IMA Journal of Numerical Analysis (2006) **26**, 779−789
doi:10.1093/imanum/dri048
Advance Access publication on January 19, 2006

A note on least-squares mixed finite elements in relation to standard and mixed finite elements

JAN BRANDTS†
Korteweg-de Vries Institute for Mathematics, Faculty of Science, University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, Netherlands

AND

YANPING CHEN AND JULIE YANG
Institute for Computational and Applied Mathematics, Xiangtan University, Xiangtan 411105, China

[Received on 4 July 2005; revised on 23 November 2005]

The least-squares mixed finite-element method for second-order elliptic problems yields an approximation $u_h \in V_h \subset H_0^1(\Omega)$ of the potential u together with an approximation $\mathbf{p}_h \in \mathbf{\Gamma}_h \subset \mathbf{H}(\mathrm{div}\,;\,\Omega)$ of the vector field $\mathbf{p} = -A\nabla u$. Comparing u_h with the standard finite-element approximation of u in V_h, and $\mathbf{p}_h$ with the mixed finite-element approximation of $\mathbf{p}$, it turns out that they are higher-order perturbations of each other. In other words, they are 'superclose'. Refined a priori bounds and superconvergence results can now be proved. Also, the local mass conservation error is of higher order than could be concluded from the standard a priori analysis.

Keywords: least-squares mixed finite-element method; standard finite-element method; mixed finite-element method; supercloseness.

1. Introduction

The least-squares mixed finite element for second-order elliptic problems for given data f and A concerns the minimization over suitable finite-dimensional subspaces $V_h \times \mathbf{\Gamma}_h \subset H_0^1(\Omega) \times \mathbf{H}(\mathrm{div};\,\Omega)$ of the weighted quadratic functional

$$\hat{J}(v, \mathbf{q}) = |f - \mathrm{div}\,\mathbf{q}|_0^2 + |A^{-1/2}(\mathbf{q} + A\nabla v)|_0^2. \tag{1.1}$$

The objective is to obtain approximations $(u_h, \mathbf{p}_h)$ of the unique pair $(u, \mathbf{p})$ for which both norms in (1.1) vanish. Optimal-order a priori error bounds hold for a wide range of subspaces (Cai *et al.*, 1994; Pehlivanov & Carey, 1994; Pehlivanov *et al.*, 1994) and also non-self-adjoint problems (Pehlivanov *et al.*, 1996) and adaptive refinement have been considered (Carey & Pehlivanov, 1997). Since the least-squares mixed approximations u_h and $\mathbf{p}_h$ are taken from usual standard (Braess, 2001; Ciarlet, 2002; Brenner & Scott, 2002) and mixed (Brezzi & Fortin, 1991; Brenner & Scott, 2002; Raviart & Thomas, 1977) finite-element spaces, respectively, it seems natural to compare them directly with their standard and mixed counterparts $u_h^{\mathrm{s}} \in V_h$ and $\mathbf{p}_h^{\mathrm{m}} \in \mathbf{\Gamma}_h$, and study their difference. The main result of this paper is Theorem 3.6, which tells that they are superclose in the sense that under some conditions,

$$\|(u_h - u_h^{\mathrm{s}}, \mathbf{p}_h - \mathbf{p}_h^{\mathrm{m}})\|_{1\times\mathrm{div},A} \leqslant Ch\|(u - u_h^{\mathrm{s}}, \mathbf{p} - \mathbf{p}_h^{\mathrm{m}})\|_{1\times\mathrm{div},A}. \tag{1.2}$$

†Email: brandts@science.uva.nl

This result allows a priori bounds for the least-squares mixed method to be refined, and also for superconvergence (Křížek *et al.*, 1998; Wahlbin, 1995) results to be transferred. Apart from this, although the least-squares mixed finite-element method does not have exact mass conservation like the mixed method, it can now easily be proved that the error is smaller by one order of h than the analysis available to date.

1.1 *Discussion and references*

The results in this paper are based on the setting of Pehlivanov *et al.* (1994) and are mainly theoretical in nature, showing that for some standard choices of finite-element spaces, the least-squares approximations are only higher-order perturbations of the standard and mixed approximations. There has been considerable development in the least-squares area in the 1990s, in particular for non-symmetric elliptic problems, and for elliptic problems other than our model problem (see, for instance, Carey *et al.*, 1998; Carey & Pehlivanov, 1997; Pehlivanov *et al.*, 1996; Pehlivanov & Carey, 1994; Bramble *et al.*, 1998, and the review paper Bochev & Gunzburger, 1998). With regard to Carey & Pehlivanov (1997), the supercloseness and superconvergence results in our paper could be employed to develop a posteriori error estimators in the usual way. Moreover, Lemma 2.1 could be very useful in a number of applications, not only in the least-squares mixed finite-element setting. In this paper, it is used to prove coercivity of the least-squares mixed bilinear form in a compact and elegant way (see Section 2.3).

In Section 2, we recall the standard and mixed finite-element methods for a model problem, and introduce the least-squares mixed method while at the same time introducing notations and settings. In Section 3, we prove our main result as indicated above in the Introduction. In Section 4, we reflect on the consequences of the supercloseness in terms of a priori bounds and superconvergence results, and on local mass conservation.

