics, Portland State University, PO Box 751, Portland, OR 97207,USA
gjay@pdx.edu

AND

PHILIP L. LEDERER* AND JOACHIM SCHÖBERL
Institute for Analysis and Scientific Computing, Technische Universität Wien, Wiedner Hauptstraße 8-10, 1040 Wien, Austria
*Corresponding author: philip.lederer@tuwien.ac.at joachim.schoeberl@tuwien.ac.at

[Received on 19 June 2019; revised on 22 March 2019]

We propose stress formulation of the Stokes equations. The velocity u is approximated with $H(\mathrm{div})$-conforming finite elements providing exact mass conservation. While many standard methods use H^1-conforming spaces for the discrete velocity $H(\mathrm{div})$-conformity fits the considered variational formulation in this work. A new stress-like variable σ equalling the gradient of the velocity is set within a new function space $H(\mathrm{curl}\,\mathrm{div})$. New matrix-valued finite elements having continuous 'normal-tangential' components are constructed to approximate functions in $H(\mathrm{curl}\,\mathrm{div})$. An error analysis concludes with optimal rates of convergence for errors in u (measured in a discrete H^1-norm), errors in σ (measured in L^2) and the pressure p (also measured in L^2). The exact mass conservation property is directly related to another structure-preservation property called *pressure robustness*, as shown by pressure-independent velocity error estimates. The computational cost measured in terms of interface degrees of freedom is comparable to old and new Stokes discretizations.

Keywords: mixed finite element methods; incompressible flows; Stokes equations.

1. Introduction

We introduce a new method for the mixed stress formulation of the Stokes equations. Let u and p be the velocity and pressure, respectively. Assume that we are given an external force f, the kinematic viscosity ν and a bounded domain $\Omega \subset \mathbb{R}^d$ ($d = 2$ or 3) with Lipschitz boundary $\partial\Omega$. The standard velocity-pressure formulation

$$\begin{cases} -\operatorname{div}(\nu\nabla u) + \nabla p = f & \text{in } \Omega, \\ \operatorname{div}(u) = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{1.1}$$

can be reformulated by introducing the variable $\sigma = \nu \nabla u$ as follows:

$$\begin{cases} \dfrac{1}{\nu}\sigma - \nabla u = 0 & \text{in } \Omega, \\ \operatorname{div}(\sigma) - \nabla p = -f & \text{in } \Omega, \\ \operatorname{div}(u) = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \Gamma. \end{cases} \tag{1.2}$$

Many authors have studied this formulation previously, e.g., Farhloul & Fortin (1993, 1997, 2002); Farhloul (1995). The initial interest in this formulation as a numerical avenue appears to be due to the fact that fluid stresses can be computed merely by algebraic operations on σ (i.e., no differentiation of computed variables is needed). In this paper we study the discretization errors and certain interesting structure-preserving features of a new numerical method based on (1.2).

Although both formulations are formally equivalent, the mixed stress formulation (1.2) requires less regularity on the velocity field u. When considering a variational formulation of the classical velocity-pressure formulation (1.1), the proper spaces for the velocity and pressure are given by $H^1_0(\Omega, \mathbb{R}^d)$ and $L^2_0(\Omega)$, respectively. Here $H^1_0(\Omega, \mathbb{R}^d)$ is the standard vector-valued Sobolev space of order one with zero boundary conditions and $L^2_0(\Omega)$ is the space of square integrable functions with zero mean value. This pair of spaces fulfils the inf-sup condition or the Ladyzenskaja-Babuška-Brezzi (LBB) condition. Moreover, the divergence operator from $H^1_0(\Omega, \mathbb{R}^d)$ to $L^2_0(\Omega)$ is surjective. Finite element discretizations of the velocity-pressure formulation (1.1) is an active area of research (John *et al.*, 2017). While many pairs of discrete velocity-pressure spaces are known to satisfy the discrete LBB condition (needed to prove stability), not all of them have the property that the divergence operator from the discrete velocity space to the discrete pressure space is surjective. Methods that have this surjectivity property are particularly interesting, because they provide numerical velocity approximations that are exactly divergence free, leading to exact mass conservation.

Exact mass conservation (and consistency) further leads to a structure-preservation property called *pressure robustness*. A feature of solutions of (1.1) is that, when the load f changes irrotationally (i.e., when f is perturbed by a gradient field), then the fluid velocity u does not change (since the additional force can be balanced solely by a pressure gradient). Indeed, since divergence-free functions are L^2-orthogonal to the irrotational part of f, and since the velocity u is uniquely determined within the divergence-free subspace of $H^1_0(\Omega, \mathbb{R}^d)$, the velocity cannot be altered by irrotational changes in f. This property is not preserved by all finite element discretizations—see Linke (2014)—leading to velocity error estimates that depend on the pressure approximation. A practical manifestation of this is a phenomenon akin to 'locking', where the velocity error increases as $\nu \to 0$ (even if the pressure error remains under control). Methods that do not exhibit this limitation are called pressure robust methods. In the recent works of Linke (2012); Brennecke *et al.* (2015); Linke *et al.* (2016); Lederer *et al.* (2017b), considering different velocity and pressure spaces, it was shown that a (nonconforming) modification of the load (right-hand side) allows one to obtain optimal pressure-independent velocity error estimates.

An alternative to this load modification approach is the use of finite element spaces, which lead to exactly divergence-free velocity approximations. In this case no load modification is needed and the velocity error does not exhibit locking. A well-known example is the H^1-conforming Scott–Vogelius element. However, it demands a special barycentric triangulation of Ω. Another approach, leading to exactly divergence-free discretizations, is to abandon full H^1-conformity and retain only the continuity of the normal component of the velocity, i.e., use $H(\mathrm{div})$-conforming finite elements for approximating u instead of H^1-conforming finite elements. Such discretizations, tailored to approximate the incompressibility constraint properly, were introduced by Cockburn *et al.* (2005, 2007) and for the Brinkman problem by Könnö & Stenberg (2012). Therein, and also in the work by Lehrenfeld & Schöberl (2016), the H^1-conformity is treated in a weak sense and a hybrid discontinuous Galerkin method is constructed. Their choice of velocity and pressure space fulfils the discrete LBB condition and, moreover, Lederer & Schöberl (2017) shows that it is robust with respect to the polynomial order.

In this work the idea of employing an $H(\mathrm{div})$-conforming velocity space is taken to an infinite dimensional variational setting to obtain insights into possible spaces for σ. Obviously, such a variational formulation cannot be derived using the standard velocity-pressure formulation (1.1) as it demands too much regularity on the velocity. In contrast, the mixed stress formulation (1.2) is a perfect fit. It leads to a variational formulation requiring less regularity for u and a new function space for σ, namely $H(\mathrm{curl}\,\mathrm{div}, \Omega)$. We call this formulation the *mass conserving mixed formulation with stresses (MCS)*. To obtain a discretization, we design new nonconforming finite elements for $H(\mathrm{curl}\,\mathrm{div}, \Omega)$, motivated by the tangential displacement normal-normal stress (TDNNS) method for structural mechanics introduced by Sinwel (2009); Pechstein & Schöberl (2011, 2017). Even though the resulting method, called the *MCS method,* includes the introduction of another variable, the computational costs are comparable to other standard methods. In two dimensions, after a static condensation step, where local element degrees of freedom (dofs) are eliminated, the approximation of the velocity with polynomials of order k requires $k + 1$ coupling dofs on each element interface for the $H(\mathrm{div})$-conforming velocity space and k for the stress space. This is the same number as for the reduced stabilized (projected jumps) $H(\mathrm{div})$-conforming hybrid discontinuous Galerkin method introduced in Lehrenfeld & Schöberl (2016). By a small modification one could even reduce the coupling of the velocity space by considering only relaxed $H(\mathrm{div})$-conformity by the same technique utilized in Lederer *et al.* (2017a, 2018). Then the costs (for $k = 1$) are the same as for the lowest-order nonconforming H^1-based method. Similar cost comparisons can be made in three dimensions.

There appears to be multiple approaches for the analysis of our new scheme. In this paper we focus on one of these possible approaches, which uses a discrete H^1-like norm for u and an L^2 norm for σ. Even though u is approximated using $H(\mathrm{div})$-conforming elements, the use of the discrete H^1-like norm for velocity errors permits easy comparison with the classical velocity-pressure formulation. An analysis in more 'natural' norms, i.e., the $H(\mathrm{div})$-norm for u and $H(\mathrm{curl}\,\mathrm{div}, \Omega)$-norm for σ) is the topic of a forthcoming work. A discrete H^1-like norm for the velocity was also used in the works by Cockburn & Sayas (2014) and Fu *et al.* (2018), where related hybridized discontinuous Galerkin (DG) discretizations of the MCS formulation were introduced and analysed.

The paper is organized as follows. We begin with Section 2 where we define the notation and prove certain preliminary results that we shall use throughout this work. In Section 3 we present the new MCS variational formulation of the Stokes problem. Section 4 defines the discrete variational formulation and the MCS method. After revealing the continuity requirements across element interfaces necessary for being conforming in $H(\mathrm{curl}\,\mathrm{div}, \Omega)$, we then define new nonconforming finite elements for the σ variable in Section 5. All technical details needed to prove stability in certain discrete norms and convergence of the new method are included in Section 6. In Section 7 we present various numerical examples to illustrate the theory.

2. Preliminaries

In this section we define the notation we use throughout and establish properties of certain Sobolev spaces we shall need later.

Let $\Omega \subset \mathbb{R}^d$ be an open bounded domain with Lipschitz boundary $\Gamma := \partial\Omega$. Throughout, d is either 2 or 3. Let $\mathscr{D}(\Omega)$ or $\mathscr{D}(\Omega, \mathbb{R})$ denote the set of infinitely differentiable compactly supported real-valued functions on Ω and let $\mathscr{D}'(\Omega)$ denote the space of distributions as usual. To indicate vector and matrix-valued functions on Ω we include the range in the notation: $\mathscr{D}(\Omega, \mathbb{R}^d) = \{u : \Omega \to \mathbb{R}^d |\ u_i \in \mathscr{D}(\Omega)\}$. Such notation is extended in an obvious fashion to other function spaces as needed. For example, while $L^2(\Omega) = L^2(\Omega, \mathbb{R})$ denotes the space of square integrable real-valued functions on Ω, analogous vector and matrix-valued function spaces are defined by

$$L^2(\Omega, \mathbb{R}^d) := \left\{u : \Omega \to \mathbb{R}^d \middle| u_i \in L^2(\Omega)\right\} \quad \text{and} \quad L^2(\Omega, \mathbb{R}^{d\times d}) := \left\{\sigma : \Omega \to \mathbb{R}^{d\times d} \middle| \sigma_{ij} \in L^2(\Omega)\right\}.$$

Similarly, $\mathscr{D}'(\Omega, \mathbb{R}^d)$ denotes the space of distributions whose components are distributions in $\mathscr{D}'(\Omega)$, $H^m(\Omega, \mathbb{R}^{d\times d})$, denotes the space of matrix-valued functions whose entries are in the standard Sobolev space $H^m(\Omega)$ for any $m \in \mathbb{R}$, etc.

Certain differential operators have different definitions depending on context. By 'curl' we mean any of the following three differential operators:

$$\begin{aligned}
&\operatorname{curl}(\phi) = (-\partial_2\phi, \partial_1\phi)^{\mathrm{T}}, && \text{for } \phi \in \mathscr{D}'(\Omega, \mathbb{R}) \text{ and } d = 2,\\
&\operatorname{curl}(\phi) = -\partial_2\phi_1 + \partial_1\phi_2, && \text{for } \phi \in \mathscr{D}'(\Omega, \mathbb{R}^2) \text{ and } d = 2,\\
&\operatorname{curl}(\phi) = (\partial_2\phi_3 - \partial_3\phi_2, \partial_3\phi_1 - \partial_1\phi_3, \partial_1\phi_2 - \partial_2\phi_1)^{\mathrm{T}}, && \text{for } \phi \in \mathscr{D}'(\Omega, \mathbb{R}^3) \text{ and } d = 3,
\end{aligned}$$

where $(\cdot)^{\mathrm{T}}$ denotes the transpose and ∂_i abbreviates $\partial/\partial x_i$. The type of the operand determines which operator definition to apply in any context, so there will be no confusion. Similarly, ∇ is to be understood from context as an operator that results in either a vector whose components are $[\nabla\phi]_i = \partial_i\phi$ for $\phi \in \mathscr{D}'(\Omega, \mathbb{R})$ or a matrix whose entries are $[\nabla\phi]_{ij} = \partial_j\phi_i$ for $\phi \in \mathscr{D}'(\Omega, \mathbb{R}^d)$. Finally, in a similar manner, we understand $\operatorname{div}(\phi)$ as either $\sum_{i=1}^d \partial_i\phi_i$ for vector-valued $\phi \in \mathscr{D}'(\Omega, \mathbb{R}^d)$, or the row-wise divergence $\sum_{j=1}^d \partial_j\phi_{ij}$ for matrix-valued $\phi \in \mathscr{D}'(\Omega, \mathbb{R}^{d\times d})$.

Let $\tilde{d} = d(d-1)/2$ (so that $\tilde{d} = 1$ and 3 for $d = 2$ and 3, respectively). The following Sobolev spaces for $d = 2, 3$ are essential in our study:

$$\begin{aligned}
H(\operatorname{div}, \Omega) &= \{u \in L^2(\Omega, \mathbb{R}^d) : \operatorname{div}(u) \in L^2(\Omega)\},\\
H(\operatorname{curl}, \Omega) &= \{u \in L^2(\Omega, \mathbb{R}^d) : \operatorname{curl}(u) \in L^2(\Omega, \mathbb{R}^{\tilde{d}})\},\\
H^{-1}(\operatorname{curl}, \Omega) &= \{\phi \in H^{-1}(\Omega, \mathbb{R}^d) : \operatorname{curl}(\phi) \in H^{-1}(\Omega, \mathbb{R}^{\tilde{d}})\},\\
H(\operatorname{curl}\operatorname{div}, \Omega) &= \{\sigma \in L^2(\Omega, \mathbb{R}^{d\times d}) : \operatorname{curl}(\operatorname{div}(\sigma)) \in H^{-1}(\Omega, \mathbb{R}^{\tilde{d}})\}.
\end{aligned}$$

A well-known trace theorem permits us to define $H_0(\operatorname{div}, \Omega) = \{u \in H(\operatorname{div}, \Omega) : u \cdot n|_\Gamma = 0\}$. Here, n denotes the outward unit normal on Γ. In other occurrences it may denote the unit outward normal on boundaries of other domains determined from context.

The action of a continuous linear functional f on an element x of a topological space X is denoted by $\langle f, x\rangle_X$, e.g., the action of a distribution $F \in \mathscr{D}'(\Omega, \mathbb{R}^d)$ on a $\phi \in \mathscr{D}(\Omega, \mathbb{R}^d)$ is denoted by $\langle F, \phi\rangle_{\mathscr{D}(\Omega,\mathbb{R}^d)}$.

We omit the subscript denoting the space X in the duality pair $\langle\cdot,\cdot\rangle_X$ when it is obvious from the context. When X is a Hilbert space we use X^* to denote its dual space. Recall that $H_0^1(\Omega)^* = H^{-1}(\Omega)$. Note that any functional $f \in H^{-1}(\Omega)$ is also a distribution and that

$$\langle f,\phi\rangle_{H_0^1(\Omega)} = \langle f,\phi\rangle_{\mathscr{D}(\Omega)} \tag{2.1}$$

for all $\phi \in \mathscr{D}(\Omega)$. The inner product of X is denoted by $(\cdot,\cdot)_X$. When X is $L^2(\Omega), L^2(\Omega,\mathbb{R}^d)$ or $L^2(\Omega,\mathbb{R}^{d\times d})$ we abbreviate $(\cdot,\cdot)_X$ to simply $(\cdot,\cdot)$.

Lemma 2.1 If $F \in H_0(\mathrm{div},\Omega)^*$, then F is in $H^{-1}(\mathrm{curl},\Omega)$ and for all $v \in H_0^1(\Omega)$,

$$\langle \mathrm{curl}(F), v\rangle_{H_0^1(\Omega)} = \langle F, \mathrm{curl}(v)\rangle_{H_0(\mathrm{div},\Omega)}.$$

Proof. For any $F \in H_0(\mathrm{div},\Omega)^*$, by the Riesz representation theorem, there exists a $q^F \in H_0(\mathrm{div},\Omega)$, satisfying

$$\langle F, v\rangle_{H_0(\mathrm{div},\Omega)} = (q^F, v) + (\mathrm{div}(q^F), \mathrm{div}(v)) \tag{2.2}$$

for $v \in H_0(\mathrm{div},\Omega)$. Choosing $v \in \mathscr{D}(\Omega,\mathbb{R}^d)$ we conclude that F is the distribution $F = q^F - \nabla\,\mathrm{div}(q^F) \in H^{-1}(\Omega,\mathbb{R}^d)$. This implies that $\mathrm{curl}(F) = \mathrm{curl}(q^F) \in H^{-1}(\Omega,\mathbb{R}^{\tilde{d}})$. Thus, $F \in H^{-1}(\mathrm{curl},\Omega)$.

Moreover, for all $\phi \in \mathscr{D}(\Omega,\mathbb{R}^d)$, using (2.1),

$$\langle \mathrm{curl}(F),\phi\rangle_{H_0^1(\Omega,\mathbb{R}^d)} = \langle \mathrm{curl}(q^F),\phi\rangle_{H_0^1(\Omega,\mathbb{R}^d)} = \langle \mathrm{curl}(q^F),\phi\rangle_{\mathscr{D}(\Omega,\mathbb{R}^d)} = (q^F, \mathrm{curl}(\phi)).$$

By the density of $\mathscr{D}(\Omega,\mathbb{R}^d)$ in $H_0^1(\Omega,\mathbb{R}^d)$, we obtain

$$\langle \mathrm{curl}(F), v\rangle_{H_0^1(\Omega,\mathbb{R}^d)} = (q^F, \mathrm{curl}(v))$$

for all $v \in H_0^1(\Omega,\mathbb{R}^d)$. The proof is now complete due to (2.2). □

In the proof of the next result we use a 'regular decomposition' of $H_0(\mathrm{div},\Omega)$. Namely, there exists a $C > 0$ such that, given any $v \in H_0(\mathrm{div},\Omega)$, there is a $\phi_v \in H_0^1(\Omega,\mathbb{R}^{\tilde{d}})$ and a $z_v \in H_0^1(\Omega,\mathbb{R}^d)$ such that

$$v = \mathrm{curl}(\phi_v) + z_v, \qquad \|\phi_v\|_{H^1(\Omega,\mathbb{R}^{\tilde{d}})} + \|z_v\|_{H^1(\Omega,\mathbb{R}^d)} \le C\|v\|_{H(\mathrm{div},\Omega)}. \tag{2.3}$$

Many authors have stated this decomposition under various assumptions on Ω. Since there are too many to list here we content ourselves by pointing to the recent work of Demlow & Hirani (2014, Lemma 5), where one can find the result under the current assumptions on Ω and further references. Results like the following are known, for example, when $d = 2$ and Ω is simply connected—see Braess (2007, p. 366); Boffi *et al.* (2013, Eq. (10.4.52)); or Brezzi & Fortin (1986). Using the regular decomposition we are able to provide a general technique for proving such results below.

