Analisys of new augmented Lagrangian formulations for mixed finite element schemes ⋆

Daniele Boffi, Carlo Lovadina

Dipartimento di Matematica, Università di Pavia, Via Abbiategrasso 209, 27100 Pavia, Italia

Summary. Two new augmented Lagrangian formulations for mixed finite element schemes are presented. The methods lead, in some cases, to an improvement in the order of the approximation. An error analysis is provided, together with some interesting examples of applications.

Mathematics Subject Classifications (1991): 65N30

1. Introduction

The problem of minimizing a functional with a linear constraint is often encountered in applied sciences. In this paper we are interested in the finite element approximation of the minimizer u of a functional

$$F(v) = \frac{1}{2}a_0(v,v) - < f, v > \tag{1.1}$$

under the linear constraint $Lu = 0$. A classical way to face this problem (cf. e.g. Brezzi-Fortin (1991)) consists in introducing a Lagrange multiplier μ and in approximating the saddle-point (u, λ) of the functional

$$G(v,\mu) = F(v) + < Lv, \mu > \quad . \tag{1.2}$$

It is well-known that the main trouble for a good discretization consists in finding spaces which satisfy some compatibility conditions. Essentially things may go wrong if at the discrete level the constraint $Lu = 0$ is imposed too strongly (locking phenomenon) or too weakly (spurious modes). For a review of this topic, we refer to Brezzi-Fortin (1991).

Correspondence to: D. Boffi

⋆ This work has been partially supported by I.A.N.-CNR, Pavia

In particular spurious modes on u can occur if the functional $F(v)$ is not coercive on the whole space, but only on the subspace $Lv = 0$. On the other hand, in many applications it happens that the functional

$$\widetilde{F}(v) = F(v) + \frac{\gamma}{2}||Lv||^2 \tag{1.3}$$

is coercive, for any constant $\gamma > 0$.

It is also clear that $\widetilde{F}$ takes its minimum at the same point u as F. Hence an equivalent mixed formulation of (1.2) consists in finding the saddle-point (u, λ) of the functional

$$\widetilde{G}(v, \mu) = \widetilde{F}(v) + < Lv, \mu > \quad . \tag{1.4}$$

Problem (1.4) is known as augmented Lagrangian formulation (cf. Fortin-Glowinski (1983) and Girault-Raviart (1986)). The advantage of this formulation is that spurious modes on u are avoided.

We will analyze the approximation of (1.3) by means of discrete augmented Lagrangian formulations in which the parameter γ depends on a power of the mesh size h. For these formulations we provide an error analysis (cf. section 3) that shows how in some cases an improvement with respect to the old augmented Lagrangian formulation can be gained.

In section 2 we state the problem, by briefly recalling the known results on standard Lagrangian and augmented Lagrangian formulation. As usual, we treat the more general case, in which the bilinear form $a_0(\cdot, \cdot)$ may be not symmetric.

In section 3 we introduce the new augmented formulations and give the details for the error analysis.

In section 4 we conclude, by showing how the new scheme can be used to improve the convergence rate in the mixed approximation of some classical problem (Stokes problem, Laplace equation, Reissner–Mindlin plates and shells).

In what follows, C will denote a constant, which possibly differs from one formula to another and which is independent of h.

Finally, we would like to thank the referees for the useful suggestions that made possible to correct and to improve the first draft of the manuscript.

2. Augmented Lagrangian formulation

Let V, W and H be real Hilbert spaces and suppose that the following dense and continuous inclusion holds

$$W \subseteq H \equiv H' \subseteq W' \; .$$

Let $a_0(\cdot, \cdot) : V \times V \to \mathbf{R}$ be a continuous bilinear form and $L : V \to W$ be a continuous linear operator. Let f be given in the space V'.

We are interested in the approximation of the following classical variational problem

$$\text{find } (u, \lambda) \in V \times W' \text{ such that}$$
$$\begin{cases} a_0(u, v) + < Lv, \lambda >_{W \times W'} = < f, v >_{V' \times V} & \forall v \in V \\ < Lu, \mu >_{W \times W'} = 0 & \forall \mu \in W' \; . \end{cases} \tag{2.1}$$

Remark 2.1 It is well-known that whenever the form $a_0(\cdot,\cdot)$ is also symmetric, problem (2.1) is equivalent to the saddle-point problem

$$\inf_{\mu\in W'}\sup_{v\in V}\left\{\frac{1}{2}a_0(v,v)+ < Lv,\mu >_{W\times W'} - < f,v >_{V'\times V}\right\} . \tag{2.2}$$

We recall that a sufficient condition for problem (2.1) to be well-posed is the coercivity of the bilinear form $a_0(\cdot,\cdot)$ on the kernel $K=\{v\in V\,|\,Lv=0\}$ and the surjectivity of the linear operator $L:V\to W$. From now on we will always assume that these hypotheses are fulfilled.

A finite element approximation of (2.1) consists in choosing finite dimensional spaces $V_h\subset V$, $W_h\subset H$ and in solving the discrete problem

$$\text{find } (u_h,\lambda_h)\in V_h\times W_h \text{ such that}$$
$$\begin{cases} a_0(u_h,v_h)+ < Lv_h,\lambda_h >_{W\times W'}=< f,v_h >_{V'\times V} & \forall v_h\in V_h \\ < Lu_h,\mu_h >_{W\times W'}=0 & \forall \mu_h\in W_h\ . \end{cases} \tag{2.3}$$

Remark 2.2 To tell the truth, a conforming approximation would only require $W_h\subset W'$. Indeed, in most applications $H=\mathrm{L}^2$, so that any natural finite element space will be included in H.