2. Model problem and approximation

As our model problem we consider the following second-order elliptic problem: given $f \in H^{-1}(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is a convex polytope, find $u \in H_0^1(\Omega)$ such that

$$-\mathrm{div}(A\nabla u) = f \ \text{ in } \Omega, \quad u = 0 \ \text{ on } \partial\Omega, \tag{2.1}$$

where A is uniformly symmetric positive definite with Lipschitz continuous coefficients and with eigenvalues in the interval $[\beta^2, \beta^{-2}]$ for some $\beta \in (0, 1]$. The formulation of (2.1) as a system of first-order equations is to find functions $u \in H_0^1(\Omega)$ and $\mathbf{p} \in \mathbf{H}(\mathrm{div}; \Omega)$ such that

$$\mathbf{p} = -A\nabla u \ \text{ in } \Omega, \quad \mathrm{div}\,\mathbf{p} = f \ \text{ in } \Omega. \tag{2.2}$$

These formulations will be used to define three different finite-element methods for u and $\mathbf{p}$.

2.1 *Notations for spaces, norms and inner products*

We use standard notations for Sobolev spaces and their norms and semi-norms; the L_2-norm and inner product we denote by $|\cdot|_0$ and $(\cdot, \cdot)$. Additionally to the usual norms on $\mathbf{H}(\mathrm{div}; \Omega)$ and $H_0^1(\Omega)$ we define

$$\|\mathbf{q}\|^2_{\mathrm{div},A} = (\mathbf{q}, \mathbf{q})_{\mathrm{div},A}, \tag{2.3}$$

where

$$(\mathbf{r}, \mathbf{q})_{\mathrm{div},A} = d(\mathbf{r}, \mathbf{q}) + (\mathrm{div}\,\mathbf{r}, \mathrm{div}\,\mathbf{q}) \quad \text{and} \quad d(\mathbf{r}, \mathbf{q}) = (A^{-1}\mathbf{r}, \mathbf{q}), \tag{2.4}$$

as well as the usual energy norm and inner product on $H_0^1(\Omega)$,

$$|v|_{1,A}^2 = a(v, v), \quad \text{where } a(w, v) = (A\nabla w, \nabla v). \tag{2.5}$$

We equip the product space $H_0^1(\Omega) \times \mathbf{H}(\mathrm{div}; \Omega)$ with the canonical inner product and norm

$$(w, \mathbf{r}; v, \mathbf{q})_{1\times\mathrm{div},A} = a(w, v) + (\mathbf{r}, \mathbf{q})_{\mathrm{div},A} \quad \text{and} \quad \|(v, \mathbf{q})\|_{1\times\mathrm{div},A}^2 = (v, \mathbf{q}; v, \mathbf{q})_{1\times\mathrm{div},A}. \tag{2.6}$$

The above weighted norms are equivalent to the usual norms and semi-norms on $H^1(\Omega)$ and $\mathbf{H}(\mathrm{div}; \Omega)$. The Poincaré–Friedrichs inequality and the assumption on the eigenvalues of A show that

$$\sup_{0\neq v\in H_0^1(\Omega)} \frac{|v|_0}{|v|_{1,A}} = d_A < \infty. \tag{2.7}$$

The constant d_A depends only on the diameter of Ω and on β. To conclude, define the bilinear form b: $H^1(\Omega) \times \mathbf{H}(\mathrm{div}; \Omega)$ by

$$b(v, \mathbf{q}) = (v, \mathrm{div}\, \mathbf{q}); \tag{2.8}$$

then with the constant γ defined by

$$0 < \gamma = \sqrt{\frac{d_A^2}{d_A^2 + 1}} < 1, \tag{2.9}$$

we are able to prove the following useful lemma.

LEMMA 2.1 For all $v \in H_0^1(\Omega)$ and $\mathbf{q} \in \mathbf{H}(\mathrm{div}; \Omega)$,

$$b(v, \mathbf{q}) \leqslant \gamma\, |v|_{1,A} \|\mathbf{q}\|_{\mathrm{div},A} \leqslant \frac{1}{2}\gamma\, \|(v, \mathbf{q})\|_{1\times\mathrm{div},A}^2. \tag{2.10}$$

Proof. Let $\mathbf{q} \in \mathbf{H}(\mathrm{div}; \Omega)$ and $v \in H_0^1(\Omega)$. Then, by Green's formula,

$$b(v, \mathbf{q})^2 = (\nabla v, \mathbf{q})^2 = (A^{1/2}\nabla v, A^{-1/2}\mathbf{q})^2 \leqslant |v|_{1,A}^2 d(\mathbf{q}, \mathbf{q}). \tag{2.11}$$

On the other hand, definition (2.7) of d_A shows that

$$b(v, \mathbf{q})^2 \leqslant |v|_0^2 |\mathrm{div}\, \mathbf{q}|_0^2 \leqslant d_A^2 |v|_{1,A}^2 |\mathrm{div}\, \mathbf{q}|_0^2. \tag{2.12}$$

Multiply (2.11) by d_A^2 and add the result to (2.12). This gives

$$(d_A^2 + 1) b(v, \mathbf{q})^2 \leqslant d_A^2 |v|_{1,A}^2 (d(\mathbf{q}, \mathbf{q}) + |\mathrm{div}\, \mathbf{q}|_0^2) = d_A^2 |v|_{1,A}^2 \|\mathbf{q}\|_{\mathrm{div},A}^2. \tag{2.13}$$

This proves the first inequality. The second follows using $2|xy| \leqslant x^2 + y^2$. $\square$

2.2 *Standard and mixed finite elements*

Let $V_h \subset H_0^1(\Omega)$ be a standard finite-element space. The solution $u \in H_0^1(\Omega)$ of (2.1) and its approximation $u_h^{\mathrm{s}} \in V_h$ satisfy

$$\forall\, v \in H_0^1(\Omega),\ \ a(u, v) = f(v) \quad \text{and} \quad \forall\, v_h \in V_h,\ \ a(u_h^{\mathrm{s}}, v_h) = f(v_h). \tag{2.14}$$