Theorem 2.2 The equality

$$H_0(\mathrm{div},\Omega)^* = H^{-1}(\mathrm{curl},\Omega)$$

holds algebraically and topologically.

Proof. Lemma 2.1 shows that $H_0(\mathrm{div},\Omega)^* \subseteq H^{-1}(\mathrm{curl},\Omega)$. To show $H^{-1}(\mathrm{curl},\Omega) \subseteq H_0(\mathrm{div},\Omega)^*$ let $g \in H^{-1}(\mathrm{curl},\Omega)$. Using the decomposition (2.3) set

$$\langle G, v\rangle_{H_0(\mathrm{div},\Omega)} := \langle \mathrm{curl}(g), \phi_v\rangle_{H_0^1(\Omega,\mathbb{R}^{\tilde{d}})} + \langle g, z_v\rangle_{H_0^1(\Omega,\mathbb{R}^d)}. \tag{2.4}$$

Due to the stability estimate of (2.3) G is a continuous linear functional in $H_0(\mathrm{div},\Omega)^*$. By Lemma 2.1 G is in $H^{-1}(\mathrm{curl},\Omega)$. It suffices to show G coincides with g (as an element of $H^{-1}(\Omega,\mathbb{R}^d)$). To this end let $w \in H_0^1(\Omega,\mathbb{R}^d)$. Since $H_0^1(\Omega,\mathbb{R}^d) \hookrightarrow H_0(\mathrm{div},\Omega)$ we have $\langle G, w\rangle_{H_0^1(\Omega,\mathbb{R}^d)} = \langle G, w\rangle_{H_0(\mathrm{div},\Omega)}$, so using decomposition (2.3),

$$\langle G, w\rangle_{H_0^1(\Omega,\mathbb{R}^d)} = \langle \mathrm{curl}(g), \phi_w\rangle_{H_0^1(\Omega,\mathbb{R}^{\tilde{d}})} + \langle g, z_w\rangle_{H_0^1(\Omega,\mathbb{R}^d)}.$$

Since both w and z_w are in $H_0^1(\Omega,\mathbb{R}^d)$, the equality $w = \mathrm{curl}(\phi_w) + z_w$ implies that $\mathrm{curl}(\phi_w) \in H_0^1(\Omega,\mathbb{R}^d)$. Hence, there is a $C > 0$ such that for any $w \in H_0^1(\Omega,\mathbb{R}^d)$ we have

$$\|\mathrm{curl}(\phi_w)\|_{H^1(\Omega,\mathbb{R}^d)} \le C\|w\|_{H^1(\Omega,\mathbb{R}^d)}. \tag{2.5}$$

Let $w_n \in \mathscr{D}(\Omega,\mathbb{R}^{\tilde{d}})$ converge to w in $H_0^1(\Omega,\mathbb{R}^{\tilde{d}})$ as $n \to \infty$, and further define the regular decomposition $w_n = \mathrm{curl}\,\phi_{w_n} + z_{w_n}$. By the construction of the regular decomposition components (see, e.g., Costabel & McIntosh, 2010), $\phi_{w_n} \in \mathscr{D}(\Omega,\mathbb{R}^{\tilde{d}})$. Moreover, using (2.1),

$$\langle \mathrm{curl}(g), \phi_{w_n}\rangle_{H_0^1(\Omega,\mathbb{R}^{\tilde{d}})} = \langle \mathrm{curl}(g), \phi_{w_n}\rangle_{\mathscr{D}(\Omega,\mathbb{R}^{\tilde{d}})} = \langle g, \mathrm{curl}(\phi_{w_n})\rangle_{\mathscr{D}(\Omega,\mathbb{R}^d)} = \langle g, \mathrm{curl}(\phi_{w_n})\rangle_{H_0^1(\Omega,\mathbb{R}^d)}.$$

Since $\mathrm{curl}(g)$ is in $H^{-1}(\Omega,\mathbb{R}^d)$ the left-most term converges to $\langle \mathrm{curl}(g), \phi_w\rangle_{H_0^1(\Omega,\mathbb{R}^d)}$. The right-most term must converge to $\langle g, \mathrm{curl}(\phi_w)\rangle_{H_0^1(\Omega,\mathbb{R}^d)}$ because (2.5) implies $\|\mathrm{curl}(\phi_{w_n} - \phi_w)\|_{H^1(\Omega,\mathbb{R}^d)} \to 0$. Thus, $\langle \mathrm{curl}(g), \phi_w\rangle_{H_0^1(\Omega,\mathbb{R}^d)} = \langle g, \mathrm{curl}(\phi_w)\rangle_{H_0^1(\Omega,\mathbb{R}^d)}$ and consequently,

$$\langle G, w\rangle_{H_0^1(\Omega,\mathbb{R}^d)} = \langle g, \mathrm{curl}(\phi_w) + z_w\rangle_{H_0^1(\Omega,\mathbb{R}^d)} = \langle g, w\rangle_{H_0^1(\Omega,\mathbb{R}^d)}.$$

This proves that $G = g$, so $g \in H_0(\mathrm{div},\Omega)^*$.

Finally, the stated topological equality follows if we show that $\|f\|_{H_0(\mathrm{div},\Omega)^*} \sim \|f\|_{H^{-1}(\mathrm{curl},\Omega)}$, where '$\sim$' denotes norm equivalence. Note that by (2.3) and triangle inequality, $\|\phi_v\|_{H^1(\Omega,\mathbb{R}^{\tilde{d}})} +$

$\|z_v\|_{H^1(\Omega,\mathbb{R}^d)} \sim \|v\|_{H(\mathrm{div},\Omega)}$. For any $f \in H_0(\mathrm{div},\Omega)^*$,

$$\begin{aligned}
\|f\|_{H_0(\mathrm{div},\Omega)^*} &= \sup_{v\in H_0(\mathrm{div},\Omega)} \frac{\langle f, v\rangle_{H_0(\mathrm{div},\Omega)}}{\|v\|_{H(\mathrm{div},\Omega)}} \\
&\sim \sup_{\phi\in H_0^1(\Omega,\mathbb{R}^{\tilde d}),\, z\in H_0^1(\Omega,\mathbb{R}^d)} \frac{\langle f, \mathrm{curl}(\phi)+z\rangle_{H_0(\mathrm{div},\Omega)}}{\|\phi\|_{H^1(\Omega,\mathbb{R}^{\tilde d})}+\|z\|_{H^1(\Omega,\mathbb{R}^d)}} && \text{by (2.3)} \\
&= \sup_{\phi\in H_0^1(\Omega,\mathbb{R}^{\tilde d}),\, z\in H_0^1(\Omega,\mathbb{R}^d)} \frac{\langle \mathrm{curl}(f), \phi\rangle_{H_0^1(\Omega,\mathbb{R}^{\tilde d})}+\langle f, z\rangle_{H_0^1(\Omega,\mathbb{R}^d)}}{\|\phi\|_{H^1(\Omega,\mathbb{R}^{\tilde d})}+\|z\|_{H^1(\Omega,\mathbb{R}^d)}} && \text{by Lemma 2.1} \\
&\sim \|f\|_{H^{-1}(\Omega,\mathbb{R}^d)}+\|\mathrm{curl}(f)\|_{H^{-1}(\Omega,\mathbb{R}^d)}.
\end{aligned}$$

Thus, the $H_0(\mathrm{div},\Omega)^*$ and $H^{-1}(\mathrm{curl},\Omega)$ norms are equivalent. □

3. Derivation of the MCS formulation of the Stokes equations

The goal of this section is to quickly derive a variational formulation of the mixed stress formulation of the Stokes system (1.2). Using the trace of a matrix $\mathrm{tr}(\tau) := \sum_{i=1}^d \tau_{ii}$ we define the deviatoric part by

$$\mathrm{dev}\,(\tau) = \tau - \frac{\mathrm{tr}(\tau)}{d}\,\mathrm{Id},$$

where Id denotes the identity matrix. Observe that due to $\mathrm{div}(u) = 0$, we have

$$\mathrm{dev}\,(\sigma) = \mathrm{dev}\,(\nu\nabla u) = \nu\nabla u - \frac{\nu}{d}\mathrm{tr}(\nabla u)\,\mathrm{Id} = \nu\left(\nabla u - \frac{1}{d}\,\mathrm{div}(u)\,\mathrm{Id}\right) = \nu\nabla u. \tag{3.1}$$

Thus, $\sigma = \nu\nabla u$ in (1.2) only represents the deviatoric part of the velocity gradient. Hence, we revise (1.2) to

$$\frac{1}{\nu}\mathrm{dev}\,(\sigma) - \nabla u = 0 \quad \text{in } \Omega, \tag{3.2a}$$

$$\mathrm{div}(\sigma) - \nabla p = -f \quad \text{in } \Omega, \tag{3.2b}$$

$$\mathrm{div}(u) = 0 \quad \text{in } \Omega, \tag{3.2c}$$

$$u = 0 \quad \text{on } \Gamma. \tag{3.2d}$$

We proceed to develop a variational formulation for (3.2).

For the reasons described in the introduction we want to derive a weak formulation, where the velocity u and the pressure p belong respectively to the following spaces.

$$V := H_0(\mathrm{div}, \Omega) = \{u \in H(\mathrm{div}, \Omega) : \; u \cdot n = 0 \text{ on } \Gamma\},$$
$$Q := L_0^2(\Omega) := \left\{ q \in L^2(\Omega) : \; \int_\Omega q \, \mathrm{d}x = 0 \right\}.$$

We begin with (3.2c). Multiplying (3.2c) with a test function $q \in Q$ and integrating over the domain Ω, we obtain the familiar equation

$$(\mathrm{div}(u), q) = 0. \tag{3.3}$$

Proceeding next to (3.2b), which must be tested with a $v \in V$, we see that σ in addition to being in $L^2(\Omega, \mathbb{R}^{d\times d})$, must also be such that $\mathrm{div}(\sigma)$ can continuously 'act' on v, i.e., $\mathrm{div}(\sigma) \in H(\mathrm{div}, \Omega)^*$. By Theorem 2.2, this is the same as requiring that

$$\mathrm{div}(\sigma) \in H^{-1}(\mathrm{curl}, \Omega). \tag{3.4}$$

Since any σ in $L^2(\Omega, \mathbb{R}^{d\times d})$ has $\mathrm{div}(\sigma) \in H^{-1}(\Omega, \mathbb{R}^d)$ the, nonredundant requirement that emerges from (3.4) is that $\mathrm{curl}(\mathrm{div}(\sigma)) \in H^{-1}(\Omega, \mathbb{R}^{\tilde{d}})$. This leads to the definition

$$\Sigma = \{\tau \in H(\mathrm{curl\,div}, \Omega) : \; \mathrm{tr}(\tau) = 0\},$$

where the requirement $\mathrm{tr}(\tau) = 0$ is motivated by (3.1). Thus, testing (3.2b) with a $v \in H_0(\mathrm{div}, \Omega)^*$ and integrating the pressure term by parts, we have

$$\langle \mathrm{div}(\sigma), v \rangle_{H_0(\mathrm{div},\Omega)} + (\mathrm{div}(v), p) = 0. \tag{3.5}$$

Finally, we multiply (3.2a) with a test function $\tau \in \Sigma$ to obtain $(\nu^{-1} \,\mathrm{dev}\,(\sigma), \tau) - (\nabla u, \tau) = 0$. Since

$$(\tau, \nabla v) = -\langle \mathrm{div}(\tau), v \rangle_{H_0(\mathrm{div},\Omega)}, \qquad \text{for all } \tau \in \Sigma, \; v \in H_0^1(\Omega, \mathbb{R}^d), \tag{3.6}$$

and $\mathrm{tr}(\tau) = 0$, using the fact that the exact velocity is in $H_0^1(\Omega, \mathbb{R}^d)$, we obtain

$$(\nu^{-1} \,\mathrm{dev}\,(\sigma), \mathrm{dev}\,(\tau)) + \langle \mathrm{div}(\tau), u \rangle_{H_0(\mathrm{div},\Omega)} = 0. \tag{3.7}$$

Note that in this derivation, while the normal trace of the velocity is an essential boundary condition included in the space V, the zero tangential velocity boundary conditions was incorporated weakly as a natural boundary condition in (3.7).

Collecting (3.7), (3.5) and (3.3) we summarize the derived weak formulation: given $f \in H_0(\mathrm{div}, \Omega)^*$ find $(\sigma, u, p) \in \Sigma \times V \times Q$ such that

$$\begin{cases} (\nu^{-1}\operatorname{dev}(\sigma), \operatorname{dev}(\tau)) + \langle \operatorname{div}(\tau), u\rangle_{H_0(\mathrm{div},\Omega)} = 0 & \text{for all } \tau \in \Sigma, \\ \langle \operatorname{div}(\sigma), v\rangle_{H_0(\mathrm{div},\Omega)} + (\operatorname{div}(v), p) = -\langle f, v\rangle_{H_0(\mathrm{div},\Omega)} & \text{for all } v \in V, \\ (\operatorname{div}(u), q) = 0 & \text{for all } p \in Q. \end{cases} \tag{3.8}$$

In the remainder of the paper we present an approximation of the weak formulation (3.8). It is possible to prove that (3.8) is well posed. However, since we shall focus on a discrete analysis of a nonconforming scheme based on (3.8), we shall not make direct use of the well posedness in this work. As a final remark on (3.8) we note that functions in Σ equal its deviatoric. Thus, we could remove 'dev' in the first term of (3.8). However, we keep it to remind ourselves that σ only approximates the deviatoric part of $\nu\nabla u$.

REMARK 3.1 (Boundary conditions). In this work we only consider homogeneous Dirichlet boundary conditions of the velocity, $u = 0$ on Γ. However, also other types of boundary conditions as for example slip boundary conditions for the velocity and homogeneous Neumann boundary conditions $(-\nu\nabla u + p\,\mathrm{Id})\cdot n = (-\sigma + p\,\mathrm{Id})\cdot n = 0$ are possible. A detailed analysis regarding this topic is included in a forthcoming work.

4. A discrete formulation

We present the discrete MCS method in this section. It is a nonconforming method based on the MCS weak formulation (3.8). We shall begin by understanding the conformity requirements of $H(\mathrm{curl\,div}, \Omega)$ and then present the method.

Suppose Ω is partitioned by a shape regular and quasiuniform triangulation $\mathscr{T}_h$ consisting of triangles and tetrahedrons in two and three dimensions, respectively. Here h denotes the maximum of the diameters of all elements in $\mathscr{T}_h$. Due to quasiuniformity $h \approx \mathrm{diam}(T)$ for any $T \in \mathscr{T}_h$. The set of element interfaces and boundaries are denoted by $\mathscr{F}_h$. This set is further split into facets on the domain boundary $F \subset \mathscr{F}_h \cap \Gamma =: \mathscr{F}_h^{\mathrm{ext}}$ and facets in the interior $F \subset \mathscr{F}_h \cap \Omega =: \mathscr{F}_h^{\mathrm{int}}$. There holds $\mathscr{F}_h = \mathscr{F}_h^{\mathrm{int}} \cup \mathscr{F}_h^{\mathrm{ext}}$. On each facet $F \in \mathscr{F}_h^{\mathrm{int}}$ we denote by $[\![\cdot]\!]$ the usual jump operator. For facets on the boundary the jump operator is just the identity. On each element boundary, and similarly on each facet on the global boundary, using the outward unit normal vector n, the normal and tangential trace of a smooth enough $u : \Omega \to \mathbb{R}^d$ is defined by

$$u_n = u \cdot n \quad \text{and} \quad u_t = u - u_n n.$$

According to this definition the normal trace is a scalar function and the tangential trace is a vector function. In two dimensions we may fix the symbol t to a unit tangent vector, obtained say by rotating n anti-clockwise by 90 degrees (thus $t = n^{\perp}$), so that $u_t = (u \cdot t)t$. In a similar manner for a smooth enough $\sigma : \Omega \to \mathbb{R}^{d\times d}$ we set

$$\sigma_{nn} = \sigma : (n \otimes n) = n^{\mathrm{T}}\sigma n \quad \text{and} \quad \sigma_{nt} = \sigma n - \sigma_{nn} n.$$

Thus, we have a scalar 'normal-normal component' and a vector-valued 'normal-tangential component', and in two dimensions t may be thought of as a unit tangent vector and $\sigma_{nt} = (t^{\mathrm{T}}\sigma n)t$.

We need to understand the conformity requirements of $H(\operatorname{curl}\operatorname{div}, \Omega)$. Just as continuity of the normal component across element interfaces is needed for $H(\operatorname{div}, \Omega)$-conformity, we shall see that *continuity of the normal-tangential component of tensors* is needed for $H(\operatorname{curl}\operatorname{div}, \Omega)$-conformity. Let

$$H^m(\mathscr{T}_h) := \{v \in L^2(\Omega) : v|_T \in H^m(T) \text{ for all } T \in \mathscr{T}_h\}.$$

For $\omega \subset \Omega$ we use $(\cdot,\cdot)_\omega$ to denote the inner product of $L^2(\omega), L^2(\omega, \mathbb{R}^d)$ or $L^2(\omega, \mathbb{R}^{d\times d})$ and, similarly, also $||\cdot||^2_\omega := (\cdot,\cdot)_\omega$.

Consider a σ in $H^1(\mathscr{T}_h, \mathbb{R}^{d\times d})$ and $\sigma_{nn}|_{\partial T} \in H^{1/2}(\partial T)$ for all elements $T \in \mathscr{T}_h$. Assume that the normal-tangential trace σ_{nt} is continuous across element interfaces. Then we claim that σ is in $H(\operatorname{curl}\operatorname{div}, \Omega)$. To see this, the definition of the distributional divergence and integration by parts yields

$$\langle \operatorname{div}(\sigma), \phi\rangle = -\int_\Omega \sigma : \nabla\phi \, \mathrm{d}x = \sum_{T\in\mathscr{T}_h} \int_T \operatorname{div}(\sigma)\cdot\phi \, \mathrm{d}x - \int_{\partial T} \sigma_n \cdot \phi \, \mathrm{d}s$$

for any $\phi \in \mathscr{D}(\Omega, \mathbb{R}^d)$. Splitting the boundary term into a tangential and a normal part, we obtain

$$\begin{aligned}\sum_{T\in\mathscr{T}_h} -\int_{\partial T} \sigma_n\cdot\phi \, \mathrm{d}s &= \sum_{T\in\mathscr{T}_h} -\int_{\partial T} \sigma_{nn}\phi_n \, \mathrm{d}s - \int_{\partial T} \sigma_{nt}\cdot\phi_t \, \mathrm{d}s \\ &= \sum_{T\in\mathscr{T}_h} -\int_{\partial T} \sigma_{nn}\phi_n \, \mathrm{d}s - \sum_{F\in\mathscr{F}_h} \int_F [\![\sigma_{nt}]\!]\cdot\phi_t \, \mathrm{d}s.\end{aligned}$$

As σ_{nt} is continuous across element interfaces the second term vanishes. Hence,

$$\begin{aligned}\langle \operatorname{div}(\sigma), \phi\rangle &= \sum_{T\in\mathscr{T}_h} \int_T \operatorname{div}(\sigma)\cdot\phi \, \mathrm{d}x - \int_{\partial T} \sigma_{nn}\phi_n \, \mathrm{d}s \\ &\le \sum_{T\in\mathscr{T}_h} ||\operatorname{div}(\sigma)||_T ||\phi||_T + ||\sigma_{nn}||_{H^{1/2}(\partial T)} ||\phi_n||_{H^{-1/2}(\partial T)} \le \mathrm{c}(\sigma)||\phi||_{H(\operatorname{div},\Omega)},\end{aligned} \tag{4.1}$$

where $\mathrm{c}(\sigma)$ is a constant depending on σ. Since $\mathscr{D}(\Omega, \mathbb{R}^d)$ is dense in $H_0(\operatorname{div}, \Omega)$ we conclude that $\operatorname{div}(\sigma)$ is in $H_0(\operatorname{div}, \Omega)^*$. Hence, by Theorem 2.2, $\sigma \in H(\operatorname{curl}\operatorname{div}, \Omega)$.

