Inf–sup conditions for twofold saddle point problems

Jason S. Howell · Noel J. Walkington

Abstract Necessary and sufficient conditions for existence and uniqueness of solutions are developed for twofold saddle point problems which arise in mixed formulations of problems in continuum mechanics. This work extends the classical saddle point theory to accommodate nonlinear constitutive relations and the twofold saddle structure. Application to problems in incompressible fluid mechanics employing symmetric tensor finite elements for the stress approximation is presented.

Mathematics Subject Classification (2000)   65N30

1 Introduction

Many problems in continuous fluid and solid mechanics lead to a variational problems with a saddle point structure of the form: $(u, p_1, p_2) \in U \times P_1 \times P_2$,

$$
\begin{bmatrix}
A & -B_1^T & -B_2^T \\
B_1 & 0 & 0 \\
B_2 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
u \\
p_1 \\
p_2
\end{bmatrix}
=
\begin{bmatrix}
f \\
g_1 \\
g_2
\end{bmatrix}
\tag{1.1}
$$

J. S. Howell (✉)
Department of Mathematics, Clarkson University, Potsdam, NY 13676, USA
e-mail: jhowell@clarkson.edu

N. J. Walkington
Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA
e-mail: noelw@andrew.cmu.edu

or

$$\begin{bmatrix} A & 0 & -B_1^T \\ 0 & 0 & -B_2^T \\ B_1 & B_2 & 0 \end{bmatrix} \begin{bmatrix} u \\ p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} f \\ g_1 \\ g_2 \end{bmatrix}. \tag{1.2}$$

Here U, P_1 and P_2 are Banach spaces, $A : U \to U'$ is typically nonlinear, and B_1 and B_2 are linear operators. In this paper we develop necessary and sufficient conditions for this class of problems to be well-posed and consider their numerical approximation.

Problems of type (1.2) can be written as

$$\begin{bmatrix} A & -B_1^T & 0 \\ B_1 & 0 & B_2 \\ 0 & -B_2^T & 0 \end{bmatrix} \begin{bmatrix} u \\ p_2 \\ p_1 \end{bmatrix} = \begin{bmatrix} f \\ g_2 \\ g_1 \end{bmatrix}. \tag{1.3}$$

These problems are often described as *twofold saddle point problems* due to the nested structure they exhibit. Each of these problems can be viewed as a single saddle point problem of the form $\begin{bmatrix} A & -B^T \\ B & 0 \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}$ where B acts on a product space. The classical theory requires B to satisfy an inf–sup condition [8,13,14]. The development of discrete spaces which inherit such inf–sup conditions for each specific problem is notoriously difficult, and this difficulty is compounded by the product structure of the twofold saddle point problems. Our main results, Theorems 3.1 and 3.2, provide several equivalent conditions on the individual operators B_1 and B_2 to guarantee that the compound operator B satisfies inf–sup conditions on the corresponding product space. These conditions are necessary and sufficient, and frequently one of the conditions is easier to verify than another for a specific problem.

The rest of this paper is organized as follows. This introductory section continues with a motivating example, followed by a review of prior results and some notation. Section 2 considers solution of single saddle point problems with linear and nonlinear operator A and their Galerkin approximation. Section 3 develops necessary and sufficient conditions on the operators B_1 and B_2 for problems of type (1.1) and (1.2) to be well posed. An application of the theory is illustrated in Sect. 4.

1.1 Examples of twofold saddle point problems

Twofold saddle point problems arise ubiquitously when mixed finite element formulations are used to approximate the stress in an incompressible fluid or solid [2,6,9, 15,25–29,31]. When the usual linear relation $S_0 = \nu D(\mathbf{u})$ for the viscous stress is replaced with a more general relation $S_0 = \nu \mathcal{A}(D(\mathbf{u}))$ the equations for the creeping (Stokes) flow of a fluid in a domain $\Omega \subset \mathbb{R}^d$ become,

$$- \operatorname{div}(S) = \mathbf{f}, \quad \operatorname{div}(\mathbf{u}) = 0, \quad S = -pI + \mathcal{A}(D(\mathbf{u})). \tag{1.4}$$

Here $D(\mathbf{u}) = (1/2)(\nabla \mathbf{u} + (\nabla \mathbf{u})^T)$, S is the stress tensor, and for incompressible fluids takes the form $S = -pI + S_0$ where $tr(S_0) = 0$ is the deviatoric part of the stress.

If $\partial\Omega = \Gamma_0 \cup \Gamma_1$ typical boundary conditions would be

$$\mathbf{u}|_{\Gamma_0} = u_0, \qquad Sn|_{\Gamma_1} = \mathbf{t}. \tag{1.5}$$

To illustrate our results we review three formulations of this problem:

(1) A single saddle point problem arises when the stress S is eliminated using the constitutive relation as in [32] (classical *mixed* formulation).
(2) Eliminating the gradient of the velocity using the constitutive relation gives rise to a *dual mixed* variational formulation [38]. A twofold saddle point problem of the form (1.1) then arises when the problem is posed in three variables [9] or the symmetry of the stress tensor is enforced weakly [2].
(3) Alternatively, independent approximation of (the symmetric part of) the velocity gradient gives rise to a twofold saddle problem of the form (1.2) (alternate *dual mixed* formulation). Dual mixed approaches that approximate the (nonsymmetric) velocity gradient directly are presented in [19,27].

When the fluid is incompressible the stress-strain relation only acts on the trace free (deviatoric) part of the strain. Letting $(\mathbb{R}^{d \times d}_{sym})_0$ denote the symmetric trace free matrices, we assume

1. $\mathcal{A} : \mathbb{R}^{d \times d} \to (\mathbb{R}^{d \times d}_{sym})_0$.
2. $\mathcal{A}(D) = \mathcal{A}(D_{sym} - (tr(D)/d)I)$.
3. The restriction $\mathcal{A} : (\mathbb{R}^{d \times d}_{sym})_0 \to (\mathbb{R}^{d \times d}_{sym})_0$ is bijective.

Classical mixed formulation The classical weak statement of this problem seeks $\mathbf{u} - u_0 \in U = \{\mathbf{u} \in H^1(\Omega)^d \mid \mathbf{u}|_{\Gamma_0} = 0\}$ and $p \in L^2(\Omega)$ satisfying

$$\int_\Omega \left(\mathcal{A}(D(\mathbf{u})) : D(\mathbf{v}) - p\,\mathrm{div}(\mathbf{v}) \right) = \int_\Omega \mathbf{f}.\mathbf{v} + \int_{\Gamma_1} \mathbf{t}.\mathbf{v}, \qquad \mathbf{v} \in U,$$

$$\int_\Omega \mathrm{div}(\mathbf{u})\,q = 0, \qquad q \in L^2(\Omega).$$

If $\Gamma_1 = \emptyset$, then p, q are required to have average zero. This is a classical saddle point problem taking the form

$$\begin{bmatrix} \mathcal{A} & -\mathrm{div}' \\ \mathrm{div} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ p \end{bmatrix} = \begin{bmatrix} \mathbf{f} + \gamma'(\mathbf{t}) \\ 0 \end{bmatrix}.$$

Here $\langle \mathrm{div}'(p), \mathbf{u} \rangle = (p, \mathrm{div}(\mathbf{u}))$ is the dual operator, and similarly $\langle \gamma'(\mathbf{t}), \mathbf{v} \rangle = (\mathbf{t}, \gamma(\mathbf{v}))_{\Gamma_1}$ is the dual of the trace operator $\gamma : U \to U/H_0^1(\Omega)^d$. Here $(\cdot, \cdot)$ is the $L^2(\Omega)$ inner product and $\langle \cdot, \cdot \rangle$ is the induced duality pairing.

Dual mixed formulation If the inverse of the stress strain relation is available, $D(\mathbf{u}) = \mathcal{A}^{-1}(S)$, it is possible to write a weak statement that evaluates the stress explicitly. Let

$$\mathbb{S} = \{S \in \mathbb{H}(\Omega; \mathrm{div})_{sym} \mid Sn|_{\Gamma_1} = 0\} \quad \text{and} \quad U = L^2(\Omega),$$

and $\mathbb{S}(\mathbf{t})$ be the elements in $S \in \mathbb{H}(\Omega; \mathrm{div})_{sym}$ for which $S\mathbf{n}|_{\Gamma_1} = \mathbf{t}$. Then $(S, \mathbf{u}) \in \mathbb{S}(\mathbf{t}) \times U$ satisfies

$$\int_\Omega \left(\mathcal{A}^{-1}(S) : T + \mathbf{u}.\mathrm{div}(T) \right) = \int_{\Gamma_0} \mathbf{u}_0.T\mathbf{n}, \qquad T \in \mathbb{S},$$

$$\int_\Omega -\mathrm{div}(S).\mathbf{v} = \int_\Omega f.\mathbf{v}. \qquad \mathbf{v} \in U.$$

This weak statement is again a classical saddle point problem. A mixed finite element method for this formulation in linear elasticity was studied in [3].

If finite element subspaces of $\mathbb{H}(\Omega; \mathrm{div})_{sym}$ are not available, it is possible to pose the previous weak statement in $\mathbb{H}(\Omega; \mathrm{div})$ and to use Raviart–Thomas elements. In this instance $\nabla\mathbf{u} = \mathcal{A}^{-1}(S) + (\nabla\mathbf{u})_{skew}$. Let

$$\mathbb{T} = \{ S \in \mathbb{H}(\Omega; \mathrm{div}) \mid S\mathbf{n}|_{\Gamma_1} = 0 \} \quad \text{and} \quad U = (L^2(\Omega))^2,$$

and $\mathbb{T}(\mathbf{t})$ be the elements in $S \in \mathbb{H}(\Omega; \mathrm{div})$ for which $S\mathbf{n}|_{\Gamma_1} = \mathbf{t}$. Next, define $W : \mathbb{R}^d \to \mathbb{R}^{d \times d}_{skew}$ by $W(\mathbf{w})_{ij} = \epsilon_{ijk} w_k$ Then $(S, \mathbf{u}, \mathbf{w}) \in \mathbb{T}(\mathbf{t}) \times U \times L^2(\Omega)$ satisfies

$$\int_\Omega \left(\mathcal{A}^{-1}(S) : T + W(\mathbf{w}) : T + \mathbf{u}.\mathrm{div}(T) \right) = \int_{\Gamma_0} \mathbf{u}_0.T\mathbf{n}, \qquad T \in \mathbb{T},$$

$$\int_\Omega -\mathrm{div}(S).\mathbf{v} = \int_\Omega f.\mathbf{v}, \qquad \mathbf{v} \in U,$$

$$\int_\Omega -S : W(\mathbf{z}) = 0, \qquad \mathbf{z} \in L^2(\Omega).$$

This is a twofold saddle point problem taking the form

$$\begin{bmatrix} \mathcal{A}^{-1} & \mathrm{div}' & W \\ -\mathrm{div} & 0 & 0 \\ -W' & 0 & 0 \end{bmatrix} \begin{bmatrix} S \\ \mathbf{u} \\ \mathbf{w} \end{bmatrix} = \begin{bmatrix} \gamma'(\mathbf{u}_0) \\ \mathbf{f} \\ 0 \end{bmatrix}. \tag{1.6}$$

Here $\langle \gamma'(\mathbf{u}_0), T \rangle = (\mathbf{u}_0, T\mathbf{n})_{\Gamma_0}$.

Finite element methods for the formulation (1.6) arising in linear elasticity were developed by Arnold et al. [2] (see [6] for a survey of approaches with weakly imposed symmetry).