Usually, $\mathbf{p}_h^{\mathrm{s}} = -A\nabla u_h^{\mathrm{s}}$, seen as an approximation of $\mathbf{p} = -A\nabla u$, is discontinuous. The system matrix of the discrete problem is symmetric and positive definite. See Braess (2001), Brenner & Scott (2002) and Ciarlet (2002) for details on the standard finite-element method. In the mixed finite-element method, the first-order formulation (2.2) is used. It can be shown that $(u, \mathbf{p})$ is the unique pair in $L^2(\Omega) \times \mathbf{H}(\mathrm{div};\,\Omega)$ such that

$$\forall\,(w, \mathbf{q}) \in L^2(\Omega) \times \mathbf{H}(\mathrm{div};\,\Omega), \quad d(\mathbf{p}, \mathbf{q}) - b(u, \mathbf{q}) + b(w, \mathbf{p}) = (f, w). \tag{2.15}$$

Let $W_h \subset L^2(\Omega)$ and $\boldsymbol{\Gamma}_h \subset \mathbf{H}(\mathrm{div};\,\Omega)$ and assume that

(A1) $W_h = \mathrm{div}\,\boldsymbol{\Gamma}_h$.

Under this assumption, the discrete version of (2.15) is to find $(u_h^{\mathrm{m}}, \mathbf{p}_h^{\mathrm{m}}) \in W_h \times \boldsymbol{\Gamma}_h$ such that

$$\forall\,(w_h, \mathbf{q}_h) \in W_h \times \boldsymbol{\Gamma}_h, \quad d(\mathbf{p}_h^{\mathrm{m}}, \mathbf{q}_h) - b(u_h^{\mathrm{m}}, \mathbf{q}_h) + b(w_h, \mathbf{p}_h^{\mathrm{m}}) = (f, w_h), \tag{2.16}$$

has a unique solution $(u_h^{\mathrm{m}}, \mathbf{p}_h^{\mathrm{m}})$. To derive optimal error bounds for a sequence of approximations, the pair $W_h, \boldsymbol{\Gamma}_h$ also needs to satisfy the Babuška–Brezzi condition (A6) further on.

The space $W_h \subset L^2(\Omega)$ generally yields discontinuous u_h^{m}. On the other hand, if $\boldsymbol{\Gamma}_h$ consists of piecewise polynomial functions, then each element from $\boldsymbol{\Gamma}_h$ has continuous normal components across the element boudaries. If W_h contains the piecewise constants and f is piecewise constant then (A1) implies that on each element of the partition of Ω,

$$\mathrm{div}\,(\mathbf{p} - \mathbf{p}_h) = 0, \tag{2.17}$$

or in other words, there is exact local mass conservation. We will return to this property in connection with our main theorem in Section 4. The mixed method, without further modifications, leads to an indefinite system matrix. We refer to Brenner & Scott (2002), Brezzi (1978) and Brezzi & Fortin (1991) for details.

2.3 *Least-squares mixed finite elements*

A method that aims to combine the favourable properties of both the standard and the mixed method is the least-squares mixed finite-element method. Define a functional $J\colon H_0^1(\Omega) \times \mathbf{H}(\mathrm{div};\,\Omega) \to \mathbb{R}$ by

$$J(v, \mathbf{q}) = (f - \mathrm{div}\,\mathbf{q},\ f - \mathrm{div}\,\mathbf{q}) + (\mathbf{q} + A\nabla v,\ A^{-1}(\mathbf{q} + A\nabla v)). \tag{2.18}$$

Clearly, $J(v, \mathbf{q}) \geqslant 0$ for all $(v, \mathbf{q}) \in H_0^1(\Omega) \times \mathbf{H}(\mathrm{div};\,\Omega)$, and the solution $(u, \mathbf{p}) \in H_0^1(\Omega) \times \mathbf{H}(\mathrm{div};\,\Omega)$ of (2.2) is the unique pair at which $J = 0$. Setting the first variation in (2.18) to zero gives

$$\forall (v, \mathbf{q}) \in H_0^1(\Omega) \times \mathbf{H}(\mathrm{div};\,\Omega), \quad B(u, \mathbf{p}|v, \mathbf{q}) = (f, \mathrm{div}\,\mathbf{q}), \tag{2.19}$$

where the bilinear form B is defined in terms of (2.6) and (2.8) by

$$B(w, \mathbf{r}; v, \mathbf{q}) = (w, \mathbf{r}; v, \mathbf{q})_{1\times \mathrm{div}, A} - b(w, \mathbf{q}) - b(v, \mathbf{r}). \tag{2.20}$$

This form is coercive on $H_0^1(\Omega) \times \mathbf{H}(\mathrm{div};\,\Omega)$. Indeed, for all non-zero $(v, \mathbf{q}) \in H_0^1(\Omega) \times \mathbf{H}(\mathrm{div};\,\Omega)$,

$$B(v, \mathbf{q}|v, \mathbf{q}) \geqslant (1 - \gamma)\|(v, \mathbf{q})\|_{1\times \mathrm{div}, A}^2 > 0, \tag{2.21}$$

which follows immediately from $B(v, \mathbf{q}|v, \mathbf{q}) = \|(v, \mathbf{q})\|^2_{1\times\mathrm{div},A} - 2b(v, \mathbf{q})$ and Lemma 2.1. Similarly, B is continuous in the sense that

$$|B(w, \mathbf{r}; v, \mathbf{q})| \leqslant (1 + \gamma)\|(w, \mathbf{r})\|_{1\times\mathrm{div},A}\|(v, \mathbf{q})\|_{1\times\mathrm{div},A}.$$