Accordingly, one of the sufficient conditions for conformity in $H(\operatorname{curl}\operatorname{div}, \Omega)$ is normal-tangential continuity. Full conformity is obtained under the further condition that $\sigma_{nn} \in H^{1/2}(\partial T)$, which demands more continuity: if the normal-normal component trace is continuous at vertices and edges in two and three dimensions, respectively, then the σ considered above would satisfy $\sigma_{nn} \in H^{1/2}(\partial T)$. If this latter constraint is relaxed much simpler elements can be constructed, as we shall see in Section 5.

The identity (4.1) also shows, due to density, that

$$\langle \operatorname{div}(\sigma), v\rangle_{H_0(\operatorname{div},\Omega)} = \sum_{T\in\mathscr{T}_h} [(\operatorname{div}(\sigma), v)_T - \langle v_n, \sigma_{nn}\rangle_{H^{1/2}(\partial T)}] \tag{4.2}$$

for all $v \in H_0(\operatorname{div}, \Omega)$. Identity (4.2) will motivate the definition of some of our bilinear forms later.

Let $\mathbb{P}^k(T)$ denote the space of polynomials of degree at most k restricted to T. Let $\mathbb{P}^k(T,\mathbb{R}^d)$ and $\mathbb{P}^k(T,\mathbb{R}^{d\times d})$ denote the space of vector and matrix-valued functions on T whose components are in $\mathbb{P}^k(T)$, and let

$$\mathbb{P}^k(\mathscr{T}_h) = \prod_{T\in\mathscr{T}_h} \mathbb{P}^k(T), \qquad \mathbb{P}^k(\mathscr{T}_h,\mathbb{R}^d) = \prod_{T\in\mathscr{T}_h} \mathbb{P}^k(T,\mathbb{R}^d), \qquad \mathbb{P}^k(\mathscr{T}_h,\mathbb{R}^{d\times d}) = \prod_{T\in\mathscr{T}_h} \mathbb{P}^k(T,\mathbb{R}^{d\times d}).$$

Define

$$\Sigma_h := \{\tau_h \in \mathbb{P}^k(\mathscr{T}_h,\mathbb{R}^{d\times d}) : \operatorname{tr}(\tau_h) = 0,\ [\![(\tau_h)_{nt}]\!] = 0,\ (\tau_h)_{nt} \in \mathbb{P}^{k-1}(F,\mathbb{R}^{d-1}) \text{ for all } F \in \mathscr{F}_h\}, \tag{4.3}$$

$$V_h := \mathbb{P}^k(\mathscr{T}_h,\mathbb{R}^d) \cap V, \tag{4.4}$$

$$Q_h := \mathbb{P}^{k-1}(\mathscr{T}_h) \cap Q. \tag{4.5}$$

Whereas the discrete velocity and pressure space are conforming subspaces of their continuous counterparts, the discrete stress space Σ_h is (slightly) nonconforming, $\Sigma_h \not\subset H(\operatorname{curl}\operatorname{div},\Omega)$. Furthermore, the normal-tangential component $(\tau_h)_{nt}|_F$ of any $\tau_h \in \mathbb{P}^k(\mathscr{T}_h,\mathbb{R}^{d\times d})$ is a tangential vector field whose values are in the tangent plane parallel to the facet F. By a slight abuse of notation we do not distinguish between this tangent plane and the isomorphic $\mathbb{R}^{d-1}$ (when we write statements like '$\tau_{nt} \in \mathbb{P}^{k-1}(F,\mathbb{R}^{d-1})$' above in (4.3)).

For the derivation of a discrete variational formulation with these spaces, we return to (3.8) and identify these bilinear forms:

$$\begin{aligned} &a : L^2(\Omega,\mathbb{R}^{d\times d}) \times L^2(\Omega,\mathbb{R}^{d\times d}) \to \mathbb{R}, && b_1 : V \times Q \to \mathbb{R},\\ &a(\sigma,\tau) := (\nu^{-1}\operatorname{dev}(\sigma), \operatorname{dev}(\tau)), && b_1(u,p) := (\operatorname{div}(u), p). \end{aligned}$$

To handle the terms with the divergence of stress variables, we define another bilinear form

$$b_2 : \{\tau \in H^1(\mathscr{T}_h,\mathbb{R}^{d\times d}) :\ [\![\tau_{nt}]\!] = 0\} \times \{v \in H^1(\mathscr{T}_h,\mathbb{R}^d) :\ [\![v_n]\!] = 0\} \to \mathbb{R}$$

motivated by the identity (4.2):

$$b_2(\tau,v) := \sum_{T\in\mathscr{T}_h} \int_T \operatorname{div}(\tau)\cdot v \,\mathrm{d}x - \sum_{F\in\mathscr{F}_h} \int_F [\![\tau_{nn}]\!] v_n \,\mathrm{d}s. \tag{4.6}$$

By integration by parts we find the equivalent representation

$$b_2(\tau,v) = -\sum_{T\in\mathscr{T}_h} \int_T \tau : \nabla v \,\mathrm{d}x + \sum_{F\in\mathscr{F}_h} \int_F \tau_{nt}\cdot [\![v_t]\!] \,\mathrm{d}s \tag{4.7}$$

since $[\![\tau_{nt}]\!] = 0$ and $[\![v_n]\!] = 0$. When trial and test functions are in the domain of these forms, the MCS weak form (3.8) can be rewritten in terms of these forms.

The discrete MCS method finds $(\sigma_h, u_h, p_h) \in \Sigma_h \times V_h \times Q_h$, satisfying

$$\begin{cases} a(\sigma_h, \tau_h) + b_2(\tau_h, u_h) = 0 & \text{for all } \tau_h \in \Sigma_h, \\ b_2(\sigma_h, v_h) + b_1(v_h, p_h) = (-f, v_h) & \text{for all } v_h \in V_h, \\ b_1(u_h, q_h) = 0 & \text{for all } q_h \in Q_h. \end{cases} \tag{MCS}$$

Note that the velocity space is the well-known BDM^k space—see for example Boffi *et al.* (2013). The pressure space is given by piecewise polynomials of one order less than the velocity space. By this we have the property $\mathrm{div}(V_h) = Q_h$. Therefore, any weakly divergence-free velocity field is also strongly divergence free:

$$(\mathrm{div}(u_h), q_h) = 0 \quad \Leftrightarrow \quad \mathrm{div}(u_h) = 0 \quad \text{in } \Omega. \tag{4.8}$$

Thus, any velocity field u_h computed from the system (MCS) is exactly divergence free.

To conclude this section we have shown that normal-tangential continuity appears to be natural for the matrix functions σ arising in fluid mechanics. There are other fields where matrix functions with tangential-tangential continuity arise naturally, as can be seen from the work of Christiansen (2011) and the dissertation of Lizao (2018). Finally, normal-normal continuity in matrix functions appears to be natural when solid mechanics is pursued in the approach of Pechstein & Schöberl (2017), as already noted previously.

5. Finite elements

The aim of this section is to construct local finite elements that yield the global finite element space Σ_h. We introduce degrees of freedom (linear functionals) on each element, which help us impose the normal-tangential continuity. We also give an explicit construction of a basis on a reference element, and provide an appropriate mapping to an arbitrary physical element of the triangulation. This is especially useful for the implementation, as there is no need to compute a dual shape function basis by biorthogonalization. The mapping technique permits easy extension to curved elements (although analysis of curved elements is beyond the scope of this work). We then complete this section by introducing an interpolation operator that we shall use in the error analysis of the next section.

The restriction of the function space Σ_h defined in (4.3) to a single element T gives the local finite element space $\Sigma_k(T) := \{\tau_h \in \mathbb{P}^k(T, \mathbb{R}^{d\times d}) : \mathrm{tr}(\tau_h) = 0, (\tau_h)_{nt} \in \mathbb{P}^{k-1}(F, \mathbb{R}^{d-1}) \text{ on all faces } F \in \mathscr{F}_T\}$, where $\mathscr{F}_T := \{F : F \subset \partial T\}$ is the set of element facets. Let

$$\mathbb{D} := \{M \in \mathbb{R}^{d\times d} : (M : \mathrm{Id}) = 0\}.$$

Then we may equivalently write

$$\Sigma_k(T) = \{\tau_h \in \mathbb{P}^k(T, \mathbb{D}) : \ (\tau_h)_{nt} \in \mathbb{P}^{k-1}(F, \mathbb{R}^{d-1}) \text{ on all faces } F \in \mathscr{F}_T\}. \tag{5.1}$$

We proceed to study this space in detail, beginning with $\mathbb{D}$.

5.1 *Trace-free matrices*

As a first step, we construct a basis for the space of matrices $\mathbb{D}$ particularly suited to study normal-tangential components on facets. Let V_i, $i \in \mathscr{V}$, denote the vertices of T, where $\mathscr{V} := \{0, 1, 2\}$ and $\mathscr{V} := \{0, 1, 2, 3\}$ in two and three dimensions, respectively. Further, let F_i be the face opposite to the vertex V_i with the normal vector given by n_i. The unit tangential vectors along edges are $t_{ij} := (V_i - V_j)/|V_i - V_j|$. Finally, let λ_i be the unique barycentric coordinate function that equals one at the vertex V_i. When $d = 2$ define three constant matrix functions, one for each $i \in \mathscr{V}$,

$$S^i := \operatorname{dev}\left(\nabla\lambda_{i+1} \otimes \operatorname{curl}(\lambda_{i+2})\right), \tag{5.2}$$

where the indices $i + 1$ and $i + 2$ are taken modulo 3. When $d = 3$, for each $i \in \mathscr{V}$, we define the following two constant matrix functions

$$S_0^i := \operatorname{dev}\left(\nabla\lambda_{i+1} \otimes (\nabla\lambda_{i+2} \times \nabla\lambda_{i+3})\right), \quad S_1^i := \operatorname{dev}\left(\nabla\lambda_{i+2} \otimes (\nabla\lambda_{i+3} \times \nabla\lambda_{i+1})\right), \tag{5.3}$$

taking the indices $i + 1$, $i + 2$ and $i + 3$ modulo 4.

Lemma 5.1 The sets $\{S^i : i \in \mathscr{V}\}$ and $\{S_q^i : i \in \mathscr{V},\ q = 0, 1\}$ form a basis of $\mathbb{D}$ when $d = 2$ and 3, respectively. Moreover, the normal-tangential component of S^i and S_q^i vanishes everywhere on the element boundary, except on F_i,

$$S^i_{nt}|_{F_j} = 0, \qquad (S^i_q)_{nt}|_{F_j} = 0, \qquad i \neq j, \ \ F_j \in \mathscr{F}_T, \ \ i,j \in \mathscr{V},$$

while on F_i it does not vanish. When $i = j \in \mathscr{V}$ and $d = 3$,

$$t^{\mathrm{T}}_{i+2,i+3} S^i_0 n_i = 0, \quad t^{\mathrm{T}}_{i+1,i+2} S^i_0 n_i \neq 0, \quad t^{\mathrm{T}}_{i+3,i+1} S^i_0 n_i \neq 0, \tag{5.4a}$$

$$t^{\mathrm{T}}_{i+2,i+3} S^i_1 n_i \neq 0, \quad t^{\mathrm{T}}_{i+1,i+2} S^i_1 n_i \neq 0, \quad t^{\mathrm{T}}_{i+3,i+1} S^i_1 n_i = 0. \tag{5.4b}$$

Proof. The first statement of the lemma follows once we prove the remaining statements. Indeed, the linear independence of the given sets follows by examining their normal-tangential components facet by facet using the remaining statements. The spanning property follows by counting.

To prove the remaining statements we start with the two-dimensional case. We define

$$s_{i,j} = \operatorname{dev}\left(\nabla\lambda_i \otimes \operatorname{curl}(\lambda_j)\right).$$

Then $s_{i+1,i+2} = S^i$. Since the nt-component of the identity vanishes for any $p \in \mathscr{V}$ and any $t_p \in \operatorname{curl}(\lambda_p)$

$$t_p^{\mathrm{T}} s_{i,j} n_p = t_p^{\mathrm{T}}\left[\nabla\lambda_i \otimes \operatorname{curl}(\lambda_j)\right] n_p = (\nabla\lambda_i \cdot t_p)(\nabla\lambda_j \cdot t_p).$$

All the stated properties in the two-dimensional case now follow easily from this identity together with the fact that T is not degenerate.

Next, consider the $d = 3$ case. Let $s_{i,j,k} = \operatorname{dev}\big(\nabla\lambda_i \otimes (\nabla\lambda_j \times \nabla\lambda_k)\big)$. If i, j, k, l is any permutation of $\mathscr{V}$, by elementary manipulations, we see that for any $p \in \mathscr{V}$ and any $t_p \in n_p^{\perp}$,

$$t_p^{\mathrm{T}} s_{i,j,k} n_p = c(n_i \cdot t_p)(t_{il} \cdot n_p) \tag{5.5}$$

for some $c \neq 0$. Therefore, on any facet F_p, we have $t_p^{\mathrm{T}}(S_0^i)n_p = t_p^{\mathrm{T}}(s_{i+1,i+2,i+3})n_p = c(n_{i+1} \cdot t_p)(t_{i+1,i} \cdot n_p)$, which vanishes for all $p \neq i$ since $n_{i+1} \cdot t_{i+1} = 0$ and $t_{i+1,i} \cdot n_{i+2} = t_{i+1,i} \cdot n_{i+3} = 0$. Similarly, we conclude that $(S_1^i)_{nt} = 0$ on all facets except F_i. Since (5.5) also implies

$$t_{jk}^{\mathrm{T}} s_{i,j,k} n_l = 0, \quad t_{ki}^{\mathrm{T}} s_{i,j,k} n_l \neq 0, \quad t_{ji}^{\mathrm{T}} s_{i,j,k} n_l \neq 0,$$

the statements in (5.4) also follow. □

5.2 *Normal-tangential bubbles*

Let the element space of interior normal-tangential bubbles be defined by

$$\mathscr{B}_k(T) := \big\{\tau_h \in \Sigma_k(T) : (\tau_h)_{nt} = 0\big\}$$

LEMMA 5.2 Any $b \in \mathscr{B}_k(T)$ can be expressed as either

$$b = \sum_{i\in\mathscr{V}} \mu_i \lambda_i S^i \quad \text{or} \quad b = \sum_{q=0}^{1} \sum_{i\in\mathscr{V}} \mu_i^q \, \lambda_i S_q^i, \tag{5.6}$$

for $d = 2$ or 3, respectively, where $\mu_i, \mu_i^0, \mu_i^1 \in \mathbb{P}^{k-1}(T)$. Consequently,

$$\dim \mathscr{B}_k(T) = \begin{cases} \dfrac{3}{2}k(k+1), & \text{if } d = 2, \\[2ex] \dfrac{8}{6}k(k+1)(k+2), & \text{if } d = 3. \end{cases}$$

Proof. We only show the proof in the $d = 2$ case as the $d = 3$ case is similar. By Lemma 5.1 applied to the matrix $b(x)$, we obtain

$$b(x) = \sum_{i\in\mathscr{V}} a_i(x) S^i, \tag{5.7}$$

and matching degrees, we conclude that $a_i \in \mathbb{P}^k(T)$. Let c_i equal the constant value of $S_{nt}^i|_{F_i}$, which is nonzero by Lemma 5.1. Then $t_i^{\mathrm{T}} b(x) n_i = c_i a_i(x) = 0$ for all $x \in F_i$. Since $a_i(x)$ vanishes on F_i, it must take the form $a_i(x) = \mu_i(x)\lambda_i(x)$ for some $\mu_i \in \mathbb{P}^{k-1}(T)$. This proves (5.6).

The dimension count follows from (5.6): again considering only the $d = 2$ case, since $\mu_i \in \mathbb{P}^{k-1}(T)$ and $\{\lambda_i S^i : i \in \mathscr{V}\}$ is a linearly independent set, the expansion in (5.6) shows that $\dim \mathscr{B}_k(T)$ is $3 \times \dim \mathbb{P}^{k-1}(T)$. □

5.3 *Mappings*

Suppose $\widehat{T}$ is the unit simplex ($d = 2$ or 3) and $T \in \mathscr{T}_h$. Let $\phi_T : \widehat{T} \to T$ be an affine homeomorphism and set $F_T := \phi_T'$. Due to the shape regularity of the mesh

$$||F_T||_\infty \approx h \quad \text{and} \quad ||F_T^{-1}||_\infty \approx h^{-1} \quad \text{and} \quad |\det(F_T)| \approx h^d. \tag{5.8}$$

The proper transformation for functions in the $H(\mathrm{div})$-conforming velocity space V_h is the Piola transformation given by $\mathscr{P}(\hat{u}_h) := (\det F_T)^{-1} F_T \hat{u}_h$, where $\hat{u}_h$ is a given polynomial on the reference element. The Piola map preserves the normal components on facets, so is useful for enforcing normal continuity. For functions demanding tangential continuity, the proper transformation is the covariant transformation given by $\mathscr{C}(\hat{u}_h) := F_T^{-\mathrm{T}} \hat{u}_h$. Therefore, to enforce the normal-tangential continuity required of tensors in Σ_h, we combine the above two transformations and define

$$\mathscr{M}(\hat{\sigma}_h) := \frac{1}{\det(F_T)} F_T^{-\mathrm{T}} (\hat{\sigma}_h \circ \phi_T^{-1}) F_T^{\mathrm{T}}, \tag{5.9}$$

where $\hat{\sigma}_h \in \Sigma_k(\hat{T})$. Of particular interest to us is how the normal-tangential components on facets $F \in \mathscr{F}_T$ map. To study this we use the restrictions of the map ϕ_T to a reference facet $\hat{F}$ as well as to a reference edge $\hat{E}$ (a $d-2$ subsimplex) in the $d = 3$ case, denoted by $\phi_T|_{\hat{F}}$ and $\phi_T|_{\hat{E}}$, respectively. Their gradients are denoted by $F_T^F = (\phi_T|_{\hat{F}})'$ and $F_T^E = (\phi_T|_{\hat{E}})'$. In the next result $\hat{n}$ and n denote the outward unit normals vector on $\hat{F}$ and F, respectively, while $\hat{t}$ denotes a unit tangent vector along $\hat{E}$ (when $d = 3$) or $\hat{F}$ (when $d = 2$ and, similarly, t denotes a unit tangent vector along E or F.