Alternative dual mixed formulation If the inverse $\mathcal{A}^{-1}$ if the constitutive relation is not available it is possible to pose a mixed formulation which computes both $D(\mathbf{u})$ and S. Let $\mathbb{S}$ and $U = L^2(\Omega)^d$ be as above and

$$\mathbb{D} = \{ D \in L^2(\Omega)^{d \times d}_{sym} \mid tr(D) = 0 \}.$$

and $\mathbb{S}(\mathbf{t})$ be the elements in $S \in \mathbb{H}(\Omega; \text{div})_{sym}$ for which $S\mathbf{n}|_{\Gamma_1} = \mathbf{t}$. Then $(D, \mathbf{u}, S) \in \mathbb{D} \times U \times \mathbb{S}(\mathbf{t})$ satisfies

$$\int_\Omega \mathcal{A}(D) : E - S : E = 0, \qquad E \in \mathbb{D},$$

$$\int_\Omega (D : T + \mathbf{u}.\text{div}(T)) = \int_{\Gamma_0} \mathbf{u}_0.T\mathbf{n}, \qquad T \in \mathbb{S}, \qquad (1.7)$$

$$\int_\Omega -\text{div}(S).\mathbf{v} = \int_\Omega f.\mathbf{v}. \qquad \mathbf{v} \in U.$$

This is a twofold saddle point problem taking the form

$$\begin{bmatrix} \mathcal{A} & 0 & -I \\ 0 & 0 & -\text{div} \\ I' & \text{div}' & 0 \end{bmatrix} \begin{bmatrix} D \\ \mathbf{u} \\ S \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{f} \\ \gamma'(\mathbf{u}_0) \end{bmatrix}. \qquad (1.8)$$

If the trace-free requirement on tensors in $\mathbb{D}$ is relaxed, a formulation which includes the pressure requires $(D, \mathbf{u}, S, p) \in L^2(\Omega)_{sym}^{d \times d} \times U \times \mathbb{S}(\mathbf{t}) \times P$ satisfying

$$\int_\Omega \mathcal{A}(D) : E - S : E - pI : E = 0, \qquad E \in L^2(\Omega)_{sym}^{d \times d},$$

$$\int_\Omega (D : T + qI : D + \mathbf{u}.\text{div}(T)) = \int_{\Gamma_0} \mathbf{u}_0.T\mathbf{n}, \qquad (T, q) \in \mathbb{S} \times P, \qquad (1.9)$$

$$\int_\Omega -\text{div}(S).\mathbf{v} = \int_\Omega f.\mathbf{v}. \qquad \mathbf{v} \in U.$$

where $P = L^2(\Omega)$. This formulation takes the form of (1.8) where the identity operator is replaced with $I \times p \, \text{tr}(.)$.

1.2 Related results

Problems (1.1) and (1.2) are simplified versions of a full twofold saddle point problem described by Brezzi and Fortin [14, p. 41, §II.1]. Existing results regarding different formulations of inf–sup conditions for problems of type (1.1) can be found in [30] and [20]. Problems with this structure arise in various applications, including elasticity [2,6], complex fluids [9,20], and hybridized mixed methods for second order elliptic problems [18]. Solvability of these problems is usually established by showing that A has the required properties and either (i) the combined operator $B_1 + B_2$ is surjective or (ii) the operators B_1 and B_2 individually satisfy appropriate inf–sup conditions.

Problems of type (1.2) are analyzed in a similar manner. In [24], problems of this type were treated as nested saddle problems, with the upper left 2×2 block of (1.3)

being treated as one operator and then showing that block is coercive over the kernel of the operator B_2. Sufficient conditions for solvability and abstract error analysis are also given in [24]. This theory has been applied to many problems in elasticity and fluids as well, see [15,25–29,31] for some examples. Problems of type (1.2) were also studied in [16,17]. In these works the twofold saddle point problems are treated as single saddle point problems, and solvability is shown by requiring $B_1 + B_2$ to satisfy a certain inf–sup condition.

1.3 Notation

For Banach spaces X and Y, X' and Y' denote their duals, and if $F : X \to Y$ is an operator, its dual operator is denoted by $F' : Y' \to X'$ or $F^T : Y' \to X'$. If $Z \subset X$ is a closed subspace the quotient space is denoted by X/Z and frequently $x \in X$ is identified with $x + Z \in X/Z$. The dual space $(X/Z)'$ will be identified with the (polar) subspace of X',

$$Z^0 = \{f \in X' \mid f(z) = 0, \ z \in Z\}.$$

Under this identification $\|f\|_{(X/Z)'} = \|f\|_{X'}$.

Standard notation is used for the Lebesgue spaces $L^p(\Omega)$, the Sobolev spaces $W^{k,p}(\Omega)$ and $H^k(\Omega) = W^{k,2}(\Omega)$. We will always assume $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain with boundary partition $\partial\Omega = \Gamma_0 \cup \Gamma_1$. The inner product $L^2(\Omega)$ will be denoted as $(\cdot, \cdot)$. Recall that $\mathbb{H}(\Omega; \mathrm{div})$ is the space of vector valued functions in $L^2(\Omega)^d$ having divergence in $L^2(\Omega)$,

$$\mathbb{H}(\Omega; \mathrm{div}) = \{v \in L^2(\Omega)^d \mid \mathrm{div}(v) \in L^2(\Omega)\},$$

equipped with inner-product

$$(v, w)_{\mathbb{H}(\Omega;\mathrm{div})} = (v, w) + (\mathrm{div}(v), \mathrm{div}(w)).$$

The space of matrix (tensor) valued functions $S \in L^2(\Omega)^{d \times d}$ having divergence $\mathrm{div}(S) = S_{ij,j} \in L^2(\Omega)^d$ is denoted by $\mathbb{H}(\Omega; \mathrm{div})$ with inner product

$$(S, T)_{\mathbb{H}(\Omega;\mathrm{div})} = (S, T) + (\mathrm{div}(S), \mathrm{div}(T)).$$

Matrix valued quantities will be denoted with upper case letters, vectors with bold face lower case letters.

2 Single saddle point problems

We give a brief overview of the solvability and numerical approximation of the single saddle point problem. In Sect. 3 it is shown that the double saddle point problems can be viewed as single saddle point problems, so the general results from this section

will be applicable. We begin by stating the Generalized Lax Milgram Theorem [8,13] which will be used ubiquitously below. The following is a particularly convenient statement of this fundamental result.

Theorem 2.1 (Generalized Lax Milgram) *Let U and V be a Banach spaces, V be reflexive, and c > 0. Let a : U × V → ℝ be bilinear and continuous; that is, there exists C > 0 such that $|a(u, v)| \leq C\|u\|_U \|v\|_V$ for all u ∈ U and v ∈ V. Then the following are equivalent:*

[C] *(Coercivity) For each u ∈ U,*

$$\sup_{0 \neq v \in V} \frac{a(u, v)}{\|v\|_V} \geq c\|u\|_U,$$

and for each v ∈ V\{0}, $\sup_{u \in U} a(u, v) > 0$.

[E] *(Existence of Solutions) For each f ∈ V′ there exists a unique u ∈ U such that*

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

and $\|u\|_U \leq (1/c)\|f\|_{V'}$.

[E′] *(Existence of Solutions for the Adjoint Problem) For each g ∈ U′ there exists a unique v ∈ V such that*

$$a(u, v) = g(u), \qquad \forall\, u \in U,$$

and $\|v\|_V \leq (1/c)\|g\|_{U'}$.

In addition, if U is reflexive then each of the above are equivalent to:

[C′] *(Adjoint Coercivity) For each u ∈ U,*

$$\sup_{0 \neq u \in U} \frac{a(u, v)}{\|u\|_U} \geq c\|v\|_V,$$

and for each u ∈ U\{0}, $\sup_{v \in V} a(u, v) > 0$.

2.1 Linear saddle point problem

The following classical result for linear saddle point problems is a special case of the Generalized Lax Milgram Theorem [14].

Corollary 2.1 *Let U and P be reflexive Banach spaces and assume that a : U × U → ℝ and b : P × U → ℝ are continuous and bilinear. Then the following are equivalent.*

1. *(Existence of solutions) For all f ∈ U′ and g ∈ P′ there exists (u, p) ∈ U × P such that*

$$a(u, v) - b(p, v) = f(v), \quad and \quad b(q, u) = g(q),$$

for all $(v, q) \in U \times P$, *and there exists* $C > 0$ *such that* $\|u\|_U + \|p\|_P \leq C(\|f\|_{U'} + \|g\|_{P'})$.

2. *There exists a constant* $c > 0$ *such that*

   (a) *The restriction* $a : Z \times Z \to \mathbb{R}$ *is coercive over* $Z \equiv \{u \in U \mid b(p, u) = 0, \ \forall p \in P\}$.

$$\sup_{v \in Z} \frac{a(u, v)}{\|v\|_U} \geq c\|u\|_U, \quad \text{and} \quad \sup_{u \in Z} a(u, v) > 0, \quad \text{for all } 0 \neq v \in Z.$$

   (b) *(inf–sup condition)* $\displaystyle\sup_{u \in U} \frac{b(p, u)}{\|u\|_U} \geq c\|p\|_P$.

3. *(Coercivity) The bilinear map* $A\big((u, p), (v, q)\big) = a(u, v) - b(p, v) + b(q, u)$, *is coercive on* $U \times P$: *there exists* $c > 0$ *such that*

$$\sup_{(v, q) \in U \times P} \frac{A\big((u, p), (v, q)\big)}{\|v\|_U + \|q\|_P} \geq c(\|u\|_U + \|p\|_P),$$

and

$$\sup_{(u, p) \in U \times P} A\big((u, p), (v, q)\big) > 0, \quad \text{for all } (0, 0) \neq (v, q) \in U \times P.$$

If $b : P \times U \to \mathbb{R}$ satisfies the inf–sup condition in (2) and Z is the kernel, then the restriction $b : P \times U/Z \to \mathbb{R}$ satisfies coercivity hypothesis [C] of the Lax Milgram Theorem. We make frequent use of the fact that, in this situation, this implies adjoint coercivity [C$'$] as well as existence for both the primal and dual problem.

2.2 Nonlinear saddle point problem

The next two theorems are a concise summary of Propositions 1.2, 2.3, 2.5, 2.6 and Theorems 1.2, 2.1 from Scheurer [39]. The structure of the nonlinear saddle problem is: find $(u, p) \in U \times P$ satisfying,

$$a(u, v) - b(p, v) = f(v),$$
$$b(q, u) = g(q),$$

for all $(v, q) \in U \times P$, where $a : U \times U \to \mathbb{R}$ is linear in the second argument and $b : P \times U \to \mathbb{R}$ is bilinear. The following theorem is the natural extension of the Corollary 2.1 for the linear problem to the situation where $a(., .)$ is nonlinear.

Theorem 2.2 *Let* U *and* P *be reflexive Banach spaces,* $b : P \times U \to \mathbb{R}$ *be continuous and bilinear, and let* $a : U \times U \to \mathbb{R}$ *be linear in its second argument. Let* $Z = \{u \in U \mid b(p, u) = 0, \ \forall p \in P\}$ *and* $c > 0$. *Then the following are equivalent.*

1. (*Existence of solutions*) *For all* $f \in U'$ *and* $g \in P'$ *there exists* $(u, p) \in U \times P$ *such that* (*a*)

$$a(u, v) - b(p, v) = f(v), \quad and \quad b(q, u) = g(q), \tag{2.1}$$

for all $(v, q) \in U \times P$, *and* (*b*) $\|u\|_{U/Z} \leq (1/c)\|g\|_{P'}$.

2. (a) *For all* $f \in U'$ *and* $u_g \in U$ *there exists* $u \in U$ *such that* $u - u_g \in Z$ *and*

$$a(u, v) = f(v), \quad v \in Z, \tag{2.2}$$

(b) (*inf–sup condition*)

$$\sup_{u \in U} \frac{b(p, u)}{\|u\|_U} \geq c\|p\|_P. \tag{2.3}$$

Remarks

1. Uniqueness of solutions will follow if, for example,

$$u_2 - u_1 \in Z \text{ and } a(u_2, v) - a(u_1, v) = 0, \ v \in Z \ \Rightarrow \ u_2 = u_1.$$

2. The above theorem is valid if $a : D(A) \times U \to \mathbb{R}$ has domain strictly contained in U provided the hypotheses in (1) and (2) require $u \in D(A)$.

Approximation properties of Galerkin schemes for nonlinear problems depend in a non-trivial fashion upon the structure of the nonlinearity [10,11,33]. However, linearity of the constraints can exploited to develop some useful formulae for the error, as shown in [39]. The following theorem (Theorem 2.1 of [39]) considers the simplest situation where the operator is (strictly) maximal monotone and Lipschitz continuous (in which case U is typically a Hilbert space).