Hence, by the Lax–Milgram Lemma, the solution $(u, \mathbf{p})$ of (2.19) is unique in $H^1_0(\Omega) \times \mathbf{H}(\mathrm{div};\, \Omega)$, and there exists a unique pair $(u_h, \mathbf{p}_h) \in V_h \times \mathbf{\Gamma}_h$ such that

$$\forall (v_h, \mathbf{q}_h) \in V_h \times \mathbf{\Gamma}_h, \quad B(u_h, \mathbf{p}_h|v_h, \mathbf{q}_h) = (f, \mathrm{div}\, \mathbf{q}_h), \tag{2.22}$$

for each pair of subspaces $V_h \subset H^1_0(\Omega)$ and $\mathbf{\Gamma}_h \subset \mathbf{H}(\mathrm{div};\, \Omega)$. The pair $(u_h, \mathbf{p}_h)$ is the least-squares mixed finite-element approximation. See Cai *et al.* (1994), Pehlivanov *et al.* (1994) and Bochev & Gunzburger (1998) for details.

3. Comparison of the approximations

In this section, we will directly compare $u_h, \mathbf{p}_h$ with u^{s}_h, and u^{m}_h and $\mathbf{p}^{\mathrm{m}}_h$. In particular, we study the differences

$$\xi_h = u_h - u^{\mathrm{s}}_h \in V_h \quad \text{and} \quad \boldsymbol{\sigma}_h = \mathbf{p}_h - \mathbf{p}^{\mathrm{m}}_h \in \mathbf{\Gamma}_h. \tag{3.1}$$

We will use the following Galerkin orthogonality relations.

PROPOSITION 3.1 For all $v_h \in V_h$, $\mathbf{q}_h \in \mathbf{\Gamma}_h$ and $w_h \in W_h$,

$$a(u - u^{\mathrm{s}}_h, v_h) = 0, \tag{3.2}$$

$$d(\mathbf{p} - \mathbf{p}^{\mathrm{m}}_h, \mathbf{q}_h) - b(u - u^{\mathrm{m}}_h, \mathbf{q}_h) = 0, \tag{3.3}$$

$$b(w_h, \mathbf{p} - \mathbf{p}^{\mathrm{m}}_h) = 0, \tag{3.4}$$

$$B(u - u_h, \mathbf{p} - \mathbf{p}_h|v_h, \mathbf{q}_h) = 0. \tag{3.5}$$

Proof. The identities follow by subtracting the continuous equations of each method (tested with discrete functions) from the discrete equations. □

Under the very weak assumption (A1) we will now compare the different approximations.

LEMMA 3.2 Assume (A1). Then we have,

$$B(\xi_h, \boldsymbol{\sigma}_h|\xi_h, \boldsymbol{\sigma}_h) = d(\mathbf{p} - \mathbf{p}^{\mathrm{m}}_h, \boldsymbol{\sigma}_h) - b(u - u^{\mathrm{s}}_h, \boldsymbol{\sigma}_h) - b(\xi_h, \mathbf{p} - \mathbf{p}^{\mathrm{m}}_h). \tag{3.6}$$

Proof. The orthogonality relation (3.5) with the choice $(v_h, \mathbf{q}_h) = (\xi_h, \boldsymbol{\sigma}_h)$ gives

$$\begin{aligned} B(\xi_h, \boldsymbol{\sigma}_h|\xi_h, \boldsymbol{\sigma}_h) &= B(\xi_h, \boldsymbol{\sigma}_h|\xi_h, \boldsymbol{\sigma}_h) + B(u - u_h, \mathbf{p} - \mathbf{p}_h|\xi_h, \boldsymbol{\sigma}_h) \\ &= B(u - u^{\mathrm{s}}_h, \mathbf{p} - \mathbf{p}^{\mathrm{m}}_h|\xi_h, \boldsymbol{\sigma}_h). \end{aligned} \tag{3.7}$$

Now use (2.20) in combination with (3.2) and (3.4). Then all terms vanish except the ones mentioned. □

3.1 *Introducing an auxiliary problem*

Consider the problem of finding $\eta \in H^1_0(\Omega)$ and $\mathbf{r} \in \mathbf{H}(\mathrm{div};\, \Omega)$ such that

$$\mathbf{r} = -A\nabla\eta, \quad \text{and} \quad \mathrm{div}\,\mathbf{r} = \mathrm{div}\,\boldsymbol{\sigma}_h \;\text{ on } \Omega. \tag{3.8}$$

Then clearly, $(\eta, \mathbf{r})$ is the unique pair in $L^2(\Omega) \times \mathbf{H}(\mathrm{div};\, \Omega)$ for which

$$\forall (w, \mathbf{q}) \in L^2(\Omega) \times \mathbf{H}(\mathrm{div};\, \Omega), \quad d(\mathbf{r}, \mathbf{q}) - b(\eta, \mathbf{q}) + b(w, \mathbf{r}) = (\mathrm{div}\, \boldsymbol{\sigma}_h, w). \tag{3.9}$$

This mixed weak formulation can be discretized in $W_h \times \boldsymbol{\Gamma}_h$ with $W_h = \mathrm{div}\, \boldsymbol{\Gamma}_h$ which means that we look for the unique pair $(\eta_h, \mathbf{r}_h)$ for which

$$\forall (w_h, \mathbf{q}_h) \in W_h \times \boldsymbol{\Gamma}_h, \quad d(\mathbf{r}_h, \mathbf{q}_h) - b(\eta_h, \mathbf{q}_h) + b(w_h, \mathbf{r}_h) = (\mathrm{div}\, \boldsymbol{\sigma}_h, w_h). \tag{3.10}$$

We will now formulate a lemma that will enable us to rewrite the terms $d(\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}, \boldsymbol{\sigma}_h)$ and $b(u - u_h^{\mathrm{s}}, \boldsymbol{\sigma}_h)$ in the right-hand side of (3.6).