LEMMA 5.3 Using the above notation and letting $\tau = \mathscr{M}(\hat{\tau})$, we have

$$c\, t^{\mathrm{T}} \tau n = \hat{t}^{\mathrm{T}} \hat{\tau} \hat{n}, \qquad \text{where } c = \begin{cases} \det(F_T^F)^2 & \text{if } d = 2, \\ \det(F_T^F)\det(F_T^E) & \text{if } d = 3. \end{cases}$$

Furthermore,

$$\mathrm{tr}(\hat{\tau}) = 0 \quad \Leftrightarrow \quad \mathrm{tr}(\tau) = 0.$$

Proof. The unit normals and tangents on the reference and mapped configurations are related by

$$n = \frac{\det(F_T)}{\det(F_T^F)} F_T^{-\mathrm{T}} \hat{n} \quad \text{and} \quad t = \frac{1}{\det(F_T^E)} F_T \hat{t},$$

with the understanding that in two dimensions we should replace F_T^E by F_T^F. Then

$$t^{\mathrm{T}} \tau n = \frac{1}{\det(F_T^E)} \hat{t}^{\mathrm{T}} F_T^{\mathrm{T}} \frac{1}{\det(F_T)} F_T^{-\mathrm{T}} \hat{\tau} F_T^{\mathrm{T}} \frac{\det(F_T)}{\det(F_T^F)} F_T^{-\mathrm{T}} \hat{n} = \frac{1}{\det(F_T^E)\det(F_T^F)} \hat{t}^{\mathrm{T}} \hat{\tau} \hat{n}.$$

Finally, the statement on traces follows from $\mathrm{tr}(F_T^{-\mathrm{T}} \hat{\tau} F_T^{\mathrm{T}}) = \mathrm{tr}(\hat{\tau})$. □

5.4 *Definition of the finite element*

We define the local finite element in the formal style of Ciarlet (2002) (also adopted in other texts, e.g., Ern & Guermond, 2004; Braess, 2013) as a triple $(T, \varSigma_k(T), \varPhi(T))$, where the geometrical element T is either a triangle or a tetrahedron, the space $\varSigma_k(T)$ is defined by (5.1) and $\varPhi(T)$ is a set of linear functionals representing the dofs defined as follows. The first group of dofs is associated to the set of element facets $\mathscr{F}_T$, the set of $d-1$ subsimplices of T: for each $F \in \mathscr{F}_T$, define

$$\varPhi^F(\tau) := \left\{ \int_F \tau_{nt} \cdot r \, \mathrm{d}s : \; r \in \mathbb{P}^{k-1}(F, \mathbb{R}^{d-1}) \right\}. \tag{5.10}$$

The next group is the set of interior dofs given by

$$\varPhi^T(\tau) := \left\{ \int_T \tau : F_T(\hat{\eta} \circ \phi_T^{-1}) F_T^{-1} \, \mathrm{d}x : \; \hat{\eta} \in \mathscr{B}_k(\hat{T}) \right\}. \tag{5.11}$$

Then set

$$\varPhi(T) := \varPhi^T \cup \{\varPhi^F : \; F \in \mathscr{F}_T\}. \tag{5.12}$$

We proceed to prove that this set of dofs is unisolvent and that the number of degrees of freedom matches the dimension of $\varSigma_k(T)$.

THEOREM 5.4 The triple $(T, \varSigma_k(T), \varPhi(T))$ defines a finite element and

$$\dim(\varSigma_k(T)) = \begin{cases} \dfrac{3}{2}(k+1)(k+2) - 3, & \text{if } d = 2, \\ \dfrac{8}{6}(k+1)(k+2)(k+3) - 8(k+1), & \text{if } d = 3. \end{cases}$$

Proof. To prove the unisolvency of the dofs, consider a $\tau_h \in \varSigma_k(T)$ satisfying $\phi(\tau_h) = 0$ for all $\phi \in \varPhi(T)$. As $(\tau_h)_{nt} \in \mathbb{P}^{k-1}(F, \mathbb{R}^{d-1})$ the facet dofs $\phi(\tau_h) = 0$ imply that $\tau_h \in \mathscr{B}_k(T)$. The interior dofs then yield

$$0 = \int_T \tau_h : F_T \hat{\eta} F_T^{-1} = \int_T F_T^{\mathrm{T}} \tau_h F_T^{-\mathrm{T}} : \hat{\eta} = \int_T (\det F_T)^{-1} \mathscr{M}^{-1}(\tau_h) : \hat{\eta} = \int_{\hat{T}} \mathscr{M}^{-1}(\tau_h) : \hat{\eta}$$

for all $\hat{\eta} \in \mathscr{B}_k(\hat{T})$. By Lemma 5.3 $\mathscr{M}^{-1}(\tau_h)$ is in $\mathscr{B}_k(\hat{T})$, so this yields $\mathscr{M}^{-1}(\tau_h) = 0$ and thus $\tau_h = 0$.

It only remains to prove the dimension count. The dimension of $\varSigma_k(T)$ is given by $\dim \mathbb{P}^k(T, \mathbb{D})$ minus the number of linearly independent conditions represented by the constraints $(\tau_h)_{nt} \in \mathbb{P}^{k-1}(F, \mathbb{R}^{d-1})$ for all $F \in \mathscr{F}_T$ that every $\tau_h \in \varSigma_k(T)$ must satisfy. Therefore,

$$\begin{aligned} \dim(\varSigma_k(T)) &\geq \dim \mathbb{P}^k(T, \mathbb{D}) - \dim \left[\mathbb{P}^k(F, \mathbb{R}^{d-1}) \setminus \mathbb{P}^{k-1}(F, \mathbb{R}^{d-1}) \right] \\ &= (d^2 - 1) \dim \mathbb{P}^k(T) - (d+1)(d-1) \dim \left[\mathbb{P}^k(F) \setminus \mathbb{P}^{k-1}(F) \right]. \end{aligned}$$

Let $N_{\varSigma_k}$ denote the number on the right-hand side. Using Lemma 5.2 to count the number of dofs in $\varPhi(T)$ we find that it coincides with $N_{\varSigma_k}$. Since $N_{\varSigma_k}$ linear functionals on $\varSigma_k(T)$ are unisolvent, we conclude that $\dim(\Sigma_k(T)) = N_{\Sigma_k}$, which after simplification agrees with the statement of the theorem. □

5.5 *Construction of shape functions*

In view of the previous results we can now write down shape functions in barycentric coordinates. It is not difficult to see that on any triangle T the set of functions

$$\lambda_{i+1}^{\alpha_1}\lambda_{i+2}^{\alpha_2}S^i, \qquad \lambda_i^{\beta_0}\lambda_{i+1}^{\beta_1}\lambda_{i+2}^{\beta_2}(\lambda_i S^i), \tag{5.13}$$

for all $i \in \mathscr{V}$, and all multi-indices (α_1, α_2) and $(\beta_0, \beta_1, \beta_2)$, with $\alpha_i \geq 0,\ \beta_i \geq 0$ having length $\alpha_1 + \alpha_2 = \beta_0 + \beta_1 + \beta_2 = k - 1$, form a basis for $\Sigma_k(T)$. Similarly, when T is a tetrahedron, the following set is a basis for $\Sigma_k(T)$:

$$\lambda_{i+1}^{\alpha_1}\lambda_{i+2}^{\alpha_2}\lambda_{i+3}^{\alpha_3}S_q^i, \qquad \lambda_i^{\beta_0}\lambda_{i+1}^{\beta_1}\lambda_{i+2}^{\beta_2}\lambda_{i+3}^{\beta_3}(\lambda_i S_q^i), \tag{5.14}$$

for all $i \in \mathscr{V}$, $q = 0, 1$, and all multi-indices $(\alpha_1, \alpha_2, \alpha_3)$ and $(\beta_0, \beta_1, \beta_2, \beta_3)$, with $\alpha_i \geq 0,\ \beta_i \geq 0$ having length $\alpha_1+\alpha_2+\alpha_3 = \beta_0+\beta_1+\beta_2+\beta_3 = k-1$. Instead of proving the linear independence of functions in (5.13) or (5.14), in the remainder of this section, we opt to do so for another set of reference element shape functions that we have implemented. By using a Dubiner basis instead of barycentric monomials the ensuing construction produces better conditioned matrices.

We start by defining some basic notation needed for the construction. The reference element is given by

$$\widehat{T} := \{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1, x_2 \text{ and } x_1 + x_2 \leq 1\} \qquad \text{for} \qquad d = 2,$$
$$\widehat{T} := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : 0 \leq x_1, x_2, x_3 \text{ and } x_1 + x_2 + x_3 \leq 1\} \qquad \text{for} \qquad d = 3.$$

For $d = 2$ we further define the reference faces and the corresponding normal and tangential vectors (see the left picture in Fig. 1) by

$$\hat{F}_0 = \{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1, x_2 \leq 1, x_1 + x_2 = 1\}, \qquad \hat{n}_0 := \frac{1}{\sqrt{2}}(1,1)^{\mathrm{T}}, \qquad \hat{t}_0 := \frac{1}{\sqrt{2}}(-1,1)^{\mathrm{T}},$$
$$\hat{F}_1 = \{(0, x_2) \in \mathbb{R}^2 : 0 \leq x_2 \leq 1\}, \qquad \hat{n}_1 := (-1,0)^{\mathrm{T}}, \qquad \hat{t}_1 := (0,-1)^{\mathrm{T}},$$
$$\hat{F}_2 = \{(x_1, 0) \in \mathbb{R}^2 : 0 \leq x_1 \leq 1\}, \qquad \hat{n}_2 := (0,-1)^{\mathrm{T}}, \qquad \hat{t}_2 := (1,0)^{\mathrm{T}}.$$

For the three-dimensional case we have

$$\hat{F}_0 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : 0 \leq x_1, x_2, x_3 \leq 1, x_1 + x_2 + x_3 = 1\},$$
$$\hat{F}_1 = \{(0, x_2, x_3) \in \mathbb{R}^3 : 0 \leq x_2, x_3 \leq 1, 0 \leq x_2 + x_3 \leq 1\},$$
$$\hat{F}_2 = \{(x_1, 0, x_3) \in \mathbb{R}^2 : 0 \leq x_1, x_3 \leq 1, 0 \leq x_1 + x_3 \leq 1\},$$
$$\hat{F}_3 = \{(x_1, x_2, 0) \in \mathbb{R}^2 : 0 \leq x_1, x_2 \leq 1, 0 \leq x_1 + x_2 \leq 1\},$$

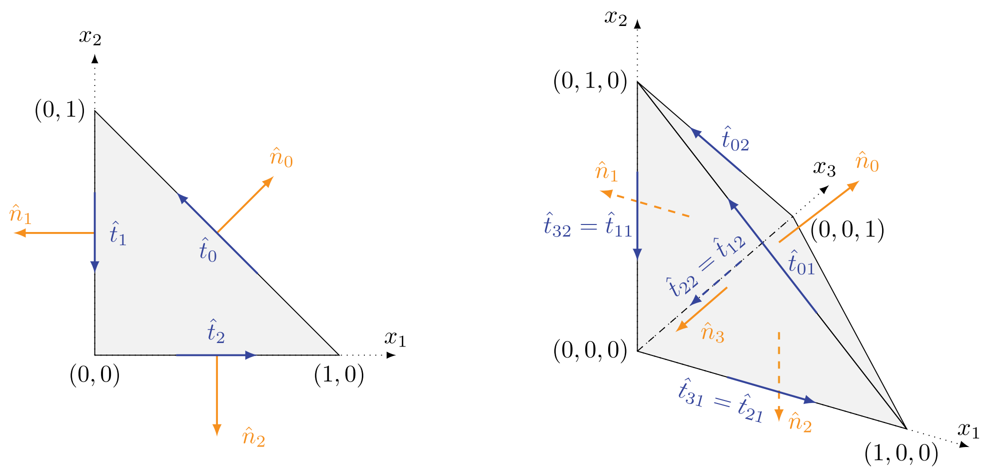

FIG. 1. The reference element and the corresponding normal and tangential vectors in two and three space dimensions.

with the associated normal and tangential vectors (see the right picture in Fig. 1)

$$\begin{aligned}
\hat{n}_0 &:= \frac{1}{\sqrt{3}}(1,1,1)^{\mathrm{T}}, & \hat{t}_{01} &:= \frac{1}{\sqrt{2}}(-1,1,0)^{\mathrm{T}}, & \hat{t}_{02} &:= \frac{1}{\sqrt{2}}(0,1,-1)^{\mathrm{T}},\\
\hat{n}_1 &:= (-1,0,0)^{\mathrm{T}}, & \hat{t}_{11} &:= (0,-1,0)^{\mathrm{T}}, & \hat{t}_{12} &:= (0,0,-1)^{\mathrm{T}},\\
\hat{n}_2 &:= (0,-1,0)^{\mathrm{T}}, & \hat{t}_{21} &:= (1,0,0)^{\mathrm{T}}, & \hat{t}_{22} &:= (0,0,-1)^{\mathrm{T}},\\
\hat{n}_3 &:= (0,0,-1)^{\mathrm{T}}, & \hat{t}_{31} &:= (1,0,0)^{\mathrm{T}}, & \hat{t}_{32} &:= (0,-1,0)^{\mathrm{T}}.
\end{aligned}$$

In Section 5.1 we presented the construction of element-wise constant matrices. Applying these techniques on the reference element (including a scaling with a proper constant) we derive for $d = 2$ the matrices given by

$$\hat{S}^0 := \sqrt{2}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \hat{S}^1 := \begin{pmatrix} 0.5 & 0 \\ 1 & -0.5 \end{pmatrix} \quad \text{and} \quad \hat{S}^2 := \begin{pmatrix} 0.5 & -1 \\ 0 & -0.5 \end{pmatrix}, \tag{5.15}$$

and for $d = 3$ the matrices

$$\begin{aligned}
&\hat{S}^0_0 = \sqrt{6}\begin{pmatrix} \frac{-2}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}, \hat{S}^1_0 = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 1 & \frac{-2}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}, \hat{S}^2_0 = \begin{pmatrix} \frac{-2}{3} & 1 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}, \hat{S}^3_0 = \begin{pmatrix} \frac{-2}{3} & 0 & 1 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix},\\
&\hat{S}^0_1 = \sqrt{6}\begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{-2}{3} \end{pmatrix}, \hat{S}^1_1 = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 1 & 0 & \frac{-2}{3} \end{pmatrix}, \hat{S}^2_1 = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 1 & \frac{-2}{3} \end{pmatrix}, \hat{S}^3_1 = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{-2}{3} & 1 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}.
\end{aligned} \tag{5.16}$$

Note that in order to follow the ideas described in Section 5.1, we took a particular choice of the numbering of the vertices of $\hat{T}$ and the corresponding tangential vectors. Similarly as in Lemma 5.1

elementary calculations show that

$$\begin{aligned} &\hat{t}_j^{\mathrm{T}}\hat{S}^i\hat{n}_j = \delta_{ij} \quad &&\text{and} \quad &&\hat{t}_j^{\mathrm{T}}\lambda_i\hat{S}^i\hat{n}_j = 0 \quad &&\text{for} \quad &&i,j = 0,1,2, \\ &\hat{t}_{jl}^{\mathrm{T}}\hat{S}_q^i\hat{n}_j = \delta_{ij}\delta_{ql} \quad &&\text{and} \quad &&\hat{t}_{jl}^{\mathrm{T}}\lambda_i\hat{S}_q^i\hat{n}_j = 0 \quad &&\text{for} \quad &&i,j \in 0,1,2,3 \quad \text{and} \quad q,l = 0,1 \end{aligned} \tag{5.17}$$

and that $\{\hat{S}^i : i = 0,1,2\}$ and $\{\hat{S}_q^i : i = 0,1,2,3; q = 0,1\}$ is a basis for $\mathbb{D}$ in two and three dimensions, respectively. Based on these constant matrices we now construct shape function for the local stress space $\Sigma_k(\hat{T})$.

We start with the two-dimensional case. Let $l_i(x_1)$ be the Legendre polynomial of order i and let $l_i^S(x_1,x_2) := x_2^i l_i(x_1/x_2)$ be the scaled Legendre polynomial of order i. Further, let $p_i^j(x_1)$ be the Jacobi polynomial of order i with coefficients $\alpha = j$, $\beta = 0$. For a detailed definition we refer to the worksby Abramowitz (1974); Andrews *et al.* (1999). We then define

$$\hat{r}_{ij}(\lambda_\alpha,\lambda_\beta,\lambda_\gamma) := l_i^S(\lambda_\beta - \lambda_\alpha, \lambda_\alpha + \lambda_\beta)p_j^{2i+1}(\lambda_\gamma - \lambda_\alpha - \lambda_\beta). \tag{5.18}$$

The polynomials $\hat{r}_{ij}(\lambda_\alpha,\lambda_\beta,\lambda_\gamma)$ with $0 \le i+j \le k$ and an arbitrary permutation (α,β,γ) of $(0,1,2)$ form a basis of the polynomial space $\mathbb{P}^k(\hat{T},\mathbb{R})$. Next note that p_0^{2i+1} is constant, thus $\hat{r}_{ij}(\lambda_\alpha,\lambda_\beta,\lambda_\gamma) = \hat{r}_{i0}(\lambda_\alpha,\lambda_\beta)$. Then there holds that for $0 \le i \le k$ the restriction of the polynomials $\hat{r}_{i0}(\lambda_{j+1},\lambda_{j+2})|_{\hat{F}_j}$, where the indices $j+1$ and $j+2$ of the barycentric coordinate functions are taken modulo 3, forms a basis of the polynomial space $\mathbb{P}^k(\hat{F}_j,\mathbb{R})$ (see Karniadakis & Sherwin, 2013, Chapter 3.2 or in Dubiner, 1991). By this we define a local basis of the stress space by

$$\begin{aligned} \hat{\Psi}_k^F &:= \{\hat{S}^j\hat{r}_{i0}(\lambda_{j+1},\lambda_{j+2}) : j = 0,1,2 \text{ and } 0 \le i \le k-1\}, \\ \hat{\Psi}_k^T &:= \{\lambda_j\hat{S}^j\hat{r}_{il}(\lambda_0,\lambda_1,\lambda_2) : j = 0,1,2 \text{ and } 0 \le i+l \le k-1\}. \end{aligned}$$

For $d = 3$ we define similarly as before

$$\begin{aligned} &\hat{r}_{ijl}(\lambda_\alpha,\lambda_\beta,\lambda_\gamma,\lambda_\delta) \\ &\quad := l_i^S(\lambda_\beta - \lambda_\alpha, \lambda_\alpha + \lambda_\beta)p_j^{2i+1,S}(\lambda_\gamma - \lambda_\alpha - \lambda_\beta, \lambda_\gamma + \lambda_\alpha + \lambda_\beta)p_l^{2i+2j+2,S}(\lambda_\delta - \lambda_\alpha - \lambda_\beta - \lambda_\gamma), \end{aligned} \tag{5.19}$$

where $p_i^{j,S}(x_1,x_2) := x_2^i p_i^j(x_1/x_2)$ is the scaled Jacobi polynomial. Again, we have that $\hat{r}_{ijl}(\lambda_\alpha,\lambda_\beta,\lambda_\gamma,\lambda_\delta)$ with $0 \le i+j+l \le k$ and an arbitrary permutation $(\alpha,\beta,\gamma,\delta)$ of $(0,1,2,3)$ defines a basis for $\mathbb{P}^k(\hat{T},\mathbb{R})$ and that for $0 \le i+l \le k$ the restriction $\hat{r}_{il0}(\lambda_{j+1},\lambda_{j+2},\lambda_{j+3})|_{\hat{F}_j}$ is a basis of $\mathbb{P}^k(\hat{F}_j,\mathbb{R})$ where the indices of the barycentric coordinate functions are now taken modulo 4. By this we define the local basis on the reference tetrahedron by

$$\begin{aligned} \hat{\Psi}_k^F &:= \{\hat{S}_q^j\hat{r}_{il0}(\lambda_{j+1},\lambda_{j+2},\lambda_{j+3}) : j = 0,1,2,3 \text{ and } q = 0,1 \text{ and } 0 \le i+l \le k-1\} \\ \hat{\Psi}_k^T &:= \{\lambda_j\hat{S}_q^j\hat{r}_{ilg}(\lambda_0,\lambda_1,\lambda_2,\lambda_3) : j = 0,1,2,3 \text{ and } q = 0,1 \text{ and } 0 \le i+l+g \le k-1\}. \end{aligned}$$

THEOREM 5.5 The set of functions $\{\hat{\Psi}_k^F \cup \hat{\Psi}_k^T\}$ is a basis for $\Sigma_k(\hat{T})$.