Theorem 2.3 *Let U and P be reflexive Banach spaces, $b : P \times U \to \mathbb{R}$ be continuous and bilinear, and let $a : U \times U \to \mathbb{R}$ be linear in its second argument. Assume that $b(., .)$ satisfies the inf–sup condition: there exists $c_b > 0$ such that*

$$\sup_{v \in U} \frac{b(p, v)}{\|v\|_U} \geq c_b\|p\|_P.$$

Assume additionally that $a(., .)$ is strictly monotone and Lipschitz in its first argument:

$$c\|u - v\|_U^2 \leq a(u, u - v) - a(v, u - v), \quad and \quad a(u, w) - a(v, w) \leq C\|u - v\|_U \|w\|_U.$$

Let $U_h \subset U$ and $P_h \subset P$ be subspaces and suppose that $b : U_h \times P_h \to \mathbb{R}$ satisfies the inf–sup condition: there exists $c_b > 0$ (independent of h) such that

$$\sup_{v_h \in U_h} \frac{b(p_h, v_h)}{\|v_h\|_U} \geq c_b\|p_h\|_P.$$

Fix $(f, g) \in U' \times P'$ and suppose that $(u, p) \in U \times P$ satisfies

$$a(u, v) - b(p, v) = f(v), \quad b(q, u) = g(q), \quad (v, q) \in U \times P,$$

and suppose $(u_h, p_h) \in U_h \times P_h$ satisfies

$$a(u_h, v_h) - b(p_h, v_h) = f(v_h), \quad b(q_h, u_h) = g(q_h), \quad (v_h, q_h) \in U_h \times P_h.$$

Then there exists a constant $C > 0$ such that

$$\|u - u_h\|_U + \|p - p_h\|_P \leq C \left\{ \inf_{v_h \in U_h} \|u - v_h\|_U + \inf_{q_h \in P_h} \|p - q_h\|_P \right\}. \quad (2.4)$$

3 Twofold saddle point problems

In this section, the twofold saddle point problems (1.1) and (1.2) are considered, and conditions for existence and uniqueness of solutions are formulated. The main results of this section are Theorems 3.1 and 3.2, in which equivalent sets of inf–sup conditions for the twofold saddle point problems are given. As one set of conditions is often easier to verify than the others, particularly for the discrete spaces used in a specific numerical scheme, these results can simplify proofs of well-posedness for both the continuous and discrete twofold saddle point problems.

3.1 Problem type 1

In this section we consider the twofold saddle problem arising from (1.1): find $(u, p_1, p_2) \in U \times P_1 \times P_2$ such that

$$\begin{aligned}
a(u, v) - b_1(p_1, v) - b_2(p_2, v) &= f(v) \quad \forall v \in U, \\
b_1(q_1, u) &= g_1(q_1) \quad \forall q_1 \in P_1, \\
b_2(q_2, u) &= g_2(q_2) \quad \forall q_2 \in P_2.
\end{aligned} \quad (3.1)$$

This can be viewed as a single saddle point problem on $U \times (P_1 \times P_2)$ with bilinear form

$$b((p_1, p_2), u) = b_1(p_1, u) + b_2(p_2, u).$$

The following theorem provides several criteria which guarantee that this bilinear form satisfies the inf–sup condition. In Theorem 3.1, statement (1) gives the inf–sup condition necessary for the application of Theorems 2.2 and 2.3; in particular, it gives the lifting necessary for the approximation properties (2.4). Statements (2) and (3) are commonly used to show existence of solutions, their inclusion extends and unifies results in [30] and [20]. Statement (4), and the equivalence of these conditions, is new.

Theorem 3.1 *Let U, P_1, and P_2 be reflexive Banach spaces, and let $b : P_1 \times U \to \mathbb{R}$, and $b_2 : P_2 \times U \to \mathbb{R}$ be bilinear and continuous. Let*

$$Z_{b_i} = \{v \in U \mid b_i(q_i, v) = 0 \;\; \forall q_i \in P_i\} \subset U, \quad i = 1, 2,$$

then the following are equivalent:

(1) *There exists $c > 0$ such that*

$$\sup_{v \in U} \frac{b_1(p_1, v) + b_2(p_2, v)}{\|v\|_U} \geq c(\|p_1\|_{P_1} + \|p_2\|_{P_2}) \quad (p_1, p_2) \in P_1 \times P_2$$

(2) *There exists $c > 0$ such that*

$$\sup_{v \in U} \frac{b_1(p_1, v)}{\|v\|_U} \geq c\|p_1\|_{P_1}, \;\; p_1 \in P_1 \quad and \quad \sup_{v \in Z_{b_1}} \frac{b_2(p_2, v)}{\|v\|_U} \geq c\|p_2\|_{P_2}, \;\; p_2 \in P_2$$

(3) *There exists $c > 0$ such that*

$$\sup_{v \in Z_{b_2}} \frac{b_1(p_1, v)}{\|v\|_U} \geq c\|p_1\|_{P_1}, \;\; p_1 \in P_1 \quad and \quad \sup_{v \in U} \frac{b_2(p_2, v)}{\|v\|_U} \geq c\|p_2\|_{P_2}, \;\; p_2 \in P_2$$

(4) *There exists $c > 0$ such that*

$$\sup_{v \in Z_{b_2}} \frac{b_1(p_1, v)}{\|v\|_U} \geq c\|p_1\|_{P_1}, \;\; p_1 \in P_1 \quad and \quad \sup_{v \in Z_{b_1}} \frac{b_2(p_2, v)}{\|v\|_U} \geq c\|p_2\|_{P_2}, \;\; p_2 \in P_2.$$

Proof The result is shown by proving $(4) \Rightarrow (2) \Rightarrow (1) \Rightarrow (4)$ and $(4) \Rightarrow (3) \Rightarrow (1) \Rightarrow (4)$. The implications $(4) \Rightarrow (2)$ and $(4) \Rightarrow (3)$ are clear.

As shown in [20], to prove $(2) \Rightarrow (1)$ note that (2) implies the existence of $v_1 \in U$ with $\|v_1\|_U = 1$ and $b_1(q_1, v_1) \geq (c/2)\|q_1\|_{P_1}$, and the existence of $v_2 \in Z_{b_2}$ with $\|v_2\|_U = 1$ and $b_2(q_2, v_2) \geq (c/2)\|q_2\|_{P_2}$. Since $b_2(\cdot, \cdot)$ is continuous, there is a $C > 0$ such that $b_2(q_2, v) \leq C\|q_2\|_{P_2}\|v\|_U$ for all $(q_2, v) \in P_2 \times U$. Set $u = v_1 + (1 + 2C/c)v_2$. Then $\|u\|_U \leq 2(1 + 2C/c)$,

$$b_1(q_1, u) = b_1(q_1, v_1) \geq \frac{c}{2}\|q_1\|_{P_1},$$

and

$$b_2(q_2, u) = b_2(q_2, v_1) + \left(1 + \frac{2C}{c}\right) b_2(q_2, v_2)$$

$$\geq -C\|q_2\|_{P_2} + \left(1 + \frac{2C}{c}\right) \frac{c}{2}\|q_2\|_{P_2}$$

$$= \frac{c}{2}\|q_2\|_{P_2}.$$

Then

$$\frac{b_1(q_1, u) + b_2(q_2, u)}{\|u\|_U} \geq \frac{c}{4(1 + 2C/c)} \left(\|q_1\|_{P_1} + \|q_2\|_{P_2} \right)$$

proving (1). The proof that (3) $\Rightarrow$ (1) is similar.

To complete the equivalence, assume (1) and let

$$Z = Z_{b_1} \cap Z_{b_2} = \{u \in U \mid b_1(p_1, u) + b_2(p_2, u) = 0, \ (p_1, p_2) \in P_1 \times P_2\}.$$

The generalized Lax Milgram theorem then shows that for each $f \in (U/Z)'$ there exists a unique $(p_1, p_2) \in P_1 \times P_2$ satisfying

$$b_1(p_1, v) + b_2(p_2, v) = f(v), \qquad v \in U,$$

with $(\|p_1\|_{P_1} + \|p_2\|_{P_2}) \leq (1/c)\|f\|_{(U/Z)'}$. Here we identify $(U/Z)' \simeq \{f \in U' \mid f(u) = 0, \ u \in Z\}$.

It follows that for each $f_1 \in (Z_{b_1}/Z)'$ there exists $p_2 \in P_2$ satisfying $b_2(p_2, v) = f_1(v)$ for all $v \in Z_{b_1}$ with $\|p_2\|_{P_2} \leq (1/c)\|f_1\|_{(Z_{b_1}/Z)'}$. The Lax-Milgram condition then shows that b_2 is coercive on $P_2 \times Z_{b_1}/Z$ so b_2 satisfies the inf–sup condition stated in (4).

Similarly for each $f \in (U/Z)'$ there exists $p_1 \in P_1$ satisfying $b_1(p_1, v) = f(v)$ for all $v \in Z_{b_2}$ with $\|p_1\|_{P_1} \leq (1/c)\|f\|_{U'}$ so b_1 satisfies the inf–sup condition stated in (4).

Combining Theorem 3.1 and Theorem 2.2 gives necessary and sufficient conditions for twofold saddle problem (3.1).

Corollary 3.1 *Let U, P_1, P_2, b_1, b_2, Z_{b_1} and Z_{b_2} be as in Lemma 3.1, and let $Z = Z_{b_1} \cap Z_{b_2}$. Assume $a : U \times U \to \mathbb{R}$ is linear in its second argument. Assume that for all $f \in U'$ and $u_g \in U$ there exists $u_0 \in Z$ such that*

$$a(u_g + u_0, v) = f(v) \qquad v \in Z,$$

and that one of the conditions (1)–(4) of Theorem 3.1 are satisfied. Then for all $f \in U'$, $g_1 \in P_1'$, and $g_2 \in P_2'$ there exists $(u, p_1, p_2) \in U \times P_1 \times P_2$ satisfying (3.1) with $\|u\|_{U/Z} \leq C(\|g_1\|_{P_1'} + \|g_2\|_{P_2'})$.

Remark 3.1 Problems of the form

$$a(u, v) - \sum_{j=1}^{k} b_j(p_j, v) = f(v), \qquad \sum_{i=1}^{k} b_i(q_i, u) = \sum_{i=1}^{k} g_i(q_i),$$

for all $(v, q_1, \ldots, q_k) \in U \times P_1 \times \cdots \times P_k$ are considered in [20] and [30].

3.2 Problem type 2

We next consider the twofold saddle problems (1.2) and (1.3) corresponding to the variational problem $(u, p_1, p_2) \in U \times P_1 \times P_2$,

$$
\begin{aligned}
a(u, v) - b_1(p_2, v) &= f(v) \quad \forall v \in U, \\
- b_2(p_2, q_1) &= g_1(q_1) \quad \forall q_1 \in P_1, \\
b_1(q_2, u) + b_2(q_2, p_1) &= g_2(q_2) \quad \forall q_2 \in P_2.
\end{aligned}
\tag{3.2}
$$

As indicated in the introduction, two different single saddle problems result with different groupings of the variables. This gives rise to different, but equivalent, statements of the inf–sup conditions required for the existence of solutions. Different characterizations of the inf–sup condition are useful when constructing numerical schemes since one form is frequently easier to verify than the other.

Define

$$
\begin{aligned}
Z_{b_2} &= \{q_2 \in P_2 \mid b_2(q_2, p_1) = 0, \ p_1 \in P_1\} \quad \text{and} \\
\mathcal{D} &= \{u \in U \mid b_1(q_2, u) = 0, \ q_2 \in Z_{b_2}\}.
\end{aligned}
$$

Grouping the variables as $(U \times P_1) \times P_2$ gives rise to a single saddle problem with operator having block form $\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}$. Coercivity of this operator on $Z \subset U \times P_1$ will require $\|p_1\|_{P_1} \leq C\|u\|_U$ when $(u, p_1) \in Z$. Alternately stated, Z is the graph of the operator $B_2^{-1} B_1 : \mathcal{D} \subset U \to P_1$, and the inf–sup condition on b_2 guarantees that this operator is well-defined and continuous. This is the content of statement (1) of Theorem 3.2 below. These conditions were shown to be sufficient for solvability in [17].

With the grouping $(U \times P_2) \times P_1$ problem (3.2) has the form of the single saddle problem, the solution of which requires $b_2(\cdot, \cdot)$ to satisfy an inf–sup condition on P_2 and the operator defined by the block $\begin{bmatrix} A & -B_1^T \\ B_1 & 0 \end{bmatrix}$ must be coercive over $U \times Z_{b_2}$. This latter condition requires $b_1(., .)$ to satisfy an inf–sup condition on Z_{b_2}. These two inf–sup conditions were shown to be sufficient for solvability in [24] and correspond to statement (2) of the following theorem.