LEMMA 3.3 Assume (A1) and let $\mathbf{r}, \mathbf{r}_h$ and η, η_h be defined through (3.9) and (3.10). Then, for all $w_h \in W_h$,

$$d(\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}, \boldsymbol{\sigma}_h) = d(\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}, \mathbf{r}_h - \mathbf{r}) + b(\eta - w_h, \mathbf{p} - \mathbf{p}_h^{\mathrm{m}}). \tag{3.11}$$

Moreover, for all $v_h \in V_h$,

$$b(u - u_h^{\mathrm{s}}, \boldsymbol{\sigma}_h) = a(u - u_h^{\mathrm{s}}, \eta - v_h). \tag{3.12}$$

Proof. By (3.10) we have $\mathrm{div}\,(\mathbf{r}_h - \boldsymbol{\sigma}_h) = 0$. Substituting this in (3.3) and adding and subtracting a dummy term gives

$$d(\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}, \boldsymbol{\sigma}_h) = d(\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}, \mathbf{r}_h - \mathbf{r}) + d(\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}, \mathbf{r}). \tag{3.13}$$

Using the symmetry of the bilinear form d and choosing $w = 0$ and $\mathbf{q} = \mathbf{p} - \mathbf{p}_h^{\mathrm{m}}$ in (3.9) gives

$$d(\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}, \mathbf{r}) = b(\eta, \mathbf{p} - \mathbf{p}_h^{\mathrm{m}}). \tag{3.14}$$

Combining (3.13) and (3.14), the first statement follows using (3.4). The second statement follows from Green's formula, which shows that

$$b(u - u_h^{\mathrm{s}}, \boldsymbol{\sigma}_h) = b(u - u_h^{\mathrm{s}}, \mathbf{r}) = (u - u_h^{\mathrm{s}}, -\mathrm{div}\, A\nabla\eta) = a(u - u_h^{\mathrm{s}}, \eta). \tag{3.15}$$

The orthogonality relation (3.2) completes the proof. □

REMARK 3.4 The proof of (3.11) uses elements from a much more general theory in Douglas & Roberts (1985), whereas (3.12) is the Aubin–Nitsche argument from the standard setting in disguise.

COROLLARY 3.5 For all $w_h \in W_h$ and $v_h \in V_h$ we have,

$$B(\xi_h, \boldsymbol{\sigma}_h | \xi_h, \boldsymbol{\sigma}_h) = d(\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}, \mathbf{r}_h - \mathbf{r}) + b(\eta - w_h - \xi_h, \mathbf{p} - \mathbf{p}_h^{\mathrm{m}}) + a(u - u_h^{\mathrm{s}}, \eta - v_h).$$

Proof. Combine Lemmas 3.2 and 3.3. □

3.2 *Main theorem*

At this point we make some further assumptions, the first of which is that the elliptic problem is $H^2(\Omega)$-regular.

(A2) For all $f \in L^2(\Omega)$, the solution u of (2.1) satisfies $\|u\|_2 \leqslant C|f|_0$.

Together with the smoothness of the coefficients of A, that

$$\|\mathbf{r}\|_1 \leqslant C\|\eta\|_2 \leqslant C|\mathrm{div}\, \boldsymbol{\sigma}_h|_0 \leqslant C\|\boldsymbol{\sigma}_h\|_{\mathrm{div},A}. \tag{3.16}$$

Furthermore, we assume that there exist constants C such that

(A3) $\forall\, v \in H_0^1(\Omega) \cap H^2(\Omega), \quad \exists\, v_h \in V_h, \quad |v - v_h|_{1,A} \leqslant Ch|v|_2.$

(A4) $\forall\, \mathbf{q} \in [H^2(\Omega)]^1, \quad \exists\, \mathbf{q}_h \in \mathbf{\Gamma}_h, \quad |\mathbf{q} - \mathbf{q}_h|_0 \leqslant Ch|\mathbf{q}|_1.$

(A5) $\forall\, \mathbf{q}$ with $\operatorname{div} \mathbf{q} \in H^1(\Omega), \quad \exists\, \mathbf{q}_h \in \mathbf{\Gamma}_h, \quad |\operatorname{div} \mathbf{q} - \operatorname{div} \mathbf{q}_h|_0 \leqslant Ch|\operatorname{div} \mathbf{q}|_1.$

Notice that (A3), (A4) and (A5) hold for all the usual finite-element spaces; (A3) for the standard nodal (Lagrangian) finite elements and the couple (A4–A5) for the Raviart–Thomas (Raviart & Thomas, 1977) and Nédélec (Nédélec, 1980, 1986; Brezzi & Fortin, 1991) spaces. Finally, we will assume that

(A6) The Babuška–Brezzi condition is satisfied uniformly in the discretization parameter h,

$$\inf_{0 \neq w \in W_h} \sup_{\mathbf{q}_h \in \mathbf{\Gamma}_h} \frac{b(w_h, \mathbf{q}_h)}{|w|_0 \|\mathbf{q}_h\|_{\operatorname{div}}} > 0, \tag{3.17}$$

THEOREM 3.6 Let (A1)–(A6) be satisfied. Then

$$\|(\xi_h, \boldsymbol{\sigma}_h)\|_{1\times\operatorname{div},A} \leqslant Ch\|(\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}, u - u_h^{\mathrm{s}})\|_{1\times\operatorname{div},A}. \tag{3.18}$$