Proof. We start with the two-dimensional case. An elementary calculation shows that the functions $\lambda_i \hat{S}^i$ with $i = 0, 1, 2$ are linearly independent. Let $\alpha_i^j \in \mathbb{R}$ and $\beta_{il}^j \in \mathbb{R}$ be arbitrary coefficients and define $\hat{S}_i^j := \hat{S}^j \hat{r}_{i0}(\lambda_{j+1}, \lambda_{j+2})$ and $\hat{B}_{il}^j := \lambda_j \hat{S}^j \hat{r}_{il}(\lambda_0, \lambda_1, \lambda_2)$. We assume that

$$\sum_{j=0}^{2} \sum_{i=0}^{k-1} \alpha_i^j \hat{S}_i^j + \sum_{j=0}^{2} \sum_{i=0}^{k-1} \sum_{l=i}^{k-1} \beta_{il}^j \hat{B}_{il}^j = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

and show that this induces that all coefficients are equal to zero. This then proves the linear independency of $\{\hat{\Psi}_k^F \cup \hat{\Psi}_k^T\}$. Let $\hat{F}_g$ with $g = 0, 1, 2$ be an arbitrary reference face. Due to (5.17) there holds

$$\hat{t}_g^T \left(\sum_{j=0}^{2} \sum_{i=0}^{k-1} \alpha_i^j \hat{S}_i^j + \sum_{j=0}^{2} \sum_{i=0}^{k-1} \sum_{l=i}^{k-1} \beta_{il}^j \hat{B}_{il}^j \right) \hat{n}_g = \hat{t}_g^T \left(\sum_{i=0}^{k-1} \alpha_i^g \hat{S}_i^g \right) \hat{n}_g = \hat{t}_g^T \left(\sum_{i=0}^{k-1} \alpha_i^g \hat{S}^g \hat{r}_{i0}(\lambda_{g+1}, \lambda_{g+2}) \right) \hat{n}_g = 0.$$

As $\hat{r}_{i0}(\lambda_{g+1}, \lambda_{g+2})$ is a polynomial basis on $\hat{F}_g$ and $\hat{S}^g, \hat{n}_g$ and $\hat{t}_g$ are constant it follows that all coefficients α_i^g have to be zero. As g was arbitrary we conclude $\alpha_i^j = 0$ for $j = 0, 1, 2$ and $0 \le i \le k-1$.

As the functions $\lambda_i \hat{S}^i$ are linearly independent we have for each $g = 0, 1, 2$ (due to the assumption at the beginning)

$$\sum_{i=0}^{k-1} \sum_{l=i}^{k-1} \beta_{il}^g \hat{B}_{il}^g = \sum_{i=0}^{k-1} \sum_{l=i}^{k-1} \beta_{il}^g \hat{r}_{il} \lambda_g \hat{S}^g = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

As $\hat{r}_{il} \lambda_g$ is a basis for $\lambda_g \mathbb{P}^{k-1}(\hat{T})$ and the last equation holds true for all points in $\hat{T}$ we conclude $\beta_{il}^g = 0$ for $0 \le i + l \le k - 1$. As g was arbitrary we conclude that all coefficients are equal to zero. Note that by $\operatorname{tr}(S^i) = 0$, all shape function in $\{\hat{\Psi}_k^F \cup \hat{\Psi}_k^T\}$ are trace free and are further tensor-valued polynomials up to order k. Further, the normal-tangential trace is only a polynomial up to order $k - 1$, thus all shape functions belong to $\Sigma_k(\hat{T})$. Counting the dimensions we have by Theorem 5.4

$$|\hat{\Psi}_k^F| + |\hat{\Psi}_k^T| = 3k + \frac{3k(k+1)}{2} = N_{\Sigma_k},$$

what concludes the proof. In three dimensions we proceed similarly. The linearly independence can be shown with the same steps. Further, with the same arguments all shape functions belong to $\Sigma_k(\hat{T})$. Again by Theorem 5.4 and

$$|\hat{\Psi}_k^F| + |\hat{\Psi}_k^T| = 8\frac{k(k+1)}{2} + 8\frac{k(k+1)(k+2)}{6} = N_{\Sigma_k},$$

we conclude the proof. □

REMARK 5.6 Note how the basis was separated into shape functions associated to faces ($\hat{\Psi}_k^F$) and shape functions associated to the element interior ($\hat{\Psi}_k^T$). The polynomial degrees in each group can be separately chosen to construct a variable-degree global finite element space (e.g., for *hp* adaptivity). For example, the span of the union of $\Psi_{k_1}^F$ and $\Psi_{k_2}^T$ gives an element space that has normal-tangential trace of degree $k_1 - 1$ and inner (bubble) shape functions of degree k_2.

5.6 *Construction of a global basis*

Using the local basis on the reference triangle $\hat{T}$ we can now simply define a global basis for the stress space Σ_h. This is done in the usual way. Using the mapping $\mathscr{M}$ and a basis function $\hat{S} \in \{\hat{\Psi}_k^T \cup \hat{\Psi}_k^F\}$, we define the restriction of a global shape function S (with support on a patch) on an arbitrary physical element $T \in \mathscr{T}_h$ by

$$S := \mathscr{M}(\hat{S}).$$

Next we identify all topological entities, vertices and faces, of the physical element T with the corresponding entities of the global mesh. This identification is needed as faces and vertices coincide for adjacent physical elements. Note that the global orientation of the faces (and edges) plays an important role in order to assure (normal-tangential) continuity. This is a well-known difficulty: see Zaglmayr (2006) for a detailed discussion regarding this topic. By this we construct global basis functions that are, restricted on a physical element $T \in \mathscr{T}_h$, always a mapped basis function of the basis defined on the reference element $\hat{T}$.

Further, note that due to Lemma 5.3 the resulting basis functions are normal-tangential continuous, thus $[\![S_{nt}]\!] = 0$. To see this let ϕ_1 be the mapping of an arbitrary element T_1 and let ϕ_2 be the mapping of an element T_2 such that $F = T_1 \cap T_2$. There exists a reference face $\hat{F} \subset \partial\hat{T}$ such that $F = \phi_1(\hat{F}) = \phi_1(\hat{F})$ (in the sense of a set) and $\phi_1|_{\hat{F}} = \phi_2|_{\hat{F}}$ (in the sense of equivalent functions). By this, and the same ideas for a reference edge $\hat{E}$ in the three-dimensional case, the constant c in Lemma 5.3 is the same for both mappings. In two dimensions we have the identity $S_{nt} = (t^{\mathrm{T}}Sn)t$, thus Lemma 5.3 implies normal-tangential continuity of S because S was a mapped basis functions of the reference element. In three dimensions S_{nt} is a tangent vector in F. Each tangent vector can be represented as a linear combination of two arbitrary edge tangent vectors $t_i \subset \partial F$. By Lemma 5.3 we deduce that the scalar values ${t_i}^{\mathrm{T}}Sn$ are preserved, thus again we have normal-tangential continuity. Taking all functions in $\{\hat{\Psi}_k^T \cup \hat{\Psi}_k^F\}$ and mapping them to each element separately results in a basis for Σ_h.

5.7 *An interpolation operator for the stress space*

We finish this section by introducing an interpolation operator for the stress space and showing an approximation result. Using the global dofs of Σ_h a canonical interpolation operator I_{Σ_h} can be defined as usual. On each $T \in \mathscr{T}_h$ the interpolant $(I_{\Sigma_h}\sigma)|_T$ coincides with the canonical local interpolant $I_T(\sigma|_T)$ defined, as usual, using the local dofs in $\Phi(T)$, by

$$\phi(\sigma - I_T\sigma) = 0 \quad \text{for all } \phi \in \Phi(T). \tag{5.20}$$

Recalling the map $\mathscr{M}$ from (5.9), note that $\mathscr{M}^{-1}(\sigma) = \det(F_T^F)F_F^{\mathrm{T}}(\sigma \circ \phi_T)F_T^{-\mathrm{T}}$.

LEMMA 5.7 For any $\sigma \in H^1(T, \mathbb{R}^{d\times d})$,

$$\mathscr{M}^{-1}(I_T\sigma) = I_{\hat{T}}(\mathscr{M}^{-1}(\sigma)).$$

Proof. Since both the left- and right-hand sides are in $\Sigma_k(\hat{T})$, it suffices to prove that

$$\hat{\phi}(\mathscr{M}^{-1}(I_T\sigma) - I_{\hat{T}}(\mathscr{M}^{-1}\sigma)) = 0 \quad \text{for all } \hat{\phi} \in \Phi(\hat{T}). \tag{5.21}$$

To see that (5.21) holds for the interior dofs on $\hat{T}$ as defined in (5.11), noting that $F_{\hat{T}}$ is the identity, we have for all $\hat{\eta} \in \mathscr{B}_k(\hat{T})$,

$$\begin{aligned}\int_{\hat{T}} \left[\mathscr{M}^{-1}(I_T\sigma) - I_{\hat{T}}(\mathscr{M}^{-1}\sigma)\right] : F_{\hat{T}}\hat{\eta}F_{\hat{T}}^{-1}\ \mathrm{d}\hat{x} &= \int_{\hat{T}} \left[\mathscr{M}^{-1}(I_T\sigma) - \mathscr{M}^{-1}\sigma\right] : \hat{\eta}\ \mathrm{d}\hat{x} \\ &= \int_T (I_T\sigma - \sigma) : F_T\hat{\eta}F_T^{-1}\ \mathrm{d}x = 0\end{aligned}$$

due to the equality of interior dofs on T in (5.20).

Next, consider the facet dofs. We only consider the $d = 3$ case (as the other case is simpler). On an arbitrary facet $\hat{F} \in \mathscr{F}_{\hat{T}}$ choose two arbitrary edges $\hat{E}_1, \hat{E}_2$ with unit tangential vectors $\hat{t}_1$ and $\hat{t}_2$. Using a dual tangential basis $\hat{s}_1$ and $\hat{s}_2$ such $\hat{s}_i \cdot \hat{t}_i = \delta_{ij}$, we expand

$$[\mathscr{M}^{-1}(I_T\sigma - \sigma)]_{nt} = [\hat{t}_1^{\mathrm{T}}\mathscr{M}^{-1}(I_T\sigma - \sigma)\hat{n}]\hat{s}_1 + [\hat{t}_2^{\mathrm{T}}\mathscr{M}^{-1}(I_T\sigma - \sigma)\hat{n}]\hat{s}_2.$$

Next we choose arbitrary $\hat{r}_1, \hat{r}_2 \in \mathbb{P}^{k-1}(\hat{F}, \mathbb{R})$ and define

$$\hat{r} := \frac{\hat{r}_1}{\det(F_{E_1})}\hat{t}_1 + \frac{\hat{r}_2}{\det(F_{E_2})}\hat{t}_2.$$

Let $r_i = \hat{r}_i \circ \phi_T$. Using a biorthogonal basis s_1, s_2 with respect to unit tangents t_1 and t_2 of mapped edges E_1 and E_2, we have $r := r_1t_1 + r_2t_2$. Using Lemma 5.3 we deduce

$$[\mathscr{M}^{-1}(I_T\sigma - \sigma)]_{nt} = \det(F_T^F)\det(F_{E_1})[t_1^{\mathrm{T}}(I_T\sigma - \sigma)n]\hat{s}_1 + \det(F_T^F)\det(F_{E_2})[t_2^{\mathrm{T}}(I_T\sigma - \sigma)n]\hat{s}_2,$$

so

$$\begin{aligned}\int_{\hat{F}} [\mathscr{M}^{-1}(I_T\sigma - \sigma)]_{nt} \cdot \hat{r}\ \mathrm{d}\hat{x} &= \int_F [t_1^{\mathrm{T}}(I_T\sigma - \sigma)n]r_1s_1 \cdot t_1\ \mathrm{d}x + \int_F [t_2^{\mathrm{T}}(I_T\sigma - \sigma)n]r_2s_2 \cdot t_2\ \mathrm{d}x \\ &= \int_F [(t_1^{\mathrm{T}}(I_T\sigma - \sigma)n)s_1 + (t_2^{\mathrm{T}}(I_T\sigma - \sigma)n)s_2] \cdot [r_1t_1 + r_2t_2]\ \mathrm{d}x \\ &= \int_F (I_T\sigma - \sigma)_{nt} \cdot r\ \mathrm{d}x = 0.\end{aligned}$$

where the last equality is due to the equality of the facet dofs in (5.20). □

THEOREM 5.8 (Interpolation operator for Σ_h). For any $m \geq 1$ and any $\sigma \in \{\tau \in H^m(\mathscr{T}_h, \mathbb{R}^{d\times d}) : [\![\tau_{nt}]\!] = 0\}$, the interpolant $I_{\Sigma_h}\sigma$ is well defined and there is a mesh-independent constant C such that

$$||\sigma - I_{\Sigma_h}\sigma||_{L^2(\Omega)} + \sqrt{\sum_{F\in\mathscr{F}_h} h||(\sigma - I_{\Sigma_h}\sigma)_{nt}||_F^2} \leq Ch^s||\sigma||_{H^s(\mathscr{T}_h)} \tag{5.22}$$

for all $s \leq \min(k, m)$.

Proof. Let $\hat{\sigma} = \mathscr{M}^{-1}(\sigma|_T)$. By Lemma 5.7 $\mathscr{M}^{-1}(\sigma - I_T\sigma) = \hat{\sigma} - I_{\hat{T}}\hat{\sigma}$. By the unisolvency of the reference element dofs (Theorem 5.4)

$$\hat{\sigma} - I_{\hat{T}}\hat{\sigma} = 0 \quad \text{for all } \hat{\sigma} \in \mathbb{P}^{k-1}(\hat{T}, \mathbb{R}^{d\times d}).$$

Now a standard argument using the Bramble–Hilbert lemma, the continuity of $I_{\hat{T}} : H^s(\hat{T}, \mathbb{R}^{d\times d}) \to L^2(\hat{T}, \mathbb{R}^{d\times d})$ and scaling arguments, finish the proof. □

6. *A priori* error analysis

In this section we show discrete inf-sup stability of the MCS method, optimal error estimates (Theorem 6.9) and pressure robustness (Theorem 6.10). The error analysis is in the following norms.

$$\begin{aligned}
\|\tau_h\|_{\Sigma_h}^2 &:= ||\tau_h||_{L^2(\Omega)}^2 = ||\mathrm{dev}\,(\tau_h)||_{L^2(\Omega)}^2, && \tau_h \in \Sigma_h,\\
||v_h||_{V_h}^2 &:= ||v_h||_{1,h}^2 := \sum_{T\in\mathscr{T}_h} ||\nabla v_h||_T^2 + \sum_{F\in\mathscr{F}_h} \frac{1}{h}\|[\![(v_h)_t]\!]\|_F^2, && v_h \in V_h,\\
||q_h||_{Q_h}^2 &:= ||q_h||_{L^2(\Omega)}^2, && q_h \in Q_h.
\end{aligned}$$

Comparing with (appropriate) norms of the infinite dimensional spaces V and Σ, these norms might seem unnatural. However, we choose these norms in order to obtain velocity error estimates in an H^1-like norm comparable to the standard velocity-pressure formulation. Since our discrete spaces do not admit H^1-conformity, our $||\cdot||_{V_h}$-norm contains a term that penalizes the tangential discontinuities (as in the analysis of discontinuous Galerkin methods). The L^2-like norm on the Σ_h is also related to an H^1-like norm of the velocity, since we expect σ_h to be an approximation of $\nu\nabla u$. From this section on, for convenience, we shall assume that ν is constant.

6.1 *Norm equivalences*

We use $A \sim B$ to indicate that there are constants $c, C > 0$ *independent of the mesh size h and the viscosity ν* such that $cA \leq B \leq CA$. We also use $A \lesssim B$ when there is a $C > 0$ independent of h and ν such that $A \leq CB$ (and $\gtrsim$ is defined similarly). Due to quasiuniformity the following estimates follow by standard scaling arguments: for any $\hat{\tau} \in \Sigma_k(\hat{T})$, letting $\tau = \mathscr{M}(\hat{\tau})$,

$$h^d\|\tau_h\|_T^2 \sim \|\hat{\tau}_h\|_{0,\hat{T}}^2. \tag{6.1}$$

On any $F \in \mathscr{F}_T$ Lemma 5.3, together with a scaling argument, yields

$$h^{d+1}\left\|t^{\mathrm{T}}\tau_h n\right\|_F^2 \sim \left\|\hat{t}^{\mathrm{T}}\hat{\tau}_h\hat{n}\right\|_{0,\hat{F}}^2. \tag{6.2}$$

LEMMA 6.1 For all $\tau_h \in \Sigma_h$,

$$\left\|\tau_h\right\|_{\Sigma_h}^2 \sim \sum_{T\in\mathscr{T}_h} ||\mathrm{dev}\,(\tau_h)||_T^2 + \sum_{F\in\mathscr{F}_h} h\left\|(\tau_h)_{nt}\right\|_F^2.$$

Proof. By finite dimensionality, for any face $\hat{F} \in \mathscr{F}_{\hat{T}}$,

$$h\|\hat{t}^{\mathrm{T}}\hat{\tau}_h\hat{n}\|_{0,\hat{F}}^2 \lesssim \|\hat{\tau}_h\|_{0,\hat{T}}^2, \qquad \text{for all } \hat{\tau}_h \in \Sigma_k(\hat{T}).$$

Due to (6.2) and (6.1) this yields

$$\sum_{F\in\mathscr{F}_h} h\|[\![(\tau_h)_{nt}]\!]\|_F^2 \lesssim \sum_{T\in\mathscr{T}_h} \|\tau_h\|_T^2, \qquad \text{for all } \tau_h \in \Sigma_k(T).$$

This proves one side of the stated equivalence. The other side is obvious. □

On each facet $F \in \mathscr{F}_h$ with normal vector n_F let Π_F^0 denote the L^2 projection onto the space of constant tangential vectors in $n_F^{\perp}$, i.e., for any vector function $v \in L^2(F, n_F^{\perp})$, the projection $\Pi_F^0 v \in n_F^{\perp}$ satisfies $(\Pi_F^0 v, t)_F = (v, t)_F$ for all $t \in n_F^{\perp}$.