The first inf–sup condition in statement (3) was used in [19] to construct projection operators with optimal approximation properties from $U \times P_1$ to the discrete kernel Z_h in $U_h \times P_{1h}$, and the second condition provides a similar projection from P_2 to Z_{b_2h}.

While each of these conditions has appeared previously, equivalence of them has not. For example, separate proofs of (2) and (3) appear in [19] for a specific problem.

Theorem 3.2 *Let U, P_1, P_2 be reflexive Banach spaces, and $b_1 : P_2 \times U \to \mathbb{R}$ and $b_2 : P_2 \times P_1 \to \mathbb{R}$ be bilinear and continuous. Define the bilinear form $b : P_2 \times (U \times P_1) \to \mathbb{R}$ by $b(p_2, (u, p_1)) = b_1(p_2, u) + b_2(p_2, p_1)$, and let*

$$
Z_{b_2} = \{q_2 \in P_2 \mid b_2(q_2, p_1) = 0, \ p_1 \in P_1\},
$$

and

$$Z_b = \{(u, p_1) \in U \times P_1 \mid b(p_2, (u, p_1)) = 0, \; p_2 \in P_2\}.$$

Then the following are equivalent:

(1) *There exist constants C and $c > 0$ such that*

$$\sup_{(v,q_1)\in U \times P_1} \frac{b(p_2, (v, q_1))}{\|(v, q_1)\|_{U \times P_1}} \geq c\|p_2\|_{P_2} \quad and \quad \|q_1\|_{P_1} \leq C\|v\|_U \; for \; (v, q_1) \in Z_b.$$

(2) *There exists $c > 0$ such that*

$$\sup_{p_2 \in P_2} \frac{b_2(p_2, q_1)}{\|p_2\|_{P_2}} \geq c\|q_1\|_{P_1}, \qquad and \qquad \sup_{v \in U} \frac{b_1(p_2, v)}{\|v\|_U} \geq c\|p_2\|_{P_2}, \quad for \; p_2 \in Z_{b_2}.$$

(3) *There exists $c > 0$ such that*

$$\sup_{(v,q_1)\in U \times P_1} \frac{b(p_2, (v, q_1))}{\|(v, q_1)\|_{U \times P_1}} \geq c\|p_2\|_{P_2}, \qquad and \qquad \sup_{p_2 \in P_2} \frac{b_2(p_2, q_1)}{\|p_2\|_{P_2}} \geq c\|q_1\|_{P_1}.$$

Moreover, when one of these conditions holds Z_b is the graph of a linear function with domain

$$\mathcal{D} = \{u \in U \mid b_1(q_2, u) = 0, \; q_2 \in Z_{b_2}\} = \{u \mid (u, p_1) \in Z_b\}.$$

Remark The inclusion $\{u \mid (u, p_1) \in Z_b\} \subset \{u \in U \mid b_1(q_2, u) = 0, \; q_2 \in Z_{b_2}\}$ always holds; the reverse inclusion requires b_2 to satisfy an inf–sup condition.

Proof We show (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1).

(1) $\Rightarrow$ (2) Setting $p_2 \in Z_{b_2}$ in the inf–sup condition satisfied by $b(., .)$ immediately gives the inf–sup condition on $b_1(., .)$ stated in (2). To establish the inf–sup condition on $b_2(., .)$ use adjoint coercivity of $b : P_2 \times (U \times P_1)/Z_b$ to obtain

$$\sup_{p_2 \in P_2} \frac{b(p_2, (u, p_1))}{\|p_2\|_{P_2}} \geq c\|(u, p_1)\|_{(U \times P_1)/Z_b}.$$

Setting $u = 0$ and expanding the definition of the quotient norm shows

$$
\begin{aligned}
\sup_{p_2 \in P_2} \frac{b_2(p_2, p_1)}{\|p_2\|_{P_2}} &\geq c \inf_{(v,q_1)\in Z_b} \left(\|v\|_U + \|p_1 - q_1\|_{P_1} \right) \\
&\geq (c/C) \inf_{(v,q_1)\in Z} \left(\|q_1\|_{P_1} + \|p_1 - q_1\|_{P_1} \right) \\
&\geq (c/C)\|p_1\|_{P_1}.
\end{aligned}
$$

The second line follows from the property assumed upon elements of Z_b and it was assumed without loss of generality that $C \geq 1$.

(2) $\Rightarrow$ (3) We first show that for each $g_2 \in P_2'$ there exists a solution of $(v, q_1) \in U \times P_1$ of

$$b(q_2, (v, q_1)) \equiv b_1(q_2, v) + b_2(q_2, q_1) = g_2(q_2), \qquad q_2 \in P_2, \qquad (3.3)$$

with $c\|(v, q_1)\|_{U \times P_1} \le \|g_2\|_{P_2'}$. The inf–sup condition assumed on $b_1(., .)$ shows that its restriction to $b_1 : Z_{b_2} \times U/\mathcal{D} \to \mathbb{R}$ is coercive where

$$\mathcal{D} = \{u \in U \mid b_1(q_2, u) = 0, \ q_2 \in Z_{b_2}\}.$$

It follows that there exists $v \in U$ such that

$$b_1(q_2, v) = g_2(q_2) \qquad q_2 \in Z_{b_2},$$

with $c\|v\|_U \le \|g_2\|_{P_2'}$. Next, the inf–sup condition on $b_2(\cdot, \cdot)$ shows that $b_2 : P_2/Z_{b_2} \times P_1 \to \mathbb{R}$ is coercive. Since $g_2(q_2) - b_1(q_2, v)$ vanishes for $q_2 \in Z_{b_2}$ there exists a solution q_1 of

$$b_2(q_2, q_1) = g_2(q_2) - b_1(q_2, v) \qquad q_2 \in P_2,$$

with

$$c\|q_1\|_{P_1} \le \|g_2\| + C\|v\|_U \le (1 + C/c)\|g_2\|_{P_2'}.$$

This gives a solution of equation (3.3) with $c\|(v, q_1)\|_{U \times P_1} \le \|g_2\|_{P_2'}$.

To establish the inf–sup condition on b in (3), fix $p_2 \in P_2$ and let $g_2 \in P_2'$ satisfy $g_2(p_2) = \|g_1\|_{P'}^2 = \|p_2\|_{P_2}^2$. Then the solution of (3.3) satisfies $c\|(v, q_1)\|_{U \times P_1} \le \|g_2\|_{P_2'} = \|p_2\|_{P_2}$ and

$$\|p_2\|_{P_2} = \frac{g_2(p_2)}{\|p_2\|_{P_2}} = \frac{b(p_2, (v, q_1))}{\|p_2\|_{P_2}} \le \frac{b(p_2, (v, q_1))}{c\|(v, q_1)\|_{U \times P_1}}.$$

(3) $\Rightarrow$ (1) Adjoint coercivity of $b : P_2 \times (U \times P_1) \to \mathbb{R}$ shows

$$\sup_{p_2 \in P_2} \frac{b(p_2, (v, q_1))}{\|p_2\|_{P_2}} \ge c\|(v, q_1)\|_{(U \times P_1)/Z_b}.$$

If $(v, q_1) \in Z_b$ it follows that $b(p_2, (v, q_1)) = 0$ for all $p_2 \in P_2$; that is, $b_2(p_2, q_1) = -b_1(p_2, v)$ for all $p_2 \in P_2$. The coercivity assumed on $b_2(., .)$ then shows

$$c\|q_1\|_{P_1} \le \sup_{p_2 \in P_2} \frac{b_2(p_2, q_1)}{\|p_2\|_{P_2}} = \sup_{p_2 \in P_2} \frac{-b_1(p_2, v)}{\|p_2\|_{P_2}} \le C\|v\|_U.$$

It follows that problems of the form (3.2) can be treated as single saddle point problems and the application of Theorem 2.2 gives rise to necessary and sufficient conditions for existence of solutions.

Corollary 3.2 *Let U, P_1, P_2, b_1, b_2, $\mathcal{D}$ be as in Theorem 3.2 and assume $a : U \times U \to \mathbb{R}$ is linear in its second argument. Assume that for all $f \in U'$ and $u_g \in U$ there exists $u_0 \in \mathcal{D}$ such that*

$$a(u_g + u_0, v) = f(v) \qquad v \in \mathcal{D},$$

and that one of the conditions (1)–(3) *of Theorem 3.2 are satisfied. Then for all $f \in U'$, $g_1 \in P_1'$, and $g_2 \in P_2'$ there exists $(u, p_1, p_2) \in U \times P_1 \times P_2$ satisfying* (3.2) *with*

$$\|u\|_{U/\mathcal{D}} \leq C \left(\|g_1\|_{P_1'} + \|g_2\|_{P_2'} \right).$$

4 Application to problem (1.4)–(1.5)

In this section the saddle point theory developed above is used to solve the nonlinear Stokes problem (1.4)–(1.5) and to formulate numerical schemes to approximate the solution. Several finite element schemes have been proposed to approximate solutions of (1.4)–(1.5) which approximate the stress S directly. All of the schemes proposed to date either need to invert the constitutive operator $\mathcal{A}$ [9,23,37], or approximate the stress using composite elements [34], or relax the symmetry of the stress tensor [19,21–23,27,35]. The method presented in this section is the first to approximate solutions of (1.4)–(1.5) using noncomposite symmetric finite elements for the stress without inverting the constitutive operator, and is also a first use of the symmetric Arnold–Winther tensors [7] for fluid problems. The main results of this section are Theorem 4.2, which gives the approximation properties of the new scheme, and Lemma 4.3, which gives an example of discrete spaces that satisfy the inf–sup conditions required for solvability of the discrete problem.

Assume $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, is bounded and simply connected and $\partial\Omega = \Gamma_0 \cup \Gamma_1$ is Lipschitz-continuous and the boundary partition is sufficiently regular for classical $H^2(\Omega)$ regularity to apply. Let

$$\mathbb{S} = \left\{ S \in \mathbb{H}(\Omega; \mathrm{div})_{sym} \mid S\mathbf{n}|_{\Gamma_1} = 0 \right\}, \qquad U = L^2(\Omega)^d,$$
$$\mathbb{D} = \left\{ D \in L^2(\Omega)_{sym}^{d \times d} \mid tr(D) = 0 \right\}.$$

For simplicity we assume a homogeneous traction boundary condition on Γ_1; non-homogeneous boundary data can be accommodated with the usual translation argument. The variational problem is: given $f \in L^2(\Omega)^d$ and $\mathbf{u}_0 \in H^{1/2}(\Gamma_0)^d$, find $(D, \mathbf{u}, S) \in \mathbb{D} \times U \times \mathbb{S}$ satisfying

$$\int_\Omega \mathcal{A}(D) : E - S : E = 0, \qquad E \in \mathbb{D},$$

$$\int_\Omega -\mathrm{div}(S).\mathbf{v} = \int_\Omega f.\mathbf{v}. \qquad \mathbf{v} \in U. \tag{4.1}$$

$$\int_\Omega (D : T + \mathbf{u}.\mathrm{div}(T)) = \int_{\Gamma_0} \mathbf{u}_0.T\mathbf{n}, \qquad T \in \mathbb{S},$$

This problem takes the form of the twofold saddle problem considered in Sect. 3.2 with operators

$$a(D, E) = \int_\Omega \mathcal{A}(D) : E, \quad b_1(S, E) = \int_\Omega S : E, \quad b_2(S, \mathbf{v}) = \int_\Omega div(S).\mathbf{v}.$$

Below we will assume $|\Gamma_1| \neq 0$, otherwise it is necessary to require $S \in \mathbb{S}$ to satisfy $\int_\Omega tr(S) = 0$.

4.1 Solution of the continuous problem

Corollary 3.2 will be used to solve problem (4.1) using condition (2) of Theorem 3.2 to establish the inf–sup hypotheses.