Proof. The coercivity (2.21) of B in combination with Corollary 3.5 means that for all $w_h \in W_h$ and all $v_h \in V_h$,

$$\begin{aligned}
(1-\gamma)&\|(\xi_h, \boldsymbol{\sigma}_h)\|_{1\times\operatorname{div},A}^2 \\
&\leqslant B(\xi_h, \boldsymbol{\sigma}_h | \xi_h, \boldsymbol{\sigma}_h) \\
&= d(\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}, \mathbf{r} - \mathbf{r}_h) + b(\eta - w_h - \xi_h, \mathbf{p} - \mathbf{p}_h^{\mathrm{m}}) + a(u - u_h^{\mathrm{s}}, \eta - v_h) \\
&\leqslant C\|\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}\|_{\operatorname{div},A}(|\mathbf{r} - \mathbf{r}_h|_0 + |\eta - w_h - \xi_h|_0) + |u - u_h^{\mathrm{s}}|_{1,A}|\eta - v_h|_{1,A}.
\end{aligned} \tag{3.19}$$

If we choose w_h to be the L^2-orthogonal projection of $\eta - \xi_h$ onto W_h, then (A5) assures us that

$$|\eta - \xi_h - w_h|_0 \leqslant Ch|\eta - \xi_h|_1 \leqslant Ch(|\eta|_{1,A} + |\xi_h|_{1,A}). \tag{3.20}$$

Furthermore, (A3) shows that v_h can be chosen such that $|\eta - v_h|_{1,A} \leqslant Ch\|\eta\|_2$, whereas (A4) and (A6) imply that $|\mathbf{r} - \mathbf{r}_h|_0 \leqslant Ch|\mathbf{r}|_1$. Together with the consequence (3.16) of the elliptic regularity (A2) this implies that (3.19) and the above imply that

$$\|(\xi_h, \boldsymbol{\sigma}_h)\|_{1\times\operatorname{div},A}^2 \leqslant Ch\|\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}\|_{\operatorname{div},A}(\|\boldsymbol{\sigma}_h\|_{\operatorname{div},A} + |\xi_h|_{1,A}) + Ch|u - u_h^{\mathrm{s}}|_{1,A}\|\boldsymbol{\sigma}_h\|_{\operatorname{div},A}. \tag{3.21}$$

This immediately gives the statement. □

The above theorem shows that under mild assumptions, the least-squares mixed finite-element approximations u_h of u and $\mathbf{p}_h$ of $\mathbf{p}$ are higher-order perturbations of the standard approximation u_h^{s} of u and the mixed approximation $\mathbf{p}_h^{\mathrm{m}}$ of $\mathbf{p}$, respectively.

4. Consequences and examples

Here we will show some direct consequences of Theorem 3.6. First of all, it provides refined a priori bounds for specific choices for the spaces V_h and $\mathbf{\Gamma}_h$. Moreover, it is also a means of transferring super-convergence results from one method to the other. Throughout this section we assume that (A1)–(A6) are all satisfied.

4.1 *Refined a priori bounds for the least-squares mixed method*

Céa's Lemma gives quasi-optimal convergence for the least-squares mixed approximations in the weighted product norm on $H^1_0(\Omega) \times \mathbf{H}(\mathrm{div};\,\Omega)$ in the sense that for all $(v_h, \mathbf{q}_h) \in H^1_0(\Omega) \times \mathbf{H}(\mathrm{div};\,\Omega)$,

$$\|(u - u_h, \mathbf{p} - \mathbf{p}_h)\|_{1\times\mathrm{div},A} \leqslant \frac{1+\gamma}{1-\gamma}\|(u - v_h, \mathbf{p} - \mathbf{q}_h)\|_{1\times\mathrm{div},A}. \tag{4.1}$$

It does not give bounds for norms of the individual errors $\|\mathbf{p} - \mathbf{p}_h\|_{\mathrm{div},A}$ and $|u - u_h|_{1,A}$ other than that each one of them is bounded by the right-hand side of (4.1). Theorem 3.6 gives additional information. By a simple triangle inequality we have

$$|u - u_h|_{1,A} \leqslant |u - u_h^{\mathrm{s}}|_{1,A} + Ch\|(u - u_h^{\mathrm{s}}, \mathbf{p} - \mathbf{p}_h^{\mathrm{m}})\|_{1\times\mathrm{div},A}.$$

If $\mathbf{\Gamma}_h$ has a higher approximation order than V_h, this results in

$$|u - u_h|_{1,A} \leqslant |u - u_h^{\mathrm{s}}|_{1,A} + Ch|u - u_h^{\mathrm{s}}|_{1,A},$$

which is no improvement over Céa's Lemma. On the other hand, if V_h has a higher approximation order than $\mathbf{\Gamma}_h$, then

$$|u - u_h|_{1,A} \leqslant |u - u_h^{\mathrm{s}}|_{1,A} + Ch\|\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}\|_{\mathrm{div},A}.$$

This improves the bound that results from Céa's Lemma, because the influence of the second term in the right-hand side is diminished by the supercloseness factor h. Of course, similar observations hold for $\|\mathbf{p} - \mathbf{p}_h\|_{\mathrm{div},A}$.

REMARK 4.1 Although these techniques allow us to derive optimal-order bounds for both variables with respect to the approximation order of the spaces, for one of the two the bound may not be optimal with respect to smoothness.