LEMMA 6.2 For all $v_h \in V_h$,

$$||v_h||_{V_h}^2 \sim \sum_{T\in\mathscr{T}_h} ||\nabla v_h||_T^2 + \sum_{F\in\mathscr{F}_h} \frac{1}{h}\left\|\Pi_F^0[\![(v_h)_t]\!]\right\|_F^2.$$

Proof. One side of the equivalence is obvious from the continuity of Π_F^0. For the other direction

$$||v_h||_{V_h}^2 \le \sum_{T\in\mathscr{T}_h} ||\nabla v_h||_T^2 + \sum_{F\in\mathscr{F}_h} \frac{2}{h}\left\|\Pi_F^0[\![(v_h)_t]\!]\right\|_F^2 + \frac{2}{h}\left\|[\![(v_h)_t]\!] - \Pi_F^0[\![(v_h)_t]\!]\right\|_F^2. \tag{6.3}$$

Now, on each facet $F \in \mathscr{F}_T$, we use the standard estimate $\left\|(v_h)_t - \Pi_F^0(v_h)_t\right\|_F \lesssim h^{1/2}\|\nabla v_h\|_T$ to complete the proof. □

6.2 *Stability analysis*

LEMMA 6.3 (Continuity of a, b_1 and b_2). The bilinear forms a, b_1 and b_2 are continuous:

$$\begin{aligned} a(\sigma_h,\tau_h) &\lesssim \frac{1}{\sqrt{\nu}}\left\|\sigma_h\right\|_{\Sigma_h}\frac{1}{\sqrt{\nu}}\left\|\tau_h\right\|_{\Sigma_h} && \text{for all } \sigma_h,\tau_h\in\Sigma_h\\ b_1(v_h,p_h) &\lesssim ||v_h||_{V_h}||p_h||_{Q_h} && \text{for all } v_h\in V_h, p_h\in Q_h\\ b_2(\sigma_h,v_h) &\lesssim \left\|\sigma_h\right\|_{\Sigma_h}||v_h||_{V_h} && \text{for all } \sigma_h\in\Sigma_h, v_h\in V_h. \end{aligned}$$

Proof. The continuity for the bilinear forms a and b_1 follows from the Cauchy–Schwarz inequality, we only consider b_2, which by (4.7) can be written as

$$b_2(\sigma_h,v_h) = -\sum_{T\in\mathscr{T}_h}\int_T \sigma_h : \nabla v_h \,\mathrm{d}x + \sum_{F\in\mathscr{F}_h}\int_F (\sigma_h)_{nt}\cdot [\![(v_h)_t]\!]\,\mathrm{d}s.$$

Since $(\sigma_h)_{nt} = (\operatorname{dev}(\sigma_h))_{nt}$ we conclude the proof by Cauchy–Schwarz inequality and Lemma 6.1. ☐

LEMMA 6.4 (Coercivity of a on the kernel). Let $K_h := \{(\tau_h,q_h)\in\Sigma_h\times Q_h : b_1(v_h,q_h)+b_2(\sigma_h,v_h)=0$ for all $v_h\in V_h\}$. For all $(\sigma_h,p_h)\in K_h$,

$$\frac{1}{\nu}(\|\sigma_h\|_{\Sigma_h}+\|p_h\|_{Q_h})^2 \lesssim a(\sigma_h,\sigma_h).$$

Proof. Let $(\sigma_h,p_h)\in K_h$ be arbitrary. As $\nu^{-1}\left\|\sigma_h\right\|^2_{\Sigma_h} = a(\sigma_h,\sigma_h)$ it is sufficient to bound only the norm of p_h. It is well known—see e.g., Boffi *et al.* (2013)—that for any $p_h\in Q_h$

$$\exists v_h\in V_h : \qquad \operatorname{div}(v_h)=p_h, \quad ||v_h||_{V_h}\lesssim ||p_h||_{Q_h}. \tag{6.4}$$

With this v_h,

$$\begin{aligned} ||p_h||^2_{Q_h} &= \sum_{T\in\mathscr{T}_h}\int_T \operatorname{div}(v_h)p_h\,\mathrm{d}x = b_1(v_h,p_h)\\ &= -b_2(\sigma_h,v_h) && \text{as } (\sigma_h,p_h)\in K_h,\\ &= \sum_{T\in\mathscr{T}_h}\int_T \sigma_h:\nabla v_h\,\mathrm{d}x - \sum_{F\in\mathscr{F}_h}\int_F(\sigma_h)_{nt}\cdot[\![(v_h)_t]\!]\,\mathrm{d}s && \text{by (4.7)},\\ &\le ||\operatorname{dev}(\sigma_h)||_{L^2(\Omega)}||v_h||_{V_h} && \text{using Lemma 6.1},\\ &\lesssim \|\sigma_h\|_{\Sigma_h}||p_h||_{Q_h} && \text{by (6.4)}. \end{aligned}$$

☐

Next, we proceed to verify the discrete LBB condition (in Theorem 6.5 below). Define

$$V_h^0 := \{w_h \in V_h : \operatorname{div}(w_h) = 0\},$$

$$\|v_h\|_{1,\mathrm{dev},h} := \left(\sum_{K \in \mathscr{T}_h} \|\operatorname{dev}(\nabla v_h)\|_T^2 + \sum_{F \in \mathscr{F}_h} \frac{1}{h} \left\| \Pi_F^0 [\![(v_h)_t]\!] \right\|_F^2 \right)^{1/2}.$$

As $\|\nabla v_h\|_T^2 \sim \|\operatorname{dev}(\nabla v_h)\|_T^2 + \|\operatorname{div}(v_h)\|_T^2$ on any $T \in \mathscr{T}_h$ and for any $v_h \in V_h$, we have by Lemma 6.2

$$\left\| v_h \right\|_{1,\mathrm{dev},h} \sim \|v_h\|_{V_h} \quad \text{for all } v_h \in V_h^0. \tag{6.5}$$

The first step towards proving the LBB condition is the construction of a specific stress function τ_h, which only depends on $\operatorname{dev}(\nabla v_h)$ for any $v_h \in V_h^0$. Using this τ_h we prove an LBB condition for b_2 on V_h^0, which is the content of the next lemma. As $\tau_h \in \Sigma_h$ has a zero trace, we cannot in general control the divergence of a general $v_h \in V_h$ solely using such a τ_h. Therefore, to complete the proof of the full inf-sup condition (in the proof of Theorem 6.6 below), we utilize an appropriate pressure test function as well.

Lemma 6.5 For any nonzero $v_h \in V_h$ there exists a nonzero $\tau_h \in \Sigma_h$, satisfying $b_2(\tau_h, v_h) \gtrsim \left\| v_h \right\|_{1,\mathrm{dev},h}^2$ and $\|\tau_h\|_{\Sigma_h} \lesssim \left\| v_h \right\|_{1,\mathrm{dev},h}$, so by (6.5),

$$\|v_h\|_{V_h} \lesssim \sup_{\tau_h \in \Sigma_h} \frac{b_2(\tau_h, v_h)}{\|\tau_h\|_{\Sigma_h}} \quad \text{for all } v_h \in V_h^0.$$

Proof. Since the ideas are the same for $d = 2$ and 3, for ease of exposition, we give the details of the proof only in the $d = 2$ case. Because of the decomposition of the dofs into face and interior dofs (see (5.10) and (5.11)) we may decompose $\Sigma_h = \Sigma_h^0 \oplus \Sigma_h^1$, where $\Sigma_h^0 = \oplus_{K \in \mathscr{T}_h} \mathscr{B}_k(T)$ and Σ_h^1 is the span of facet shape functions (see also Remark 5.6). In particular, Σ_h^1 contains the lowest-order shape function S^F with the property that $S^F_{nt} \in n_F^{\perp}$ and $||S^F_{nt}||_2 = 1$ on the facet F and equals $(0, 0)$ on all other facets in $\mathscr{F}_h$. (S^F can be explicitly written down by mapping (5.15) or by appropriately scaling (5.2).) Given any $v_h \in V_h^0$, define

$$\tau_h^0 := \sum_{T \in \mathscr{T}_h} \sum_{F \in \mathscr{F}_T} -(S^F : \operatorname{dev}(\nabla v_h)) \lambda_T^F S^F, \qquad \tau_h^1 := \sum_{F \in \mathscr{F}_h} \frac{1}{\sqrt{h}} (\Pi_F^0 [\![(v_h)_t]\!]) S^F, \tag{6.6}$$

where λ_T^F is the barycentric coordinate of T that vanishes on F (thus $\lambda_T^F S^F$ is a linear inner *nt*-bubble). Below, we shall construct a linear combination of these functions to obtain the τ_h stated in the lemma.

By (6.1) and (6.2) a scaling argument (like in Lemma 6.1) shows that there is a mesh-independent C_1 such that

$$\left\| \tau_h^1 \right\|_{\Sigma_h}^2 \le C_1 \sum_{F \in \mathscr{F}_h} \frac{1}{h} \left\| \Pi_F^0 [\![(v_h)_t]\!] \right\|_F^2. \tag{6.7}$$

A similar scaling argument also shows that

$$\left\|\tau_h^0\right\|_{\Sigma_h}^2 \lesssim \sum_{T\in\mathscr{T}_h} \|\mathrm{dev}\,(\nabla v_h)\|_T^2. \tag{6.8}$$

By construction $(\tau_h^0)_{nt}$ vanishes and

$$b_2(\tau_h^0, v_h) = \sum_{T\in\mathscr{T}_h} -\int_T \tau_h^0 : \nabla v_h \,\,\mathrm{d}x = \sum_{T\in\mathscr{T}_h}\int_T \sum_{F\in\mathscr{F}_T} (S^F : \mathrm{dev}\,(\nabla v_h))^2 \lambda_T^F.$$

Since the functions S^F form a basis for $\mathbb{D}$ by Lemma 5.1, a scaling argument shows that

$$b_2(\tau_h^0, v_h) \gtrsim \sum_{T\in\mathscr{T}_h} \|\mathrm{dev}\,(\nabla v_h)\|_T^2. \tag{6.9}$$

Next, set $\tau_h = \gamma_0\tau_h^0 + \gamma_1\tau_h^1$ where γ_0 and γ_1 are positive constants to be chosen. Then

$$\begin{aligned}
b_2(\tau_h, v_h) &\gtrsim \gamma_0 \sum_{T\in\mathscr{T}_h} \|\mathrm{dev}\,(\nabla v_h)\|_T^2 + \gamma_1 b_2(\tau_h^1, v_h) && \text{by (6.9)}\\
&= \gamma_0 \sum_{T\in\mathscr{T}_h} \|\mathrm{dev}\,(\nabla v_h)\|_T^2 + \gamma_1 \left(\sum_{T\in\mathscr{T}_h} -\int_T \tau_h^1 : \nabla v_h \,\,\mathrm{d}x + \sum_{F\in\mathscr{F}_h}\int_F (\tau_h^1)_{nt}\cdot [\![(v_h)_t]\!] \,\,\mathrm{d}s\right)\\
&= \gamma_0 \sum_{T\in\mathscr{T}_h} \|\mathrm{dev}\,(\nabla v_h)\|_T^2 - \gamma_1 \sum_{T\in\mathscr{T}_h}\int_T \tau_h^1 : \mathrm{dev}\,(\nabla v_h)\,\mathrm{d}x + \gamma_1 \sum_{F\in\mathscr{F}_h} \frac{1}{h}\left\|\varPi_F^0 [\![(v_h)_t]\!]\right\|_F^2 && \text{by (6.6).}
\end{aligned}$$

Applying the Cauchy–Schwarz inequality and also Young's inequality with $\delta > 0$ we further have

$$\begin{aligned}
b_2(\tau_h, v_h) &\gtrsim \gamma_0 \sum_{T\in\mathscr{T}_h} \|\mathrm{dev}\,(\nabla v_h)\|_T^2 - \gamma_1 \|\tau_h^1\|_{\Sigma_h} \sqrt{\sum_{T\in\mathscr{T}_h} \|\mathrm{dev}\,(\nabla v_h)\|_T^2} + \gamma_1 \sum_{F\in\mathscr{F}_h} \frac{1}{h}\left\|\varPi_F^0 [\![(v_h)_t]\!]\right\|_F^2\\
&\gtrsim \left(\gamma_0 - \frac{\gamma_1\delta}{2}\right) \sum_{T\in\mathscr{T}_h} \|\mathrm{dev}\,(\nabla v_h)\|_T^2 + \left(1 - \frac{C_1}{2\delta}\right)\frac{\gamma_1}{h} \sum_{F\in\mathscr{F}_h} \left\|\varPi_F^0 [\![(v_h)_t]\!]\right\|_F^2,
\end{aligned}$$

where in the last step we also used (6.7). Choosing $\delta = C_1$, $\gamma_1 = 1/\delta = 1/C_1$ and $\gamma_0 = 1$,

$$b_2(\tau_h, v_h) \gtrsim \sum_{T\in\mathscr{T}_h} \|\,\mathrm{dev}\,(\nabla v_h)\|_T^2 + \sum_{F\in\mathscr{F}_h} \frac{1}{h}\left\|\varPi_F^0 [\![(v_h)_t]\!]\right\|_F^2. \tag{6.10a}$$

Let us also note that (6.7) and (6.8) yield

$$\|\tau_h\|_{\Sigma_h} \lesssim \sum_{T\in\mathscr{T}_h} \|\mathrm{dev}\,(\nabla v_h)\|_T^2 + \sum_{F\in\mathscr{F}_h} \frac{1}{h}\left\|\Pi_F^0[\![(v_h)_t]\!]\right\|_F^2. \tag{6.10b}$$

The estimates (6.10) and the norm equivalences of (6.5) and Lemma 6.2 complete the proof. □

THEOREM 6.6 (Discrete LBB condition). For all $v_h \in V_h$,

$$\sup_{(\tau_h,q_h)\in\Sigma_h\times Q_h} \frac{b_1(v_h,q_h)+b_2(\tau_h,v_h)}{\left\|\tau_h\right\|_{\Sigma_h}+||q_h||_{Q_h}} \gtrsim ||v_h||_{V_h}. \tag{6.11}$$

Proof. By Lemma 6.5, for any $v_h \in V_h$, there is a $\tau_h \in \Sigma_h$, satisfying $b_2(\tau_h, v_h) \gtrsim \left\|v_h\right\|_{1,\mathrm{dev},h}^2$ and $\|\tau_h\|_{\Sigma_h} \lesssim \left\|v_h\right\|_{1,\mathrm{dev},h}$. Next we choose the pressure variable $q_h = \mathrm{div}(v_h)$, which is possible due to the specific choice of V_h and Q_h, so that $b_1(v_h, q_h) = \|\mathrm{div}(v_h)\|_{Q_h}^2$. With these choices of τ_h and q_h, we have

$$\frac{b_1(v_h,q_h)+b_2(\tau_h,v_h)}{\left\|\tau_h\right\|_{\Sigma_h}+||q_h||_{Q_h}} \ge \frac{\left\|v_h\right\|_{1,\mathrm{dev},h}^2+\|\mathrm{div}(v_h)\|_{Q_h}^2}{\left\|\tau_h\right\|_{\Sigma_h}+||q_h||_{Q_h}} \gtrsim \|v_h\|_{V_h}.$$

□

REMARK 6.7 (Residual stabilization alternative). A crucial ingredient in the proof of the LBB condition was the choice made in (6.6). The choice of τ_h^0 in terms of $(S^F : \mathrm{dev}\,(\nabla v_h))\lambda_T^F S^F$ was admissible as $\mathrm{dev}\,(\nabla u_h)$ is a polynomial of degree $k-1$, and Σ_h contains the element-wise bubbles of degree k in $\mathscr{B}_k(T)$. This choice would not be admissible if we had used bubbles in $\mathscr{B}_{k-1}(T)$ instead of $\mathscr{B}_k(T)$. Therefore, if we replace the stress space by the lower-degree space

$$\widetilde{\Sigma}_h := \{\tau_h \in \mathbb{P}^{k-1}(\mathscr{T}_h, \mathbb{R}^{d\times d}) :\ \mathrm{tr}(\tau_h) = 0,\ [\![(\tau_h)_{nt}]\!] = 0\},$$

the above proof can no longer be used to conclude stability of the resulting method. Yet, it is possible to get a good method (with optimal error convergence results) using $\widetilde{\Sigma}_h$ by a residual-based stabilization term. Define $c : \left[L^2(\Omega,\mathbb{R}^{d\times d})\times V\right] \times \left[L^2(\Omega,\mathbb{R}^{d\times d})\times V\right] \to \mathbb{R}$ by

$$c((\sigma,u),(\tau,v)) := -\sum_{T\in\mathscr{T}_h} \frac{\nu}{2}\int_T \left(\frac{1}{\nu}\sigma - \nabla u\right) : \left(\frac{1}{\nu}\tau - \nabla v\right)\,\mathrm{d}x.$$

When this form is added to the system (MCS) and Σ_h is replaced by $\widetilde{\Sigma}_h$, it is possible to prove stability.

THEOREM 6.8 (Consistency). The mass conserving mixed stress formulation (MCS) is consistent in the following sense. If the exact solution of the mixed Stokes problem (3.2) is such that $u \in H^1(\Omega,\mathbb{R}^d)$, $\sigma \in H^1(\Omega,\mathbb{R}^{d\times d})$ and $p \in L_0^2(\Omega,\mathbb{R})$, then

$$a(\sigma,\tau_h) + b_2(\tau_h,u) + b_2(\sigma,v_h) + b_1(v_h,p) + b_1(u,q_h) = (-f,v_h)_\Omega$$

for all $v_h \in V_h, q_h \in Q_h$ and $\tau_h \in \Sigma_h$.