First inf–sup condition The inf–sup condition on $b_2(., .)$ becomes

$$\sup_{S \in \mathbb{S}} \frac{\int_\Omega \mathrm{div}(S).\mathbf{v}}{\|S\|_{\mathbb{H}(\Omega;\mathrm{div})}} \geq c\|\mathbf{v}\|_{L^2(\Omega)}, \quad \forall \mathbf{v} \in U. \tag{4.2}$$

A proof of this inequality may be found in [12, §11.1, 11.2]. Briefly, given $\mathbf{v} \in U$ select $S = D(\mathbf{u}) \equiv (1/2)(\nabla\mathbf{u} + (\mathbf{u})^T)$ where $\mathbf{u}$ is the solution of the linear elasticity problem,

$$div(D(\mathbf{u})) = \mathbf{v}, \qquad \mathbf{u}|_{\Gamma_0} = 0, \qquad D(\mathbf{u})\mathbf{n}|_{\Gamma_1} = \mathbf{0}.$$

Then $\int_\Omega div(S).\mathbf{v} = \|\mathbf{v}\|_{L^2(\Omega)}^2$, and the Poincaré and Korn inequalities may be used to verify that there exists $C > 0$ such that $\|S\|_{\mathbb{H}(\Omega;\mathrm{div})} \leq C\|\mathbf{v}\|_{L^2(\Omega)}$, and (4.2) follows. For the discrete problem we use the property that regularity theory for the linear elastic problem shows that if $\mathbf{v} \in L^2(\Omega)^d$ then $\mathbf{u} \in H^2(\Omega)^d$ so $S = D(\mathbf{u}) \in H^1(\Omega)^{d \times d}$.

Second inf–sup condition Setting $Z_{b_2} = \{S \in \mathbb{S} \mid divS = 0\}$, the inf–sup condition required of $b_1(., .)$ becomes

$$\sup_{D \in \mathbb{D}} \frac{\int_\Omega D : S}{\|D\|_{L^2(\Omega)}} \geq c\|S\|_{L^2(\Omega)}, \quad \forall S \in Z_{b_2}. \tag{4.3}$$

Given $S \in Z_{b_2}$ the inf–sup condition is established upon setting $D = S - (1/d)tr(S)I \in \mathbb{D}$ to obtain $\int_\Omega D : S = \|D\|_{L^2(\Omega)}^2$. Lemma 5.2 in the appendix shows that there is a constant $C > 0$ such that $\|S\|_{L^2(\Omega)} \leq C\|D\|_{L^2(\Omega)}$ when $S \in Z_{b_2}$ and (4.3) follows.

Define the set

$$\mathcal{D} = \left\{ D \in \mathbb{D} \mid \int_{\Omega} D : S = 0 \ \ \forall S \in Z_{b_2} \right\} \subset \mathbb{D}.$$

This set is not easily characterized so minimal conditions on $\mathcal{A} : \mathbb{R}^{d \times d} \to \mathbb{R}^{d \times d}$ to guarantee coercivity over $\mathcal{D}$ are not available. Fortunately $\mathcal{A}$ is typically a maximal monotone operator on all of $\mathbb{R}^{d \times d}$ so this is not an issue. Granted this, Theorem 3.2 establishes existence of solutions to problem (4.1).

Theorem 4.1 *Let $\Omega \subset \mathbb{R}^d$ and the spaces $\mathbb{S}$, U, and $\mathbb{D}$ be given at the beginning of this section. Assume $\mathcal{A} : \mathbb{R}^{d \times d} \to \mathbb{R}^{d \times d}$ maps bounded sets to bounded sets, and satisfies*

$$(\mathcal{A}(E) - \mathcal{A}(D)) : (E - D) \geq 0 \qquad \text{and} \qquad \lim_{|D| \to \infty} \mathcal{A}(D) : D/|D| \to \infty.$$

Then for all $f \in L^2(\Omega)^d$ and $\mathbf{u}_0 \in H^{1/2}(\Gamma_0)^d$, problem (4.1) has a solution $(D, \mathbf{u}, S) \in \mathbb{D} \times U \times \mathbb{S}$ satisfying

$$\|D\|_{\mathbb{D}/\mathcal{D}} \leq C \left(\|\mathbf{f}\|_{L^2(\Omega)} + \|\mathbf{u}_0\|_{H^{1/2}(\Gamma_0)} \right),$$

where

$$\mathcal{D} = \left\{ D \in \mathbb{D} \mid \int_{\Omega} D : S = 0, \ \forall S \in \mathbb{S} \text{ with } div(S) = 0 \right\}.$$

Moreover, if $\mathcal{A}$ is strictly monotone,

$$(\mathcal{A}(E) - \mathcal{A}(D)) : (E - D) > 0, \qquad E \neq D,$$

the solution is unique.

4.2 Galerkin approximation

Let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of Ω and let $\mathbb{D}_h$, U_h, $\mathbb{S}_h$ be finite dimensional subspaces of $\mathbb{D}$, U, $\mathbb{S}$, respectively. We consider Galerkin approximations of problem (4.1) satisfying $(D_h, \mathbf{u}_h, S_h) \in \mathbb{D}_h \times U_h \times \mathbb{S}_h$,

$$\int_{\Omega} \mathcal{A}(D_h) : E_h - S_h : E_h = 0, \qquad E_h \in \mathbb{D}_h,$$

$$\int_{\Omega} -\text{div}(S_h) . \mathbf{v}_h = \int_{\Omega} f . \mathbf{v}_h, \qquad \mathbf{v}_h \in U_h, \tag{4.4}$$

$$\int_{\Omega} (D_h : T_h + \mathbf{u}_h.\mathrm{div}(T_h)) = \int_{\Gamma_0} \mathbf{u}_0.T_h\mathbf{n}, \qquad T_h \in \mathbb{S}_h,$$

In this context establishing discrete versions of (4.2) and (4.3) can be considerably less complicated and technical than showing a condition of the form

$$\sup_{(D_h,\mathbf{u}_h)\in\mathbb{D}_h\times U_h} \frac{\int_{\Omega} (D_h : T_h + \mathbf{u}_h.\mathrm{div}(T_h))}{\|D_h\|_{\mathbb{D}} + \|u_h\|_U} \geq c\|T_h\|_{\mathbb{H}(\Omega;\mathrm{div})},$$

since the discrete form of (4.3) only needs to be verified for divergence-free T_h. This illustrates the utility of the equivalent inf–sup conditions given in Theorem 3.2.

First inf–sup condition The issue of finding conforming finite elements for symmetric tensors satisfying an inf–sup condition of the form (4.2) is well-documented [4,5,14,38]. We consider the finite element pairs $(\mathbb{S}_h, U_h)$ of symmetric tensors and vectors constructed by Arnold et al. [1,7] which satisfy the inf–sup condition.

Let $k \geq 1$ and define $\mathbb{P}_k(K)$ to be the set of all polynomials of degree at most k on the simplex $K \in \mathcal{T}_h$. On K, define $\mathbb{S}_K$ to be the symmetric Arnold–Winther tensors

$$\mathbb{S}_K = \left\{ S_h \in \mathbb{H}(\mathrm{div}, K)_{sym} \mid S_h \in \mathbb{P}_{k+d}(K)^{d\times d}_{sym} \text{ and } \mathrm{div}(S_h) \in \mathbb{P}_k(K)^d \right\}.$$

The space $\mathbb{S}_h$ is the union of $\mathbb{S}_K$ over all $K \in \mathcal{T}_h$, subject to the condition that the normal components are continuous across mesh edges (faces for $d = 3$) and all components are continuous at vertices. Define

$$U_h = \left\{ \mathbf{u} \in L^2(\Omega)^d \mid \mathbf{u}|_K \in \mathbb{P}_k(K)^d \quad \forall K \in \mathcal{T}_h \right\}$$

and note that there is no interelement continuity requirement for U_h. Figure 1 gives a diagram of the degrees of freedom on each triangle for the lowest order Arnold–Winther (S_h, u_h) pair $(k = 1)$ in two dimensions. In [1,7] an interpolation operator $\Pi_h : \mathbb{S} \cap H^1(\Omega)^{d\times d} \to \mathbb{S}_h$ is constructed for which

$$\mathrm{div}(\Pi_h S) = \mathcal{P}_h^U \mathrm{div}(S), \qquad
\begin{array}{ccc}
\mathbb{S} \cap H^1(\Omega)^{d\times d} & \xrightarrow{\ div\ } & U \\
\Pi_h \downarrow & & \downarrow P_h^U \\
\mathbb{S}_h & \xrightarrow{\ div\ } & U_h
\end{array}$$

where $\mathcal{P}_h^U : U \to U_h$ is the orthogonal projection.

The discrete analog of (4.2) follows since for any $\mathbf{v}_h \in U_h$ there is a $S \in \mathbb{S} \cap H^1(\Omega)^{d\times d}$ satisfying $div(S) = \mathbf{v}_h$ and $\|S\|_{L^2(\Omega)} \leq C\|\mathbf{v}_h\|_{L^2(\Omega)}$. Then $\mathrm{div}(\Pi_h S) = \mathcal{P}_h^U \mathrm{div}(S) = \mathbf{v}_h$ and $\|\Pi_h S\|_{\mathbb{H}(\Omega;\mathrm{div})} \leq C\|\mathbf{v}_h\|_U$ and the discrete inf–sup condition follows.
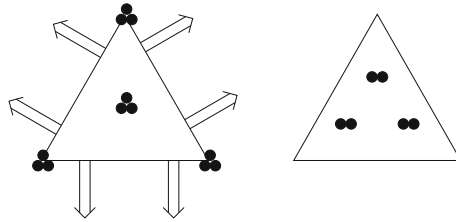
Fig. 1 Lowest-order Arnold–Winther elements in two dimensions for S_h (*left*) and u_h (*right*). For S_h, the points represent values of the components of S_h (vertices) and the value of the three components of the moment of degree 0 of S_h (interior), and the *arrows* represent the values of the moments of degree 0 and 1 of the two normal components of S_h on each edge. For u_h, the points represent the value of the two components at the three interior nodes

Second inf–sup condition The commuting diagram for the Arnold–Winther spaces shows

$$Z_h \equiv \left\{ S_h \in \mathbb{S}_h \mid \int_\Omega \mathrm{div}(S_h).\mathbf{u}_h = 0, \ \ \mathbf{u}_h \in U_h \right\} = \{ S_h \in \mathbb{S}_h \mid \mathrm{div}(S_h) = \mathbf{0} \}$$

The second inf–sup condition for the discrete spaces becomes

$$\sup_{D_h \in \mathbb{D}_h} \frac{\int_\Omega D_h : S_h}{\|D_h\|_{L^2(\Omega)}} \geq c \|S_h\|_{L^2(\Omega)}, \qquad S_h \in Z_h. \tag{4.5}$$

When $div(S) = 0$, Lemma 5.2 of the appendix shows that $\|S\|_{L^2(\Omega)} \leq C \|S_0\|_{L^2(\Omega)}$ where $S_0 = S - (tr(S)/d)I$ is the trace free part of S. Since $S_h \in \mathbb{S}_h$ is piecewise polynomial of degree $k + d$ on $\mathcal{T}_h$, so it suffices to let

$$\mathbb{D}_h = \left\{ D_h \in L^2(\Omega)^{d \times d} \mid tr(D_h) = 0 \text{ and } D_h|_K \in \mathbb{P}_{k+d}(K)^{d \times d}, \ \ K \in \mathcal{T}_h \right\};$$

however, typically much smaller spaces (and hence cheaper numerical schemes) suffice. Note too that smaller spaces will not necessarily result in a loss of accuracy; the theory only requires the identity operator from $\mathbb{D}_h$ to the space of divergence-free Arnold–Winther tensors to be surjective.

For smooth functions the Arnold–Winther spaces exhibit the following approximation properties,

$$\|S - \Pi_h S\|_{\mathbb{H}(\Omega;\mathrm{div})} \leq Ch^{k+1}, \qquad \|\mathbf{u} - \mathcal{P}_h^U \mathbf{u}\|_{L^2(\Omega)} \leq Ch^{k+1},$$

and this rate would be achieved for $\|D_h - D\|_{L^2(\Omega)}$ if $\mathbb{D}_h$ contained the piecewise polynomials of degree k.

Having constructed discrete subspaces which satisfy these inf–sup conditions, Theorem 2.3 provides error estimates for the Galerkin approximations.

Theorem 4.2 *Let $\Omega \subset \mathbb{R}^d$ and the spaces $\mathbb{S}$, U, and $\mathbb{D}$ be given at the beginning of this section. Assume there exist constants $C, c > 0$ such that $\mathcal{A} : \mathbb{R}^{d \times d} \to \mathbb{R}^{d \times d}$ satisfies*

$$c|D - E|^2 \le (\mathcal{A}(E) - \mathcal{A}(D)) : (E - D) \quad and \quad (\mathcal{A}(E) - \mathcal{A}(D)) : F \le |E - D| \, |F|.$$

Let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of Ω, $\mathbb{S}_h \subset \mathbb{S}$ be the Arnold–Winther space of index k over $\mathcal{T}_h$, $U_h \subset U$ be the discontinuous piecewise polynomial space of degree k, and $\mathbb{D}_h \subset \mathbb{D}$ be the discontinuous piecewise polynomial space of degree $k + d$.