For example, let V_h be the space of continuous piecewise quadratic functions relative to a simplicial partition of the domain, and $\mathbf{\Gamma}_h$ the lowest order Raviart–Thomas (Raviart & Thomas, 1977) space with respect to the same partition. Then

$$\|u - u_h\|_{1,A} \leqslant Ch^2\|(u, \mathbf{p})\|_{3\times 2} \quad \text{and} \quad \|\mathbf{p} - \mathbf{p}_h\|_{\mathrm{div},A} \leqslant Ch\|(u, \mathbf{p})\|_{3\times 2}, \tag{4.2}$$

where $\|(\cdot,\cdot)\|_{s,t}$ is the product norm on $H^s(\Omega) \times H^t(\Omega)$. Céa's Lemma (4.1) only gives an $\mathcal{O}(h)$ bound in the norm $\|\cdot\|_{1\times\mathrm{div}}$. Full approximation quality for $\mathbf{p}_h$ is obtained, but more smoothness was required to establish this. Similarly, take for V_h the space of continuous piecewise linears and for $\mathbf{\Gamma}_h$ the one-but-lowest order Raviart–Thomas space, then we get

$$\|u - u_h\|_{1,A} \leqslant Ch\|(u, \mathbf{p})\|_{2\times 3} \quad \text{and} \quad \|\mathbf{p} - \mathbf{p}_h\|_{\mathrm{div},A} \leqslant Ch^2\|(u, \mathbf{p})\|_{2\times 3}. \tag{4.3}$$

Again, both are optimal approximation orders for the given spaces, but the first is not optimal with respect to smoothness. Combining the continuous piecewise linears for V_h with the lowest order Raviart–Thomas space is also optimal with respect to smoothness. This is exactly the situation for which Céa's Lemma suffices.

4.2 *Superconvergence for least-squares mixed elements*

Theorem 3.6 also gives us the possibility to transfer superconvergence results (Křížek *et al.*, 1998; Wahlbin, 1995) from the standard and mixed method to the least-squares mixed method. We will consider superconvergence in the sense of interpolants, i.e. a method is 'superconvergent' if it is superclose to some local interpolant of the exact solution, and if this interpolant allows a simple post-processing. Roughly speaking, this is the case for u_h^{s} and $\mathbf{p}_h^{\mathrm{m}}$ if there exist local interpolation schemes L_h into V_h and $\mathbf{\Pi}_h$ into $\mathbf{\Gamma}_h$ such that

- $|L_h u - u_h^{\mathrm{s}}|_{1,A} = z(h)|u - u_h^{\mathrm{s}}|_{1,A}$ with $\lim_{h\to 0} z(h) = 0$,
- $|\mathbf{\Pi}_h \mathbf{p} - \mathbf{p}_h^{\mathrm{m}}|_0 = y(h)\|\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}\|_{\mathrm{div},A}$ with $\lim_{h\to 0} y(h) = 0$,

respectively. These properties are, completely or partially, preserved in the least-squares approximations $\mathbf{p}_h$ and u_h, since the triangle inequality means that

$$|L_h u - u_h|_{1,A} \leqslant |L_h u - u_h^{\mathrm{s}}|_{1,A} + Ch\|(u - u_h^{\mathrm{s}}, \mathbf{p} - \mathbf{p}_h^{\mathrm{m}})\|_{1\times\mathrm{div},A}. \tag{4.4}$$

Therefore, if $\|\mathbf{p} - \mathbf{p}_h^{\mathrm{m}}\|_{\mathrm{div},A}$ is at least of the same approximation order as $|u - u_h^{\mathrm{s}}|_{1,A}$, we also have that

$$|L_h u - u_h|_{1,A} = z(h)|u - u_h|_{1,A} \quad \text{with } \lim_{h\to 0} z(h) = 0, \tag{4.5}$$

because in that case also $|u - u_h|_{1,A}$ has the same order as $|u - u_h^{\mathrm{s}}|_{1,A}$. Again, similar remarks apply to the vector variable $\mathbf{p}_h$.

As a first example, let V_h be the space of continuous piecewise linear or quadratic functions relative to a uniform simplicial partition of the domain in N space dimensions. If $\mathbf{\Gamma}_h$ has at least the same approximation order as V_h, then superconvergence for u_h follows from the similar property for u_h^{s} (Andreev & Lazarov, 1988; Brandts & Křížek, 2003, 2005). As a second example, let $\mathbf{\Gamma}_h$ be the space of lowest, or one-but-lowest order Raviart–Thomas (Raviart & Thomas, 1977) elements relative to a uniform triangular partition of a planar domain. If V_h has at least the same approximation order as V_h, then superconvergence for $\mathbf{p}_h$ follows from the corresponding property for $\mathbf{p}_h^{\mathrm{m}}$ (Brandts, 1994, 2000). Notice that the condition of uniform meshes comes from the superconvergence theory itself, and not from our theory in this paper. Indeed, superconvergence results of other types and for other elements (Douglas & Wang, 1989; Duran, 1990; Ewing *et al.*, 1991; Wang, 1991) can be transferred in a similar fashion.

4.3 *Higher-order local mass conservation*

The mixed finite-element method in which W_h contains the piecewise constants and in which f is assumed piecewise constant has exact local mass conservation in the sense that

$$\mathrm{div}(\mathbf{p} - \mathbf{p}_h) = 0. \tag{4.6}$$

Using Theorem (3.6), we find that

$$|\mathrm{div}(\mathbf{p} - \mathbf{p}_h)|_0 = |\mathrm{div}(\mathbf{p}_h^{\mathrm{m}} - \mathbf{p}_h)|_0 \leqslant Ch\|(u - u_h^{\mathrm{s}}, \mathbf{p} - \mathbf{p}_h^{\mathrm{m}})\|_{1\times\mathrm{div},A}, \tag{4.7}$$

which shows that even though there is no local mass conservation for the least-squares mixed finite-element method, the mass conservation error is smaller by a factor h than would be expected from the straightforward a priori bound.