Proof. As the exact solutions σ and u are continuous, we have $[\![\sigma_{nn}]\!] = 0$ and $[\![u_t]\!] = 0$ on all faces $F \in \mathscr{F}_h$, and thus using representations (4.6) and (4.7), we have

$$b_2(\sigma, v_h) = \sum_{T\in\mathscr{T}_h} \int_T \operatorname{div}(\sigma) \cdot v_h \, \mathrm{d}x - \sum_{F\in\mathscr{F}_h} \int_F [\![\sigma_{nn}]\!](v_h)_n \, \mathrm{d}s = \sum_{T\in\mathscr{T}_h} \int_T \operatorname{div}(\sigma) \cdot v_h \, \mathrm{d}x$$

and

$$b_2(\tau_h, u) = -\sum_{T\in\mathscr{T}_h} \int_T \tau_h : \nabla u \, \mathrm{d}x + \sum_{F\in\mathscr{F}_h} \int_F (\tau_h)_{nt} \cdot [\![u_t]\!] \, \mathrm{d}s = -\sum_{T\in\mathscr{T}_h} \int_T \tau_h : \nabla u \, \mathrm{d}x.$$

Using $\operatorname{div}(u) = 0$ we further get that $b_1(u, q_h) = 0$, so all together we have

$$\begin{aligned} &a(\sigma, \tau_h) + b_2(\tau_h, u) + b_2(\sigma, v_h) + b_1(v_h, p) + b_1(u, q_h) \\ &\quad = \int_\Omega \frac{1}{\nu} \operatorname{dev}(\sigma) : \operatorname{dev}(\tau_h) \, \mathrm{d}x - \sum_{T\in\mathscr{T}_h} \int_T \tau_h : \nabla u \, \mathrm{d}x + \sum_{T\in\mathscr{T}_h} \int_T \operatorname{div}(\sigma) \cdot v_h \, \mathrm{d}x + \int_\Omega \operatorname{div}(v_h) p \, \mathrm{d}x. \end{aligned}$$

For the exact solution we have $\operatorname{dev}(\sigma) = \nu \nabla u$. Further, as $\operatorname{div}(u) = 0$, a simple calculation shows that $\tau_h : \nabla u = \tau_h : \operatorname{dev}(\nabla u) = \operatorname{dev}(\tau_h) : \nabla u$. Using integrating by parts for the last integral, we conclude

$$\begin{aligned} &a(\sigma, \tau_h) + b_2(\tau_h, u) + b_2(\sigma, v_h) + b_1(v_h, p) + b_1(u, q_h) \\ &\quad = \int_\Omega \nabla u : \operatorname{dev}(\tau_h) \, \mathrm{d}x - \sum_{T\in\mathscr{T}_h} \int_T \operatorname{dev}(\tau_h) : \nabla u \, \mathrm{d}x + \sum_{T\in\mathscr{T}_h} \int_T \operatorname{div}(\sigma) \cdot v_h \, \mathrm{d}x + \int_\Omega \operatorname{div}(v_h) p \, \mathrm{d}x \\ &\quad = \int_\Omega \operatorname{div}(\sigma) \cdot v_h \, \mathrm{d}x + \int_\Omega \operatorname{div}(v_h) p \, \mathrm{d}x = \int_\Omega \left[\operatorname{div}(\sigma) - \nabla p\right] \cdot v_h \, \mathrm{d}x = \int_\Omega -f v_h \, \mathrm{d}x. \end{aligned}$$

□

6.3 *Error estimates*

THEOREM 6.9 (Optimal convergence rates). Let $u \in H^1(\Omega, \mathbb{R}^d) \cap H^m(\mathscr{T}_h, \mathbb{R}^d)$, $\sigma \in H^1(\Omega, \mathbb{R}^{d\times d}) \cap H^{m-1}(\mathscr{T}_h, \mathbb{R}^{d\times d})$ and $p \in L^2_0(\Omega, \mathbb{R}) \cap H^{m-1}(\mathscr{T}_h, \mathbb{R})$ be the exact solution of the mixed Stokes problem (3.2). Further, let σ_h, u_h and p_h be the solution of the mass conserving mixed stress formulation (MCS). For $s = \min(m-1, k)$ there holds

$$||u - u_h||_{V_h} + \frac{1}{\nu} \left\|\sigma - \sigma_h\right\|_{\Sigma_h} + \frac{1}{\nu}||p - p_h||_{Q_h} \lesssim h^s \left(||u||_{H^{s+1}(\mathscr{T}_h)} + \frac{1}{\nu}||\sigma||_{H^s(\mathscr{T}_h)} + \frac{1}{\nu}||p||_{H^s(\mathscr{T}_h)} \right).$$

Proof. The proof is based on the discrete stability established above, which we shall use after bounding the error by triangle inequality into interpolation error and a discrete measure of error, as follows:

$$\begin{aligned} ||u-u_h||_{V_h} + \frac{1}{\nu}\left\|\sigma-\sigma_h\right\|_{\Sigma_h} + \frac{1}{\nu}||p-p_h||_{Q_h} \\ \lesssim ||u-I_{V_h}u||_{V_h} + \frac{1}{\nu}\left\|\sigma-I_{\Sigma_h}\sigma\right\|_{\Sigma_h} + \frac{1}{\nu}||p-I_{Q_h}p||_{Q_h} \\ + \ ||I_{V_h}u-u_h||_{V_h} + \frac{1}{\nu}\left\|I_{\Sigma_h}\sigma-\sigma_h\right\|_{\Sigma_h} + \frac{1}{\nu}||I_{Q_h}p-p_h||_{Q_h}. \end{aligned} \tag{6.12}$$

Here I_{Σ_h} is the interpolation operator studied in Theorem 5.8, I_{V_h} is the standard $H(\mathrm{div})$-conforming interpolant—see Brezzi *et al.* (1985); Raviart & Thomas (1977)—and I_{Q_h} is the L^2 projection into Q_h. Note that for $s=\min(m-1,k)$ we have the approximation results

$$||u-I_{V_h}u||_{V_h} \lesssim h^s||u||_{H^{s+1}(\mathscr{T}_h)} \quad \text{and} \quad ||p-I_{Q_h}p||_{Q_h} \lesssim h^s||p||_{H^s(\mathscr{T}_h)}. \tag{6.13}$$

When this is combined with (5.22) of Theorem 5.8, the first three terms on the right-hand side (6.12) can be bounded as needed.

To bound the remaining terms of (6.12), we first define the following norm on the product space $V_h \times \Sigma_h \times Q_h$ given by

$$||(u_h,\sigma_h,p_h)||_* := \sqrt{\nu}||u_h||_{V_h} + \frac{1}{\sqrt{\nu}}\left(\left\|\sigma_h\right\|_{\Sigma_h} + ||p_h||_{Q_h}\right).$$

Using the Brezzi theorem—see for example in Boffi *et al.* (2013)—the LBB condition of the bilinear forms b_1 and b_2 (Theorem 6.6), the coercivity of a (Lemma 6.4) and the continuity (Lemma 6.3) imply inf-sup stability of the bilinear form

$$B(u_h,\sigma_h,p_h;v_h,\tau_h,q_h) := a(\sigma_h,\tau_h) + b_1(u_h,q_h) + b_1(v_h,p_h) + b_2(\sigma_h,v_h) + b_2(\tau_h,u_h),$$

with respect to the product space norm $||(\cdot,\cdot,\cdot)||_*$, i.e.,

$$\begin{aligned} ||(I_{V_h}u-u_h, I_{\Sigma_h}\sigma-\sigma_h, I_{Q_h}p-p_h)||_* &\leq \sup_{(v_h,\tau_h,q_h)\in V_h\times\Sigma_h\times Q_h} \frac{B(I_{V_h}u-u_h, I_{\Sigma_h}\sigma-\sigma_h, I_{Q_h}p-p_h; v_h,\tau_h,q_h)}{||(v_h,\tau_h,q_h)||_*} \\ &\leq \sup_{(v_h,\tau_h,q_h)\in V_h\times\Sigma_h\times Q_h} \frac{B(I_{V_h}u-u, I_{\Sigma_h}\sigma-\sigma, I_{Q_h}p-p; v_h,\tau_h,q_h)}{||(v_h,\tau_h,q_h)||_*}, \end{aligned}$$

where we used the consistency result of Theorem 6.8 in the last step.

Next, we estimate the terms that form $B(I_{V_h}u-u, I_{\Sigma_h}\sigma-\sigma, I_{Q_h}p-p; v_h, \tau_h, q_h)$. Using the Cauchy–Schwarz inequality,

$$\begin{aligned}
&a(I_{\Sigma_h}\sigma-\sigma,\tau_h)+b_1(I_{V_h}u-u,q_h)+b_1(v_h,I_{Q_h}p-p)\\
&\lesssim\left(\frac{1}{\sqrt{\nu}}\left\|I_{\Sigma_h}\sigma-\sigma\right\|_{\Sigma_h}\frac{1}{\sqrt{\nu}}\left\|\tau_h\right\|_{\Sigma_h}\right)+\left(\sqrt{\nu}\|I_{V_h}u-u\|_{V_h}\frac{1}{\sqrt{\nu}}\|q_h\|_{Q_h}\right)+\left(\sqrt{\nu}\|v\|_{V_h}\frac{1}{\sqrt{\nu}}\|I_{Q_h}p-p\|_{Q_h}\right)\\
&\leq ||(I_{V_h}u-u,I_{\Sigma_h}\sigma-\sigma,I_{Q_h}p-p)||_*||(v_h,\tau_h,q_h)||_*.
\end{aligned}$$

For the terms including the bilinear form b_2, we also have by the Cauchy–Schwarz inequality applied on each element and each facet

$$\begin{aligned}
&b_2(I_{\Sigma_h}\sigma-\sigma,v_h)+b_2(\tau_h,I_{V_h}u-u)\\
&\qquad\lesssim\sum_{F\in\mathscr{F}_h}\sqrt{h}||(I_{\Sigma_h}\sigma-\sigma)_{nt}||_F\frac{1}{\sqrt{h}}||[\![(v_h)_t]\!]||_F+\sum_{T\in\mathscr{T}_h}||I_{\Sigma_h}\sigma-\sigma||_T||\nabla v_h||_T\\
&\qquad+\sum_{F\in\mathscr{F}_h}\sqrt{h}||(\tau_h)_{nt}||_F\frac{1}{\sqrt{h}}||[\![(I_{V_h}u-u)_t]\!]||_F+\sum_{T\in\mathscr{T}_h}||\tau_h||_T||\nabla(I_{V_h}u-u)||_T.
\end{aligned}$$

Scaling with $\sqrt{\nu}$ and applying the norm equivalence Lemma 6.1 finally yield

$$\begin{aligned}
&b_2(I_{\Sigma_h}\sigma-\sigma,v_h)+b_2(\tau_h,I_{V_h}u-u)\\
&\quad\lesssim\left(\frac{1}{\sqrt{\nu}}\left\|I_{\Sigma_h}\sigma-\sigma\right\|_{\Sigma_h}+\frac{1}{\sqrt{\nu}}\sqrt{\sum_{F\in\mathscr{F}_h}h||(I_{\Sigma_h}\sigma-\sigma)_{nt}||_F^2}+\sqrt{\nu}||I_{V_h}u-u||_{V_h}\right)||(v_h,\tau_h,0)||_*.
\end{aligned}$$

All together this leads to the estimate

$$\begin{aligned}
&||(I_{V_h}u-u_h,I_{\Sigma_h}\sigma-\sigma_h,I_{Q_h}p-p_h)||_*\\
&\qquad\lesssim||(I_{V_h}u-u,I_{\Sigma_h}\sigma-\sigma,I_{Q_h}p-p)||_*+\frac{1}{\sqrt{\nu}}\sqrt{\sum_{F\in\mathscr{F}_h}h||(I_{\Sigma_h}\sigma-\sigma)_{nt}||_F^2}.
\end{aligned}$$

Again, with (6.13) and (5.22), we conclude the proof. □

6.4 *Pressure robustness*

We define the continuous Helmholtz projector $\mathbb{P}$ as the rotational part of a Helmholtz decomposition (see Girault & Raviart, 2012) of a given load f

$$f=\nabla\theta+\xi=:\nabla\theta+\mathbb{P}(f),$$

with $\theta \in H^1(\Omega)/\mathbb{R}$ and $\xi =: \mathbb{P}(f) \in \{v \in H_0(\mathrm{div}, \Omega) : \mathrm{div}(v) = 0\}$. Testing the second line of (3.8) with an arbitrary divergence free test function $v \in \{v \in H_0(\mathrm{div}, \Omega) : \mathrm{div}(v) = 0\}$, we see that

$$\langle \mathrm{div}\ \sigma, v\rangle_{H_0(\mathrm{div},\Omega)} = -\langle \mathbb{P}(f), v\rangle_{H_0(\mathrm{div},\Omega)};$$

hence $\sigma = \nu \nabla u$ is steered only by a part of f, namely $\mathbb{P}(f)$. If the right-hand side is perturbed by a gradient field $\nabla\alpha$, then σ and u should not change as $\mathbb{P}(f + \nabla\alpha) = \mathbb{P}(f)$. In the work by Linke (2014) this relation was discussed in a discrete setting. If a discrete method fulfils this property, it is called pressure robust because one can then deduce an H^1-velocity error that is independent of the pressure. The convergence estimate of Theorem 6.9 includes the scaled term $1/\nu||p||_{H^s(\mathscr{T}_h)}$, which blows up as $\nu \to 0$. However, the mass conserving mixed stress formulation (MCS) is pressure robust, allowing us to conclude that velocity errors do not blow up as $\nu \to 0$ by virtue of the next theorem.

THEOREM 6.10 (Pressure robustness). Let $u \in H^1(\Omega, \mathbb{R}^d) \cap H^m(\mathscr{T}_h, \mathbb{R}^d)$ and let $\sigma \in H^1(\Omega, \mathbb{R}^{d\times d}) \cap H^{m-1}(\mathscr{T}_h, \mathbb{R}^{d\times d})$ be the exact solution of the mixed Stokes problem (3.2). Further, let σ_h, u_h be the solution of the mass conserving mixed stress formulation (MCS). For $s = \min(m-1, k)$ there holds

$$||u - u_h||_{V_h} + \frac{1}{\nu}\left\|\sigma - \sigma_h\right\|_{\Sigma_h} \lesssim h^s||u||_{H^{s+1}(\mathscr{T}_h)}.$$

Proof. The proof follows along the lines of the proof of Theorem 6.9. Using the triangle inequality

$$||u - u_h||_{V_h} + \frac{1}{\nu}\left\|\sigma - \sigma_h\right\|_{\Sigma_h} \lesssim ||u - I_{V_h}u||_{V_h} + \frac{1}{\nu}\left\|\sigma - I_{\Sigma_h}\sigma\right\|_{\Sigma_h} + ||I_{V_h}u - u_h||_{V_h} + \frac{1}{\nu}\left\|I_{\Sigma_h}\sigma - \sigma_h\right\|_{\Sigma_h}.$$

The first two terms can be estimated using the approximation results (6.13) and (5.22). Next note that from the LBB condition of Lemma 6.5 on V_h^0 and the trivial coercivity inequality $a(\sigma_h, \sigma_h) \geq (1/\nu)\left\|\sigma_h\right\|^2_{\Sigma_h}$ for all $\sigma_h \in \Sigma_h$, we conclude inf-sup stability of the bilinear form $B(u_h, \sigma_h, 0; v_h, \tau_h, 0)$ with respect to the product space norm $||(\cdot, \cdot, 0)||_*$ on the subspace $V_h^0 \times \Sigma_h \times \{0\}$, i.e.,

$$\begin{aligned} ||I_{V_h}u - u_h||_{V_h} + \frac{1}{\nu}\left\|I_{\Sigma_h}\sigma - \sigma_h\right\|_{\Sigma_h} &= \frac{1}{\sqrt{\nu}}||(I_{V_h}u - u_h, I_{\Sigma_h}\sigma - \sigma_h, 0)||_* \\ &\leq \sup_{(v_h,\tau_h)\in V_h^0\times\Sigma_h} \frac{B(I_{V_h}u - u_h, I_{\Sigma_h}\sigma - \sigma_h, 0; v_h, \tau_h, 0)}{\sqrt{\nu}||(v_h, \tau_h, 0)||_*}. \end{aligned}$$

Note that the form is continuous by Lemma 6.3. By steps similar to those in the proof of the consistency result of Theorem 6.8, we have

$$B(u, \sigma, 0; v_h, \tau_h, 0) = \int_\Omega \mathrm{div}(\sigma)\cdot v_h = \int_\Omega -f\cdot v_h + \int_\Omega \nabla p\cdot v_h = \int_\Omega -f\cdot v_h$$

for all $v_h, \tau_h \in V_h^0 \times \Sigma_h$, where we used $\mathrm{div}(\sigma) = -f + \nabla p$ and integration by parts for ∇p. This shows that the method is also consistent on the subspace of divergence-free velocity test functions, a key ingredient to obtain pressure robustness. We now have

$$\sup_{(v_h,\tau_h)\in V_h^0\times\Sigma_h}\frac{B(I_{V_h}u-u_h,I_{\Sigma_h}\sigma-\sigma_h,0;v_h,\tau_h,0)}{\sqrt{\nu}||(v_h,\tau_h,q_h)||_*}=\sup_{(v_h,\tau_h)\in V_h^0\times\Sigma_h}\frac{B(I_{V_h}u-u,I_{\Sigma_h}\sigma-\sigma,0;v_h,\tau_h,0)}{\sqrt{\nu}||(v_h,\tau_h,q_h)||_*}.$$

The rest of the proof follows along the previous lines using the identity $\sigma=\nu\nabla u$, and we obtain

$$||u-u_h||_{V_h}+\frac{1}{\nu}\left\|\sigma-\sigma_h\right\|_{\Sigma_h}\lesssim h^s\left(||u||_{H^{s+1}(\mathscr{T}_h)}+\frac{1}{\nu}||\sigma||_{H^s(\mathscr{T}_h)}\right)\le h^s||u||_{H^{s+1}(\mathscr{T}_h)}.$$

□

7. Numerical examples

In the following we present a numerical example to validate the results of Section 6. All numerical examples were implemented within the finite element library NGSolve/Netgen, see Schöberl (1997, 2014). Let $\Omega=[0,1]^d$ and choose the right-hand side $f=-\operatorname{div}(\sigma)+\nabla p$ with the exact solution given by

$$\sigma=\nu\nabla\operatorname{curl}(\psi_2),\quad\text{and}\quad p:=x^5+y^5-\frac{1}{3}\quad\text{for } d=2$$

$$\sigma=\nu\nabla\operatorname{curl}(\psi_3,\psi_3,\psi_3),\quad\text{and}\quad p:=x^5+y^5+z^5-\frac{1}{2}\quad\text{for } d=3,$$

where $\psi_2:=x^2(x-1)^2y^2(y-1)^2$ and $\psi_3:=x^2(x-1)^2y^2(y-1)^2z^2(z-1)^2$ defines velocity through a vector and scalar potential in two and three dimensions, respectively. In Fig. 2 different errors are plotted for varying polynomial orders $k=2,3,4,5$ in the two-dimensional case with a fixed viscosity $\nu=10^{-3}$. As predicted by Theorem 6.9 the H^1-seminorm error of the velocity, the L^2-norm error of the stress and the L^2-norm error of the pressure have the same optimal convergence rate.