Let $(D, \mathbf{u}, S)$ satisfy (4.1) and $(D_h, \mathbf{u}_h, S_h)$ satisfy (4.4). Then there is a constant $c > 0$ such that, for $1 \le m \le k + 1$,

$$\|D - D_h\|_{\mathbb{D}} + \|\mathbf{u} - \mathbf{u}_h\|_U + \|S - S_h\|_{\mathbb{S}} \le Ch^m \{\|D\|_m + \|\mathbf{u}\|_m + \|S\|_m + \|\mathrm{div}(S)\|_m\}.$$

4.3 Finite element subspaces for $\mathbb{D}$

In this section a macroelement construction [40] is employed to determine subspaces $\mathbb{D}_h \subset \mathbb{D}$ that will satisfy the inf–sup condition (4.5) and are smaller than naive ones considered in Theorem 4.2.

4.3.1 Macroelement construction

We employ the concept of a macroelement from [40].

Definition 4.1 (1) A parent macroelement $\hat{M}$ is a connected finite union of simplices in $\mathbb{R}^d$.

(2) A macroelement M in $\mathcal{T}_h$ affine equivalent to $\hat{M}$ is a set of simplices of $\mathcal{T}_h$ whose union is homeomorphic to $\hat{M}$ by a homeomorphism which is piecewise affine on each simplex $K \subset M$.

The following lemma relates local inf–sup conditions on macroelements to the corresponding condition on a parent macroelement

Lemma 4.1 *Let M and $\hat{M}$ be affine equivalent macroelements and let $\chi : \hat{M} \to M$ be the corresponding piecewise affine homeomorphism. If $\hat{K} \subset \hat{M}$ is a simplex and $\chi(\hat{K}) = K$, write*

$$\chi(\xi)|_{\hat{K}} = x_{0K} + F_K \xi, \quad \xi \in \hat{K},$$

where $F_K \in \mathbb{R}^{d \times d}$, and let $J = det(F)$. Let $\hat{\mathbb{S}}(\hat{M}), \hat{\mathbb{D}}(\hat{M}) \subset L^2(\hat{M})^{d \times d}$ and define the mappings $\hat{\,} : \hat{\mathbb{S}}(\hat{M}) \to \mathbb{S}_h(M) \subset L^2(M)^{d \times d}$ and $\hat{\,} : \hat{\mathbb{D}}(\hat{M}) \to \mathbb{D}_h(M) \subset L^2(M)^{d \times d}$ by

$$S_h = (1/\sqrt{J}) F \hat{S} F^T \quad and \quad D_h = (1/\sqrt{J}) F^{-T} \hat{D} F^{-1}. \tag{4.6}$$

Under these transformations,

1. *Symmetric matrices on $\hat{M}$ are mapped to symmetric matrices on M.*
2. *Piecewise polynomials of degree k on $\hat{M}$ map to piecewise polynomials of degree k on M.*
3. *If all of the Jacobian matrices F_K are equal to a scalar multiple of an orthogonal matrix then trace free functions on $\hat{M}$ to trace free functions on M.*
4. *The mapping $\hat{\mathbb{S}}(\hat{M}) \mapsto \mathbb{S}_h(M)$ maps functions in $\mathbb{H}(div, \hat{M})$ to functions in $\mathbb{H}(div, M)$; moreover, divergence free functions on $\hat{M}$ map to divergence free functions on M. In particular, the (divergence free) Arnold–Winther space on $\hat{M}$ is mapped to the (divergence free) Arnold–Winther space on M.*
5. *For each $\hat{K} \subset \hat{M}$ and $K = \chi(\hat{K})$*

$$\|S_h\|_{L^2(K)} \leq \|F_K\|^2 \|\hat{S}\|_{L^2(\hat{K})}, \qquad \|D_h\|_{L^2(K)} \leq \|F_K^{-1}\|^2 \|\hat{D}\|_{L^2(\hat{K})},$$

   and

$$\int_{K_i} S_h : D_h = \int_{\hat{K}_i} \hat{S} : \hat{D}.$$

6. *Define the "condition number" to be $\kappa_M = \|F\|_{L^\infty(\hat{M})} \|F^{-1}\|_{L^\infty(\hat{M})}$. If $\hat{S} \in \hat{\mathbb{S}}\backslash\{0\}$, then*

$$\sup_{D_h \in \mathbb{D}_h(M)} \frac{\int_M S_h : D_h}{\|D_h\|_{L^2(M)} \|S_h\|_{L^2(M)}} \geq \frac{1}{\kappa_M^2} \sup_{\hat{D} \in \hat{\mathbb{D}}(\hat{M})} \frac{\int_{\hat{M}} (\hat{S}, \hat{D})}{\|\hat{D}\|_{L^2(\hat{M})} \|\hat{S}\|_{L^2(\hat{M})}}. \quad (4.7)$$

The mapping $\hat{S} \mapsto S_h$ is the symmetric version of the Piola transform [14,36] and the transformation $\hat{D} \mapsto D_h$ was constructed to be the dual operator preserving the inner product.

The next step is to show that a global inf–sup condition in $\mathcal{T}_h$ will follow from the local inf–sup condition on the right hand side of (4.7).

Theorem 4.3 *Let $\hat{\mathcal{M}}$ be a family of parent macroelements and $\Omega = \cup_H M_H$ be a covering of Ω by macroelements of $\mathcal{T}_h$ each affine equivalent to some $\hat{M} \in \hat{\mathcal{M}}$. For each $\hat{M} \in \hat{\mathcal{M}}$ let $\hat{\mathbb{D}}(\hat{M}), \hat{\mathbb{S}}(\hat{M}) \subset L^2(\hat{M})^{d\times d}$ be specified and assume that there exists $\hat{c} > 0$*

$$\sup_{\hat{D} \in \hat{\mathbb{D}}(\hat{M})} \frac{\int_{\hat{M}} (\hat{S}, \hat{D})}{\|\hat{D}\|_{L^2(\hat{M})}} \geq \hat{c} \|\hat{S}\|_{L^2(\hat{M})}, \qquad \hat{S} \in \hat{\mathbb{S}}(\hat{M}), \quad \hat{M} \in \hat{\mathcal{M}}.$$

Let $\mathbb{S}_h \subset L^2(\Omega)^{d\times d}$ and assume that the local spaces $\mathbb{S}_h(M)$ formed by restriction to M are the images of $\hat{\mathbb{S}}(\hat{M})$ under the (symmetric) Piola transform (4.6). If $\mathbb{D}_h$ is the (discontinuous) finite element space on $\mathcal{T}_h$ spanned by the local spaces $\mathbb{D}_h(M)$, then

$$\sup_{D_h \in \mathbb{D}_h} \frac{\int_\Omega S_h : D_h}{\|D_h\|_{L^2(\Omega)}} \geq \frac{\hat{c}}{\gamma\kappa^2} \|S_h\|_{L^2(\Omega)},$$

where $\kappa = \max_M \kappa_M$ is the maximal condition number on each macroelement and γ is the ply of the covering $\Omega = \cup_H M_H$; that is, $x \in \Omega$ belongs to at most γ macroelements.

Proof Fix $S_h \in \mathbb{S}_h$ and for each H let $D_H \in \mathbb{D}_h(M_H)$ satisfy $\|D_H\|_{L^2(M_H)} = \|S_h\|_{L^2(M_H)}$ and

$$\int_{M_H} D_H : S_h \geq (\hat{c}/\kappa^2) \|S_h\|_{L^2(M_H)}^2.$$

Then define $D_h = \sum_H D_H$, and use the property that $\{M_H\}$ cover Ω to conclude

$$\int_{\Omega} D_h : S_h \geq (\hat{c}/\kappa^2) \|S_h\|_{L^2(\Omega)}^2.$$

Next, since the covering has finite ply we compute

$$\|D_h\|_{L^2(\Omega)}^2 = \sum_{K \in \mathcal{T}_h} \int_K \left| \sum_H D_H \right|^2$$

$$\leq \sum_{K \in \mathcal{T}_h} \int_K \gamma \sum_H |D_H|^2$$

$$= \gamma \sum_{K \in \mathcal{T}_h} \sum_H \int_{K \cap M_H} |D_H|^2$$

$$= \gamma \sum_H \sum_{K \subset \mathcal{T}_h} \int_{K \cap M_H} |D_H|^2$$

$$= \gamma \sum_H \int_{M_H} |D_H|^2$$

$$= \gamma \sum_H \int_{M_H} |S_h|^2$$

$$\leq \gamma^2 \int_{\Omega} |S_h|^2.$$

It follows that $\|D_h\|_{L^2(\Omega)} \leq \gamma \|S_h\|_{L^2(\Omega)}$, and

$$\frac{\int_{\Omega} D_h : S_h}{\|D_h\|_{L^2(\Omega)}} \geq (\hat{c}/\gamma\kappa^2) \|S_h\|_{L^2(\Omega)}.$$

In practice it is trivial to construct a covering of Ω by macroelements with finite ply, and a bound upon the condition number follows from the bound upon the aspect

ratio of the simplices in $\mathcal{T}_h$. In this situation the global inf–sup condition will follow provided the local inf–sup condition holds on each parent macroelement with constant independent of $\hat{M} \in \hat{\mathcal{M}}$. The following lemma shows that verification of the inf–sup condition on each macroelement reduces to linear algebra.

Lemma 4.2 *Let $\hat{M}$ be a parent macroelement and suppose $\hat{\mathbb{D}}, \hat{\mathbb{S}} \subset L^2(\hat{M})_{sym}^{d \times d}$ are finite dimensional subspaces and that for all $\hat{S} \in \hat{\mathbb{S}}$*

$$\int_{\hat{M}} \hat{D} : \hat{S} = 0 \quad \forall \hat{D} \in \hat{\mathbb{D}} \quad \Rightarrow \quad \hat{S} = 0.$$

Then there exists $\hat{c} > 0$ such that

$$\sup_{\hat{D} \in \hat{\mathbb{D}}} \frac{\int_{\hat{M}} \hat{D} : \hat{S}}{\|\hat{D}\|_{L^2(\hat{M})}} \geq \hat{c} \|\hat{S}\|_{L^2(\hat{M})}, \quad \hat{S} \in \hat{\mathbb{S}}. \tag{4.8}$$

Proof Let the linear map $\phi : \hat{\mathbb{D}} \to \hat{\mathbb{S}}$ be characterized by $(\phi(\hat{D}), \hat{S})_{L^2} = (\hat{D}, \hat{S})_{L^2}$ for all $\hat{S}$. The hypothesis of the theorem shows $Rg(\phi)^{\perp} = \{0\}$, and finite dimensionality guarantees $Rg(\phi)$ is closed so is all of $\hat{S}$. Then ϕ is surjective and is an isomorphism on $Ker(\phi)^{\perp}$. Selecting $\hat{D} = \phi^{-1}(\hat{S}) \in Ker(\phi)^{\perp}$ establishes the inf–sup condition (4.8). $\qquad \square$

If $\hat{\mathcal{M}}$ is finite the inf–sup condition then reduces to linear algebra. When $\hat{\mathcal{M}}$ is infinite the macroelements can usually be parameterized by a compact set; for example, the vertices if the simplices on $\hat{M}$ typically lie in a compact set of $\mathbb{R}^d$. If the constants $\hat{c} = \hat{c}(\hat{M}) > 0$ depend continuously upon the parameters a uniform bound from below will follow (positive functions on compact sets achieve their minimum).