Acknowledgements

The authors thank the anonymous referee who suggested the result in Section 4.3. JB was supported by a fellowship (1999–2004) of the Royal Netherlands Academy of Arts and Sciences. He gratefully acknowledges this support.

REFERENCES

ANDREEV, A. B. & LAZAROV, R. D. (1988) Superconvergence of the gradient for quadratic triangular finite elements. *Numer. Methods Partial Differential Equations*, **4**, 15–32.

BOCHEV, P. B. & GUNZBURGER, M. D. (1998). Finite elements of least-squares type. *SIAM Rev.*, **40**, 789–837.

BRAESS, D. (2001) *Finite Element: Theory, Fast Solvers, and Applications in Solid Mechanics*, 2nd edn. Cambridge: Cambridge University Press.

BRAMBLE, J. H., LAZAROV, R. D. & PASCIAK, J. E. (1998) Least-squares for second order elliptic problems. *Comput. Methods Appl. Mech. Engin.*, **152**, 195–210.

BRANDTS, J. H. (1994) Superconvergence and a posteriori error estimation in triangular mixed finite elements. *Numer. Math.*, **68**, 311–324.

BRANDTS, J. H. (2000) Superconvergence for triangular order $k = 1$ Raviart–Thomas mixed finite elements and for triangular standard quadratic finite element methods. *Appl. Numer. Anal.*, **34**, 39–58.

BRANDTS, J. H. & KŘÍŽEK, M. (2003) Gradient superconvergence on uniform simplicial partitions of polytopes. *IMA J. Numer. Anal.*, **23**, 1–17.

BRANDTS, J. H. & KŘÍŽEK, M. (2005) Superconvergence of tetrahedral quadratic finite elements. *J. Comp. Math.*, **23**, 27–36.

BRENNER, S. C. & SCOTT, L. R. (2002) *The Mathematical Theory of Finite Element Methods*, 2nd edn. New York: Springer.

BREZZI, F. (1978) On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *RAIRO*, **R2**, 129–151.

BREZZI, F. & FORTIN, M. (1991) *Mixed and Hybrid Finite Element Methods*. Berlin: Springer.

CAI, Z., LAZAROV, R., MANTEUFFEL, T. A. & MCCORMICK, S. F. (1994) First-order system least squares for second-order partial differential equations: part I. *SIAM J. Numer. Anal.*, **31**, 1785–1799.

CAREY, G. F. & PEHLIVANOV, A. I. (1997) Local error estimation and adaptive remeshing scheme for least-squares mixed finite elements. *Comput. Methods Appl. Mech. Engin.*, **150**, 125–131.

CAREY, G. F., PEHLIVANOV, A. I., SHEN, Y., BOSE, A. & WANG, K. C. (1998) Least-squares finite elements for fluid flow and transport. *Int. J. Numer. Methods Fluids*, **27**, 97–107.

CIARLET, P. (2002) *The Finite Element Method for Elliptic Problems*, 2nd edn. SIAM Classics in Applied Mathematics, vol. 40.

DOUGLAS, J. & ROBERTS, J. E. (1985) Global estimates for mixed methods for 2nd order elliptic problems. *Math. Comp.*, **44**, 39–52.

DOUGLAS, J. & WANG, J. (1989) Superconvergence of mixed finite element methods on rectangular domains. *Calcolo*, **26**, 121–134.

DURAN, R. (1990) Superconvergence for rectangular mixed finite elements. *Numer. Math.*, **58**, 2–15.

EWING, R. E., LAZAROV, R. D. & WANG, J. (1991) Superconvergence of the velocity along the Gauss lines in mixed finite element methods. *SIAM J. Numer. Anal.*, **28**, 1015–1029.

KŘÍŽEK, M., NEITTAANMÄKI, P. & STENBERG, R., EDS (1998) *Finite Element Methods: Superconvergence, Post-Processing and A Posteriori Estimates, Proceedings of the Conference at University of Jyväskylä, 1996.* Lecture Notes in Pure and Applied Mathematics, vol. 196. New York: Marcel Dekker.

NÉDÉLEC, J. C. (1980) Mixed finite elements in R^3. *Numer. Math.*, **35**, 315–341.

NÉDÉLEC, J. C. (1986) A new family of mixed finite elements in R^3. *Numer. Math.*, **50**, 57–81.

PEHLIVANOV, A. I. & CAREY, G. F. (1994) Error estimates for least-squares mixed finite elements. *RAIRO Math. Model. Numer. Anal.*, **28**, 499–516.

PEHLIVANOV, A. I., CAREY, G. F. & LAZAROV, R. D. (1994) Least-squares mixed finite elements for second order elliptic problems. *SIAM J. Numer. Anal.*, **31**, 1368–1377.

PEHLIVANOV, A. I., CAREY, G. F. & VASSILEVSKI, P. S. (1996) Least-squares mixed finite element methods for non-self-adjoint elliptic problems. *Numer. Math.*, **72**, 501–522.

RAVIART, P. A. & THOMAS, J. M. (1977) *A Mixed Finite Element Method for 2nd Order Elliptic Problems*. Lecture Notes in Mathematics, vol. 606. Berlin: Springer, pp. 292–315.

WAHLBIN, L. B. (1995) *Superconvergence in Galerkin Finite Element Methods*. Lecture Notes in Mathematics, vol. 1605. New York: Springer.

WANG, J. (1991) Superconvergence and extrapolation for mixed finite element methods on rectangular domains. *Math. Comput.*, **56**, 477–503.