The L^2-norm of the velocity error converges at one higher order as shown in the bottom right plot of Fig. 2. This can be explained by the standard Aubin–Nitsche duality argument, by which we can prove

$$||u-u_h||_{L^2(\Omega)}\le h^{k+1}||u||_{H^{k+1}(\mathscr{T}_h)}$$

whenever the problem admits full elliptic regularity and the exact solution u is smoother. This argument works in both two and three dimensions. The higher observed rate of convergence in three dimensions (for $\nu=10^{-3}$), given by the estimated order of convergence (eoc), can be seen in Table 1.

Next, we study pressure robustness. The above-mentioned right-hand side f consists of an irrotational part (the gradient of the pressure) and a part with curl. We study how the velocity error (in H^1 seminorm) varies as $\nu\to 0$ for the presented MCS method and the standard Taylor–Hood method—see e.g., Brezzi & Falk (1991) and Girault & Raviart (2012)—using the same polynomial approximation order for the velocity in the two-dimensional case. We observe in Fig. 3 that the error of the Taylor–Hood method increases as $\nu\to 0$ and behaves as if it were scaled by a factor $1/\nu$ for small values of ν. This is the locking phenomenon we discussed earlier: clearly the Taylor–Hood method is not pressure robust (and does not provide exactly divergence-free numerical velocity). In contrast, the velocity errors in the MCS method (also in Fig. 3) appear to be not influenced by varying values

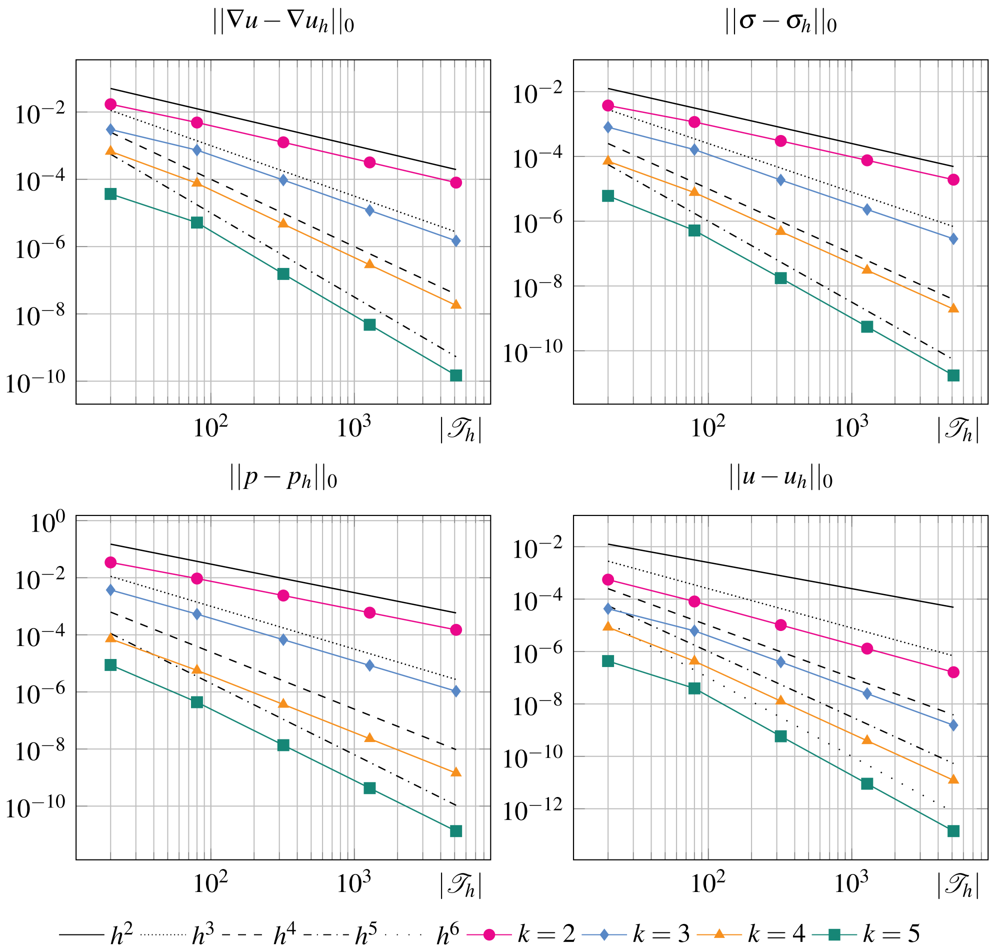

FIG. 2. Convergence plots for the two-dimensional case with a fixed viscosity $\nu = 10^{-3}$.

of ν. This behaviour is observed for several polynomial orders $k = 2, 3, 4$. These observations match the predictions of Theorem 6.10.

We conclude with a few remarks on the cost of solving the discrete system (MCS). After an element-wise static condensation step there are two different types of dofs that couple at element interfaces. These coupling dofs determine the costs for the factorization step of the assembled matrix. In the $d = 2$ case, the normal continuity of the $H(\mathrm{div})$-conforming velocity space demands $k + 1$ dofs per interface, while the normal-tangential continuity of the stress space Σ_h requires k dofs, i.e., we have $2k + 1$ dofs per interface. This is comparable to the number of interface dofs for standard methods. In fact, it is identical to the number of dofs per interface of an advanced method (with a reduced stabilization called 'projected jumps') presented in the recent work of Lehrenfeld & Schöberl (2016). Similar cost comparison observations apply for the $d = 3$ case.

TABLE 1 *The H^1-seminorm error of the velocity, the L^2-norm error of the pressure and the stress and the L^2-norm error of the velocity for different polynomial orders $k = 1, 2, 3$ for the three-dimensional case and a fixed viscosity $\nu = 10^{-3}$*

$\lvert\mathscr{T}\rvert$	$\lVert\nabla u - \nabla u_h\rVert_0$	(eoc)	$\lVert\sigma - \sigma_h\rVert_0$	(eoc)	$\lVert p - p_h\rVert_0$	(eoc)	$\lVert u - u_h\rVert_0$	(eoc)
$k = 1$								
28	$4.6 \cdot 10^{-3}$	(–)	$3.5 \cdot 10^{-3}$	(–)	$2.4 \cdot 10^{-1}$	(–)	$4.3 \cdot 10^{-4}$	(–)
224	$3.9 \cdot 10^{-3}$	(0.2)	$2.7 \cdot 10^{-3}$	(0.4)	$1.7 \cdot 10^{-1}$	(0.6)	$2.6 \cdot 10^{-4}$	(0.7)
1792	$2.3 \cdot 10^{-3}$	(0.8)	$1.3 \cdot 10^{-3}$	(1.1)	$8.9 \cdot 10^{-2}$	(0.9)	$7.6 \cdot 10^{-5}$	(1.8)
14336	$1.1 \cdot 10^{-3}$	(1.0)	$6.3 \cdot 10^{-4}$	(1.0)	$4.6 \cdot 10^{-2}$	(1.0)	$1.9 \cdot 10^{-5}$	(2.0)
114688	$5.6 \cdot 10^{-4}$	(1.0)	$3.1 \cdot 10^{-4}$	(1.0)	$2.3 \cdot 10^{-2}$	(1.0)	$4.9 \cdot 10^{-6}$	(2.0)
$k = 2$								
28	$2.8 \cdot 10^{-3}$	(–)	$1.9 \cdot 10^{-3}$	(–)	$7.5 \cdot 10^{-2}$	(–)	$1.4 \cdot 10^{-4}$	(–)
224	$1.5 \cdot 10^{-3}$	(0.9)	$4.6 \cdot 10^{-4}$	(2.0)	$3.1 \cdot 10^{-2}$	(1.3)	$3.4 \cdot 10^{-5}$	(2.0)
1792	$5.8 \cdot 10^{-4}$	(1.4)	$1.8 \cdot 10^{-4}$	(1.4)	$9.5 \cdot 10^{-3}$	(1.7)	$8.2 \cdot 10^{-6}$	(2.1)
14336	$1.7 \cdot 10^{-4}$	(1.8)	$4.9 \cdot 10^{-5}$	(1.9)	$2.5 \cdot 10^{-3}$	(1.9)	$1.3 \cdot 10^{-6}$	(2.7)
114688	$4.4 \cdot 10^{-5}$	(2.0)	$1.3 \cdot 10^{-5}$	(2.0)	$6.4 \cdot 10^{-4}$	(2.0)	$1.6 \cdot 10^{-7}$	(2.9)
$k = 3$								
28	$1.0 \cdot 10^{-3}$	(–)	$2.9 \cdot 10^{-4}$	(–)	$6.7 \cdot 10^{-3}$	(–)	$2.5 \cdot 10^{-5}$	(–)
224	$4.8 \cdot 10^{-4}$	(1.1)	$9.2 \cdot 10^{-5}$	(1.6)	$1.6 \cdot 10^{-3}$	(2.1)	$6.3 \cdot 10^{-6}$	(2.0)
1792	$1.5 \cdot 10^{-4}$	(1.7)	$1.7 \cdot 10^{-5}$	(2.4)	$2.6 \cdot 10^{-4}$	(2.6)	$1.0 \cdot 10^{-6}$	(2.6)
14336	$2.0 \cdot 10^{-5}$	(2.9)	$2.4 \cdot 10^{-6}$	(2.8)	$3.5 \cdot 10^{-5}$	(2.9)	$7.2 \cdot 10^{-8}$	(3.9)
114688	$2.6 \cdot 10^{-6}$	(2.9)	$3.1 \cdot 10^{-7}$	(2.9)	$4.5 \cdot 10^{-6}$	(3.0)	$4.7 \cdot 10^{-9}$	(3.9)

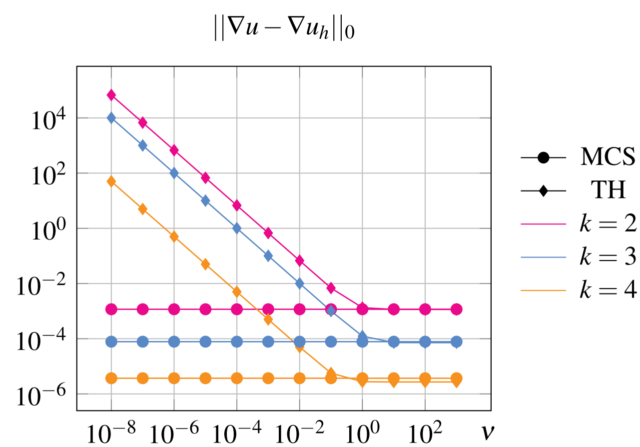

FIG. 3. The H^1-seminorm error for the MCS method and a Taylor–Hood approximation for $k = 2, 3, 4$ and varying viscosity ν.

Acknowledgements

Part of this work was completed while two of the authors were at Basque Center for Applied Mathematics (BCAM). We gratefully acknowledge the hospitality of BCAM.

Funding

Austrian Science Fund (FWF) through the research programme ‘Taming Complexity in Partial Differential Systems’ (F65)—project ‘Automated Discretization in Multiphysics’(P10) (P.L.); Basque Center for Applied Mathematics (MINECO: BCAM Severo Ochoa Excellence Accreditation SEV-2013-0323) National Science Foundation (USA).

References

Abramowitz, M. (1974) *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*. Dover Publications, Incorporated.

Andrews, G., Askey, R. & Roy, R. (1999) *Special Functions*. Encyclopedia of Mathematics and its Applications. Cambridge, England: Cambridge University Press.

Boffi, D., Brezzi, F. & Fortin, M. (2013) *Mixed Finite Element Methods and Applications*. Berlin, Germany: Springer Science & Business Media.

Braess, D. (2007) *Finite Elements: Theory, Fast Solvers, and Applications in Elasticity Theory*, 3rd edn. Cambridge: Cambridge University Press, pp. xviii+365. (Translated from the German by Larry L. Schumaker.)

Braess, D. (2013) *Finite Elemente—Theorie, Schnelle Löser und Anwendungen in der Elastizitätstheorie*. Berlin, Germany: Springer.

Brennecke, C., Linke, A., Merdon, C. & Schöberl, J. (2015) Optimal and pressure-independent L^2 velocity error estimates for a modified Crouzeix–Raviart Stokes element with BDM reconstructions. *J. Comput. Math.*, **33**, 191–208.

Brezzi, F., Douglas Jr., J. & Marini, L. D. (1985) Two families of mixed finite elements for second order elliptic problems. *Numer. Math.*, **47**, 217–235.

Brezzi, F. & Falk, R. S. (1991) Stability of higher-order Hood–Taylor method. *SIAM J. Numer. Anal.*, **28**.

Brezzi, F. & Fortin, M. (1986) Numerical approximation of Mindlin–Reissner plates. *Math. Comp.*, **47**, 151–158.

Christiansen, S. H. (2011) On the linearization of Regge calculus. *Numer. Math.*, **119**, 613–640.

Ciarlet, P. (2002) *The Finite Element Method for Elliptic Problems*. Classics in Applied Mathematics. University City, Philadelphia, USA: Society for Industrial and Applied Mathematics.

Cockburn, B., Kanschat, G. & Schötzau, D. (2005) A locally conservative LDG method for the incompressible Navier–Stokes equations. *Math. Comp.*, **74**, 1067–1095.

Cockburn, B., Kanschat, G. & Schötzau, D. (2007) A note on discontinuous Galerkin divergence-free solutions of the Navier–Stokes equations. *J. Sci. Comput.*, **31**, 61–73.

Cockburn, B. & Sayas, F.-J. (2014) Divergence-conforming HDG methods for Stokes flows. *Math. Comp.*, **83**, 1571–1598.

Costabel, M. & McIntosh, A. (2010) On Bogovskiĭ and regularized Poincaré integral operators for de Rham complexes on Lipschitz domains. *Math. Z.*, **265**, 297–320.

Demlow, A. & Hirani, A. N. (2014) A posteriori error estimates for finite element exterior calculus: the de Rham complex. *Found. Comput. Math.*, **14**, 1337–1371.

Dubiner, M. (1991) Spectral methods on triangles and other domains. *J. Sci. Comput.*, **6**, 345–390.

Ern, A. & Guermond, J.-L. (2004) *Theory and Practice of Finite Elements*. Applied Mathematical Sciences, vol. 159, 1st edn. New York: Springer.

Farhloul, M. (1995) Mixed and nonconforming finite element methods for the stokes problem. *Can. Appl. Math. Q.*, **3**, 399–418.

Farhloul, M. & Fortin, M. (1993) A new mixed finite element for the Stokes and elasticity problems. *SIAM J. Numer. Anal.*, **30**, 971–990.

Farhloul, M. & Fortin, M. (1997) Dual hybrid methods for the elasticity and the Stokes problems: a unified approach. *Numer. Math.*, **76**, 419–440.

Farhloul, M. & Fortin, M. (2002) Review and complements on mixed-hybrid finite element methods for fluid flows. *Proceedings of the 9th International Congress on Computational and Applied Mathematics (Leuven, 2000)*, vol. **140**, pp. 301–313.

Fu, G., Jin, Y. & Qiu, W. (2018) Parameter-free superconvergent H(div)-conforming HDG methods for the Brinkman equations. *IMA J. Numer. Anal.*, **00**, 1–26.

Girault, V. & Raviart, P.-A. (2012) *Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms*, vol. **5**. Berlin: Springer Science & Business Media.

John, V., Linke, A., Merdon, C., Neilan, M. & Rebholz, L. (2017) On the divergence constraint in mixed finite element methods for incompressible flows. *SIAM Rev.*, **59**, 492–544.

Karniadakis, G. & Sherwin, S. (2013) *Spectral/hp Element Methods for Computational Fluid Dynamics*. Oxford, England, UK : Oxford University Press.

Könnö, J. & Stenberg, R. (2012) Numerical computations with H(div)-finite elements for the Brinkman problem. *Comput. Geosci.*, **16**, 139–158.

Lederer, P. L., Lehrenfeld, C. & Schöberl, J. (2017a) Hybrid discontinuous Galerkin methods with relaxed H(*div*)-conformity for incompressible flows. Part I. *SIAM J Numer. Anal.* (submitted). Preprint available at arXiv:1707.02782.

Lederer, P. L., Linke, A., Merdon, C. & Schöberl, J. (2017b) Divergence-free reconstruction operators for pressure-robust Stokes discretizations with continuous pressure finite elements. *SIAM J. Numer. Anal.*, **55**, 1291–1314.

Lederer, P. L., Lehrenfeld, C. & Schöberl, J. (2018) Hybrid discontinuous Galerkin methods with relaxed H(*div*)-conformity for incompressible flows. Part II. *SIAM J. Numer. Anal.*, **56**, 2070–2094.

Lederer, P. L. & Schöberl, J. (2017) Polynomial robust stability analysis for H(div)-conforming finite elements for the Stokes equations. *IMA J. Numer. Anal.*, **drx051**.

Lehrenfeld, C. & Schöberl, J. (2016) High order exactly divergence-free hybrid discontinuous Galerkin methods for unsteady incompressible flows. *Comput. Methods Appl. Mech. Engrg.*, **307**, 339–361.

Linke, A. (2012) A divergence-free velocity reconstruction for incompressible flows. *C. R. Math. Acad. Sci. Paris*, **350**, 837–840.

Linke, A. (2014) On the role of the Helmholtz decomposition in mixed methods for incompressible flows and a new variational crime. *Comput. Methods Appl. Mech. Engrg.*, **268**, 782–800.

Linke, A., Matthies, G. & Tobiska, L. (2016) Robust arbitrary order mixed finite element methods for the incompressible Stokes equations with pressure independent velocity errors. *ESAIM: M2AN*, **50**, 289–309.

Lizao, L. (2018) Regge finite elements with applications in solid mechanics and relativity. *Ph.D. Thesis*, Minneapolis, Minnesota, USA: University of Minnesota.

Pechstein, A. S. & Schöberl, J. (2011) Tangential-displacement and normal-normal-stress continuous mixed finite elements for elasticity. *Math. Models Methods Appl. Sci.*, **21**, 1761–1782.

Pechstein, A. S. & Schöberl, J. (2017) The TDNNS method for Reissner–Mindlin plates. *Numer. Math.*, **137**, 713–740.

Raviart, P.-A. & Thomas, J. M. (1977) A mixed finite element method for 2nd order elliptic problems. *Mathematical Aspects of Finite Element Methods (Proceedings of the Conference Held in Rome, 1975)*. Lecture Notes in Mathematics, vol. **606**. Berlin: Springer, pp. 292–315.

Schöberl, J. (1997) NETGEN an advancing front 2D/3D-mesh generator based on abstract rules. *Comput. Vis. Sci.*, **1**, 41–52.

Schöberl, J. (2014) C++11 implementation of finite elements in NGSolve. Institute for Analysis and Scientific Computing. Vienna, Austria: Vienna University of Technology.

Sinwel, A. S. (2009) A new family of mixed finite elements for elasticity. *Ph.D. Thesis*, Linz, Austria: Johannes Kepler University Linz.

Zaglmayr, S. (2006) High order finite element methods for electromagnetic field computation. *Ph.D. Thesis*, Linz, Austria: Johannes Kepler University Linz.