4.3.2 Two dimensional example

In this section the macroelement technique is used to show that in two dimensions the inf–sup condition (4.5) holds for the lowest order Arnold–Winther elements, $k = 1$, when $\mathbb{D}_h$ is the space of discontinuous "quadratic plus bubble" symmetric trace free matrices. For this purpose we select the macroelements to be pairs of triangles which have an edge in common, and the reference macro elements be pairs of simplices with common edge lying on the unit interval of the x-axis (see Fig.2),

$$\hat{\mathcal{M}} = \{\hat{M} = \hat{K}_1 \cup \hat{K}_2 \mid \hat{K}_1 \cap \hat{K}_2 = [0, 1] \times \{0\},$$
$$\text{and } \hat{K}_1, \hat{K}_2 \text{ have aspect ratio bounded by } C\}.$$

If $\{\mathcal{T}_h\}$ is a regular family of triangulations of Ω then each pair of triangles $M = K_1 \cup K_2$ sharing an edge of length h_M can be mapped to $\hat{M} \in \hat{\mathcal{M}}$ with an affine homeomorphism of the form,

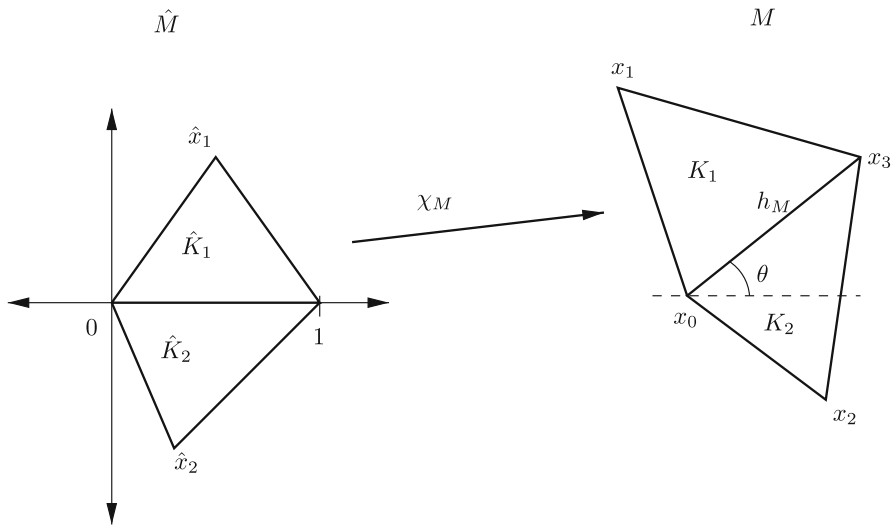$$x = \chi(\hat{x}) = x_0 + h_M Q \hat{x}, \qquad Q^T Q = I.$$

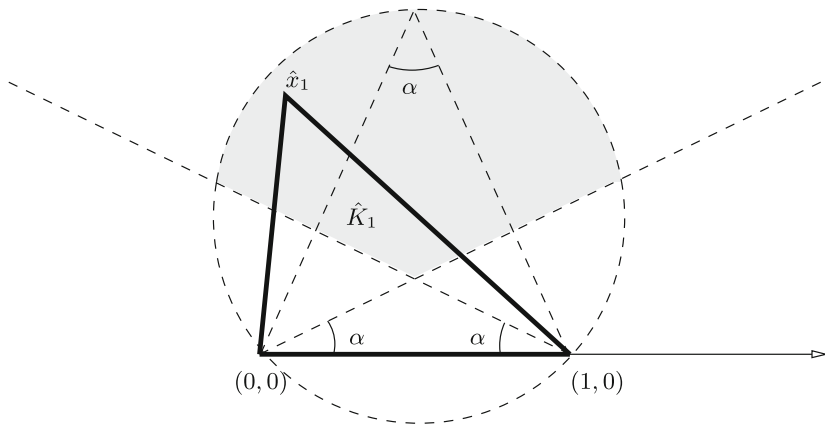**Fig. 2** Reference $\hat{M}$ and affine equivalent M macroelements

Fig. 3 If $\hat{K}_1$ has minimum angle α then $\hat{x}_1$ lies in a compact subset of $\mathbb{R}^2$

Selecting the Jacobians to be orthogonal guarantees that trace free matrices are mapped to trace free matrices under the Piola transformations (4.6), and that condition numbers $\kappa_M = 1$ are trivially bounded.

It remains to verify the inf–sup condition on each macroelement $\hat{M} \in \hat{\mathcal{M}}$. To do this we parameterize $\hat{M} = \hat{K}_1 \cup \hat{K}_2 \in \hat{\mathcal{M}}$ by the coordinates of the vertices $\hat{x}_1$ and $\hat{x}_2$ of $\hat{K}_1$ and $\hat{K}_2$ not on the $\hat{x}_1$-axis. Since each simplex has bounded aspect ratio it follows that the pair $(\hat{x}_1, \hat{x}_2)$ lie in a compact subset of $\mathbb{R}^2$ as indicated in Fig. 3. Moreover, since the inf–sup condition only involves integrals over $\hat{M} = \hat{M}(\hat{x}_1, \hat{x}_2)$ it is clear that the constant depends continuously upon the two parameters. In this situation it suffices to establish the inf–sup condition on each $\hat{M}(\hat{x}_1, \hat{x}_2)$ using, for example, Lemma 4.2.

The linear algebra required to complete the proof of the inf–sup condition involves large matrices with symbolic components, $(\hat{x}_1, \hat{x}_2)$. Maple was used to verify the hypothesis in Lemma 4.2 as follows:

1. On each of $\hat{K}_1$ and $\hat{K}_2$ the divergence free functions in the lowest order Arnold–Winther space take the form

$$
\begin{aligned}
\hat{S}_{11} &= 1/2\, a_0 x^2 + 1/6\, a_1 x^3 + 1/2\, a_2 y x^2 \\
&\quad + \left(a_3 + a_4 y + a_5 y^2 \right) x + a_6 + a_7 y + a_8 y^2 + a_9 y^3 \\
\hat{S}_{12} &= -a_0 x y - 1/2\, a_1 x^2 y - 1/2\, a_2 y^2 x - a_3 y - 1/2\, a_4 y^2 - 1/3\, a_5 y^3 \\
&\quad - a_{10} x - 1/2\, a_{11} x^2 - 1/3\, a_{12} x^3 + a_{17} \\
\hat{S}_{22} &= 1/2\, a_0 y^2 + 1/2\, a_1 x y^2 + 1/6\, a_2 y^3 \\
&\quad + \left(a_{10} + a_{11} x + a_{12} x^2 \right) y + a_{13} + a_{14} x + a_{15} x^2 + a_{16} x^3
\end{aligned}
$$

2. Accumulate the set of equations to enforce the following:
   (a) $\hat{S}$ is continuous at the vertices $(0, 0)$ and $(1, 0)$,

   $$
   \hat{S}(0, 0_-) = \hat{S}(0, 0_+), \quad \text{and} \quad \hat{S}(1, 0_-) = \hat{S}(1, 0_+).
   $$

   (b) $\hat{S}\mathbf{n}$ is continuous across the x-axis where $\mathbf{n} = (0, 1)^T$. Along the x-axis the normal components are univariate polynomials of degree three which agree at the end points, so it suffices to require continuity at two more points.

   $$
   \hat{S}(1/3, 0_-)\mathbf{n} = \hat{S}(1/3, 0_+)\mathbf{n}, \quad \text{and} \quad \hat{S}(2/3, 0_-)\mathbf{n} = \hat{S}(2/3, 0_+)\mathbf{n}.
   $$

   (c) On each of the triangles, K_i, $\hat{S}$ is orthogonal to a basis of the trace free matrices with components in $\mathcal{P}_2(\hat{K}_i) \oplus b_{\hat{K}_i}$.

   $$
   \int_{\hat{K}_i} \left(\hat{S}_{11} - \hat{S}_{22} \right) p = 0, \quad \int_{\hat{K}_i} \hat{S}_{12} p = 0, \quad p \in \{1, x, y, x^2, xy, y^2, b_{K_i}(x, y)\}.
   $$

3. Solve the system of linear equations. If the solution is a constant diagonal matrix, $\hat{S}(x, y) = \alpha I$, conclude the inf–sup condition holds. Since the coefficient matrix of the linear system has symbolic entries corresponding to the components of $\hat{x}_1$ and $\hat{x}_2$ fraction free Gauss elimination should be employed for their solution.

Granted the integrity of the linear algebra system, this procedure provides a proof of the following lemma.

Lemma 4.3 *Let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of a bounded domain $\Omega \subset \mathbb{R}^2$ and let $Z_h \subset \mathbb{S}_h$ be the subspace of the lowest order $(k = 1)$ Arnold–Winther tensors with divergence zero and let*

$$
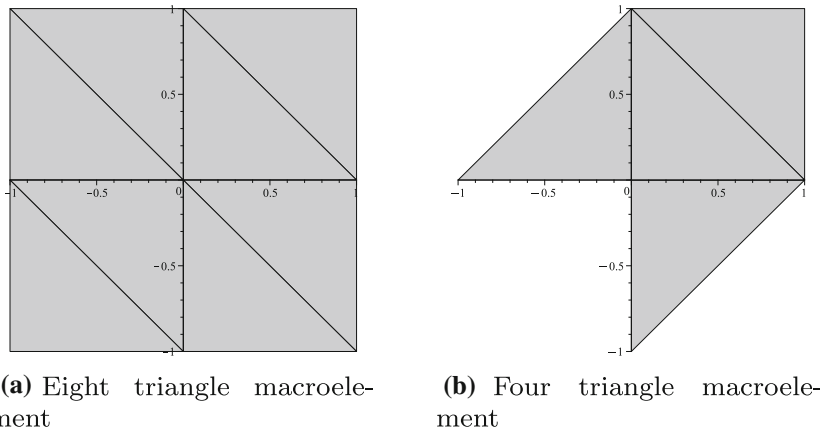\mathbb{D}_h = \left\{ D \in \mathbb{D} \mid D_{ij}|_K \in \mathcal{P}_2(K) \oplus b_K \quad \forall K \in \mathcal{T}_h \right\}.
$$

(a) Eight triangle macroelement

(b) Four triangle macroelement

Fig. 4 Parent macroelements used for inf–sup conditions for the trace-free D formulation (**a**) and for the formulation with pressure (**b**)

Then there is a constant $c > 0$ independent of h such that

$$\sup_{D_h \in \mathbb{D}_h} \frac{\int_\Omega D_h : S_h}{\|D_h\|_\mathbb{D}} \geq c\|S_h\|, \qquad S_h \in Z_h.$$

The bubble component was necessary for the two element macroelement construction. However, numerical experiments suggest that it may not be required if a larger macroelement is used. For example, if Ω is a square (or L-shaped etc.) and is triangulated by dividing uniform square grids along the diagonal, then groups of 8 triangles (four squares) can be mapped to the macroelement shown in Fig. 4a by maps of the form $x = \chi_M(\hat{x}) = x_0 + h\xi$ so that the Jacobian is $F = hI$. In this situation the linear system only contains numerical entries, and the macroelement calculations shows that the bubble is not required; that is the inf–sup condition is satisfied when

$$\mathbb{D}_h = \left\{ D \in \mathbb{D} \mid D_{ij}|_K \in \mathbb{P}_2(K) \ \ \forall K \in \mathcal{T}_h \right\}.$$

4.4 Alternate formulation with pressure

The solution (1.9) requires

$$\sup_{D \in \mathbb{D}} \frac{\int_\Omega D : S + pI : D}{\|D\|} \geq c(\|S\|_\mathbb{S} + \|p\|_P), \quad \forall (S, p) \in Z \times P. \qquad (4.9)$$

The proof of (4.9) is shown by setting $D = S - (1/d)tr(S)I$ (as for (4.3)) if $\|S\|_\mathbb{S} \geq \|p\|_P$ and choosing $D = pI - S$ when $\|S\|_\mathbb{S} \leq \|p\|_P$.

The macroelement construction is vastly simplified in this situation since the mappings $\chi : \hat{M} \to M$ do not need to preserve the trace. In particular, a single parent

element suffices and the inf–sup condition reduces to showing a single (numerical) matrix has full rank. For the parent element with four triangles shown in Fig. 4b the macroelement construction shows that the inf–sup condition (4.9) is satisfied when $\mathbb{S}_h$ and U_h are as in Sect. 4.2, and

$$\mathbb{D}_h = \left\{ D \in L^2(\Omega)_{sym}^{d \times d} \mid D_{ij}|_K \in \mathbb{P}_1(K) \oplus b_K \quad \forall K \in \mathcal{T}_h \right\},$$

and

$$P_h = \{p \in P \mid p|_K \in \mathbb{P}_1(K) \oplus b_K \quad \forall K \in \mathcal{T}_h\}.$$

If $\mathcal{P}_h^{\mathbb{D}}$ and $\mathcal{P}_h^P$ denote the orthogonal projections, onto these spaces we have

$$\|D - \mathcal{P}_h^{\mathbb{D}} D\|_{L^2(\Omega)} \le ch^m \|D\|_m, \quad 0 \le m \le 2,$$
$$\|p - \mathcal{P}_h^P p\|_{L^2(\Omega)} \le ch^m \|p\|_m, \quad 0 \le m \le 2.$$

The error estimate for the discrete problem becomes

$$\|D - D_h\|_{\mathbb{D}} + \|\mathbf{u} - \mathbf{u}_h\|_U + \|S - S_h\|_{\mathbb{S}} + \|p - p_h\|_P$$
$$\le Ch^m \{\|D\|_m + \|\mathbf{u}\|_m + \|S\|_m + \|\operatorname{div}(S)\|_m + \|p\|_m\}.$$

5 Appendix A: Auxiliary results

Lemma 5.1 *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain and let $\partial\Omega = \bar{\Gamma}_0 \cup \bar{\Gamma}_1$ be a decomposition into two open sets with $|\Gamma_1| \ne 0$. Let $U = \{\mathbf{u} \in H^1(\Omega)^d \mid \mathbf{u}|_{\Gamma_0} = 0\}$. Then there exists $c > 0$ such that*

$$\sup_{\mathbf{u} \in U} \frac{\int_\Omega p\, div(\mathbf{u})}{\|\mathbf{u}\|_{H^1(\Omega)}} \ge c \|p\|_{L^2(\Omega)}, \qquad p \in L^2(\Omega).$$

Proof Recall that

$$\sup_{\mathbf{u}_0 \in H_0^1(\Omega)^d} \frac{\int_\Omega p\, div(\mathbf{u}_0)}{\|\mathbf{u}_0\|_{H^1(\Omega)}} \ge c_0 \|p\|_{L^2(\Omega)/\mathbb{R}},$$

and $\|p\|_{L^2(\Omega)}^2 = \|p\|_{L^2(\Omega)/\mathbb{R}}^2 + |\Omega|\bar{p}^2 = \|p - \bar{p}\|_{L^2(\Omega)}^2 + |\Omega|\bar{p}^2$ where $\bar{p}$ is the average value of p. Let $U = H_0^1(\Omega)^d \oplus U_1$ be the orthogonal decomposition. Then writing

$\mathbf{u} \in U$ as $\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_1$ we compute

$$\int_\Omega p\,\mathbf{u} = \int_\Omega (p - \bar{p})\,div(\mathbf{u}_0 + \mathbf{u}_1) + \bar{p}\,div(\mathbf{u}_1)$$

$$= \int_\Omega (p - \bar{p})\,div(\mathbf{u}_0 + \mathbf{u}_1) + \bar{p}\int_{\Gamma_1} \mathbf{u}_1.\mathbf{n}.$$

Since the trace operator $\gamma : U \to H^{1/2}(\Gamma_1)$ is surjective, there exists $\mathbf{u}_1 \in U_1$ with $\|\mathbf{u}_1\|_{H^1(\Omega)} = 1$ such that the integral on the right is positive. Then

$$\int_\Omega p\,\mathbf{u} \geq \int_\Omega (p - \bar{p})\,div(\mathbf{u}_0) - \|p - \bar{p}\|_{L^2(\Omega)} + c_1|\bar{p}|$$

$$\geq (c_0\|\mathbf{u}_0\| - 1)\|p - \bar{p}\|_{L^2(\Omega)} + c_1|\bar{p}|$$

$$= (c_0\|\mathbf{u}_0\| - 1)\|p\|_{L^2(\Omega)/\mathbb{R}} + c_1|\bar{p}|.$$

Selecting $\|\mathbf{u}_0\| = (1 + c_1)/c_0$ gives $\|\mathbf{u}\| \leq C(c_0, c_1)$ which completes the proof.

Lemma 5.2 *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain and let $\partial\Omega = \bar{\Gamma}_0 \cup \bar{\Gamma}_1$ be a decomposition into two open sets with $|\Gamma_1| \neq 0$. Let $\mathbb{S} = \{S \in \mathbb{H}(\Omega; div) \mid S\mathbf{n}|_{\Gamma_1} = 0\}$ and $U = \{\mathbf{u} \in H^1(\Omega)^d \mid \mathbf{u}|_{\Gamma_0} = 0\}$. Then there exists $c > 0$ such that*

$$\|tr(S)\|_{L^2(\Omega)} \leq C\left(\|S_0\|_{L^2(\Omega)} + \|div(S)\|_{U'}\right),$$

where $S_0 = S - (tr(S)/d)I$ is the deviatoric part of S.

This is a variation of Lemma 3.1 in [3] and the proof follows analogously.

References

1. Arnold, D.N., Awanou, G., Winther, R.: Finite elements for symmetric tensors in three dimensions. Math. Comput. **77**(263), 1229–1251 (2008)
2. Arnold, D.N., Brezzi, F., Douglas, J. Jr.: PEERS: a new mixed finite element for plane elasticity. Jpn. J. Appl. Math. **1**(2), 347–367 (1984)
3. Arnold, D.N., Douglas, J. Jr., Gupta, C.P.: A family of higher order mixed finite element methods for plane elasticity. Numer. Math. **45**(1), 1–22 (1984)
4. Arnold, D.N., Falk, R.S., Winther R.: Differential complexes and stability of finite element methods. I. The de Rham complex. In: Compatible Spatial Discretizations. Volume 142 of IMA Vol. Math. Appl., pp 24–46. Springer, New York (2006)
5. Arnold, D.N., Falk, R.S., Winther R.: Differential complexes and stability of finite element methods. II. The elasticity complex. In: Compatible Spatial Discretizations. Volume 142 of IMA Vol. Math. Appl., pp 47–67. Springer, New York (2006)
6. Arnold, D.N., Falk, R.S., Winther, R.: Mixed finite element methods for linear elasticity with weakly imposed symmetry. Math. Comput. **76**(260), 1699–1723 (2007)
7. Arnold, D.N., Winther, R.: Mixed finite elements for elasticity. Numer. Math. **92**(3), 401–419 (2002)
8. Babuška, I.: The finite element method with Lagrangian multipliers. Numer. Math. **20**, 179–192 (1973)
9. Baranger, J., Najib, K., Sandri, D.: Numerical analysis of a three-fields model for a quasi-Newtonian flow. Comput. Methods Appl. Mech. Eng. **109**(3–4), 281–292 (1993)

10. Barrett, J.W., Liu, W.B.: Finite element error analysis of a quasi-Newtonian flow obeying the Carreau or power law. Numer. Math. **64**(4), 433–453 (1993)
11. Barrett, J.W., Liu, W.B.: Quasi-norm error bounds for the finite element approximation of a non-Newtonian flow. Numer. Math. **68**(4), 437–456 (1994)
12. Brenner, S.C., Scott, L.R.: The mathematical theory of finite element methods. In: Texts in Applied Mathematics, vol. 15, 2nd edn. Springer-Verlag, New York (2002)
13. Brezzi, F.: On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge **8**(R-2), 129–151 (1974)
14. Brezzi, F., Fortin, M.: Mixed and hybrid finite element methods. In: Springer Series in Computational Mathematics, vol. 15. Springer-Verlag, New York (1991)
15. Bustinza, R., Gatica, G.N., González, M., Meddahi, S., Stephan, E.P.: Enriched finite element subspaces for dual-dual mixed formulations in fluid mechanics and elasticity. Comput. Methods Appl. Mech. Eng. **194**(2–5), 427–439 (2005)
16. Chen, Z.: Analysis of expanded mixed methods for fourth-order elliptic problems. Numer. Methods Partial Differ. Equ. **13**(5), 483–503 (1997)
17. Chen, Z.: Expanded mixed finite element methods for linear second-order elliptic problems. I. RAIRO Modél. Math. Anal. Numér. **32**(4), 479–499 (1998)
18. Cockburn, B., Gopalakrishnan, J.: Error analysis of variable degree mixed methods for elliptic problems via hybridization. Math. Comput. **74**(252), 1653–1677 (2005)
19. Ervin, V.J., Howell, J.S., Stanculescu, I.: A dual-mixed approximation method for a three-field model of a nonlinear generalized Stokes problem. Comput. Methods Appl. Mech. Eng. **197**(33–40), 2886–2900 (2008)
20. Ervin, V.J., Jenkins, E.W., Sun, S.: Coupled generalized non-linear Stokes flow with flow through a porous media. SIAM J. Numer. Anal. **47**(2), 929–952 (2009)
21. Farhloul, M., Fortin, M.: A new mixed finite element for the Stokes and elasticity problems. SIAM J. Numer. Anal. **30**(4), 971–990 (1993)
22. Farhloul, M., Fortin, M.: Dual hybrid methods for the elasticity and the Stokes problems: a unified approach. Numer. Math. **76**(4), 419–440 (1997)
23. Farhloul, M., Zine, A.M.: A posteriori error estimation for a dual mixed finite element approximation of non-Newtonian fluid flow problems. Int. J. Numer. Anal. Model. **5**(2), 320–330 (2008)
24. Gatica, G.N.: Solvability and Galerkin approximations of a class of nonlinear operator equations. Z. Anal. Anwendungen **21**(3), 761–781 (2002)
25. Gatica, G.N., Gatica, L.F.: On the a priori and a posteriori error analysis of a two-fold saddle-point approach for nonlinear incompressible elasticity. Int. J. Numer. Methods Eng. **68**(8), 861–892 (2006)
26. Gatica, G.N., Gatica, L.F., Stephan, E.P.: A dual-mixed finite element method for nonlinear incompressible elasticity with mixed boundary conditions. Comput. Methods Appl. Mech. Eng. **196**(35–36), 3348–3369 (2007)
27. Gatica, G.N., González, M., Meddahi, S.: A low-order mixed finite element method for a class of quasi-Newtonian Stokes flows. I. A priori error analysis. Comput. Methods Appl. Mech. Eng. **193**(9–11), 881–892 (2004)
28. Gatica, G.N., Márquez, A., Meddahi, S.: A new dual-mixed finite element method for the plane linear elasticity problem with pure traction boundary conditions. Comput. Methods Appl. Mech. Eng. **197**(9–12), 1115–1130 (2008)
29. Gatica, G.N., Meddahi, S.: A fully discrete Galerkin scheme for a two-fold saddle point formulation of an exterior nonlinear problem. Numer. Funct. Anal. Optim. **22**(7–8), 885–912 (2001)
30. Gatica, G.N., Sayas, F.J.: Characterizing the inf-sup condition on product spaces. Numer. Math. **109**(2), 209–231 (2008)
31. Gatica, G.N., Stephan, E.P.: A mixed-FEM formulation for nonlinear incompressible elasticity in the plane. Numer. Methods Partial Differ. Equ. **18**(1), 105–128 (2002)
32. Girault, V., Raviart, P.A.: Finite element methods for Navier-Stokes equations. In: Springer Series in Computational Mathematics. Theory and Algorithms, vol. 5. Springer-Verlag, Berlin (1986)
33. Glowinski, R., Marroco, A.: Sur l'approximation par éléments finis d'order un et la résolution par pénalisation dualité d'une classe de problèmes de Dirichlet non linéaires. RAIRO 9ème année, pp. 41–76 (1975)
34. Johnson, C.: A mixed finite element method for the Navier-Stokes equations. RAIRO Anal. Numér. **12**(4), 335–348 (1978)

35. Manouzi, H., Farhloul, M.: Mixed finite element analysis of a non-linear three-fields Stokes model. IMA J. Numer. Anal. **21**(1), 143–164 (2001)
36. Marsden, J.E., Hughes, T.J.R.: Mathematical Foundations of Elasticity. Prentice-Hall, Englewood Cliffs (1983)
37. Ming, P., Shi, Z.: Dual combined finite element methods for non-Newtonian flow (II) parameter-dependent problem. ESAIM-M2AN **34**(5), 1051–1067 (2000)
38. Roberts, J.E., Thomas, J.M.: Mixed and hybrid methods. In: Handbook of Numerical Analysis, vol. II. Handb. Numer. Anal., II, pp. 523–639. North-Holland, Amsterdam (1991)
39. Scheurer, B.: Existence et approximation de points selles pour certains problèmes non linéaires. RAIRO Anal. Numér. **11**(4), 369–400 (1977)
40. Stenberg, R.: Analysis of mixed finite elements methods for the Stokes problem: a unified approach. Math. Comput. **42**(165), 9–23 (1984)