It is standard to obtain (cf. e.g. Brezzi-Fortin (1991)) that (2.3) is uniquely solvable if there exist two positive constants α_h and β_h such that

$$a_0(v_h,v_h)\ge\alpha_h||v_h||_V^2 \quad \forall v_h\in K_h \tag{2.4}$$

$$\inf_{\mu_h\in W_h}\sup_{v_h\in V_h}\frac{(Lv_h,\mu_h)_H}{||v_h||_V||\mu_h||_{W'}}\ge\beta_h\ , \tag{2.5}$$

where $K_h=\{v_h\in V_h\,|\,(Lv_h,\mu_h)_H=0\quad\forall\mu_h\in W_h\}$.

Moreover if both α_h and β_h are uniformly bounded away from zero with respect to h, the following error estimate holds

$$||u-u_h||_V+||\lambda-\lambda_h||_{W'}\le C\inf_{v_h,\mu_h}\{||u-v_h||_V+||\lambda-\mu_h||_{W'}\}\ . \tag{2.6}$$

It is easily seen that finding finite element spaces satisfying simultaneously both (2.4) and (2.5) is not, in general, a trivial task. Indeed condition (2.5) essentially requires that the space V_h is sufficiently "rich" with respect to W_h; on the other hand if V_h is too large, K_h may consequently grow and condition (2.4) might fail.

Remark 2.3 If the bilinear form is coercive on the whole space V, then condition (2.4) is automatically fulfilled, due to the fact that V_h is contained in V. In this case the only hypothesis to check is the inf-sup condition (2.5), with $\beta_h\ge\beta>0$.

Let us now consider the following problem

$$\text{find } (u,\lambda)\in V\times W' \text{ such that}$$
$$\begin{cases} a_0(u,v)+\gamma(Lu,Lv)_W+ < Lv,\lambda >_{W\times W'}=< f,v >_{V'\times V} & \forall v\in V \\ < Lu,\mu >_{W\times W'}=0 & \forall\mu\in H\ , \end{cases} \tag{2.7}$$

where γ is a positive constant.

It is clear that (2.7) admits the same unique solution of problem (2.1), thanks to the fact that $Lu = 0$ in W. The discrete counterpart of formulation (2.7) reads as follows

(2.8)

$$\text{find } (u_h, \lambda_h) \in V_h \times W_h \text{ such that}$$
$$\begin{cases} a_0(u_h, v_h) + \gamma(Lu_h, Lv_h)_W + < Lv_h, \lambda_h >_{W\times W'} = < f, v_h > & \forall v_h \in V_h \\ < Lu_h, \mu_h >_{W\times W'} = 0 & \forall \mu_h \in W_h \ . \end{cases}$$

It is not hard to show that, under the usual hypotheses for problem (2.1) to be well-posed, there exists a positive constant η such that

$$a_0(v, v) + \eta||Lv||_W^2 \geq \delta||v||_V^2 \quad \forall v \in V \ , \tag{2.9}$$

with $\delta > 0$. If moreover the bilinear form $a_0(\cdot,\cdot)$ is positive-semidefinite (as we will always suppose in the following), then one can choose $\eta = 1$. It follows that condition (2.4) is automatically verified, by substituting $a_0(\cdot,\cdot)$ with $\widetilde{a}(\cdot,\cdot) = a_0(\cdot,\cdot) + \gamma||L\cdot||_W^2$. Hence any discretization based on the augmented formulation (2.7) will lead to a stable and convergent method (cf. error estimate (2.6)) provided that the only inf-sup condition (2.5) is fulfilled. This makes things easier in designing performant mixed schemes.

Remark 2.4 When $a_0(\cdot,\cdot)$ is symmetric, problem (2.7) is equivalent to the standard augmented Lagrangian saddle-point problem

$$\inf_{\mu\in W'} \sup_{v\in V} \left\{ \frac{1}{2} a_0(v, v) + \frac{\gamma}{2}||Lv||_W^2 + < Lv, \mu >_{W\times W'} - < f, v >_{V'\times V} \right\} . \tag{2.10}$$

3. New formulations and error analysis

In the previous section, we were dealing with the augmented formulation (2.7), in which the parameter γ has been supposed to be a certain positive constant. In this section we propose two ways of modifying that formulation: the common feature of our modifications consists in choosing $\gamma = \gamma_h = h^{-\alpha}$, with $\alpha \geq 0$. As we will see, our modifications are able to improve, in some cases, the error estimate (2.6).

First formulation

The first scheme we propose consists in solving the following variational system

$$\text{find } (u_h, \lambda_h) \in V_h \times W_h \text{ such that}$$
$$\begin{cases} a_0(u_h, v_h) + h^{-\alpha}(Lu_h, Lv_h)_W + < Lv_h, \lambda_h >_{W\times W'} \\ \qquad\qquad = < f, v_h >_{V'\times V} & \forall v_h \in V_h \\ < Lu_h, \mu_h >_{W\times W'} = 0 & \forall \mu_h \in W_h \ . \end{cases} \tag{3.1}$$

Before dealing with the error analysis of problem (3.1), let us briefly recall our notation:

$$K = \{v \in V \,|\, Lv = 0\}$$

$$K_h = \{v_h \in V_h \,|\, (Lv_h, \mu_h)_H = 0 \;\forall \mu_h \in W_h\}$$

and hypotheses:

there exist positive constants δ_1 and δ_2 independent of h such that

$$\begin{array}{lll} i) & a_0(v,v) \geq 0 & \forall v \in V \\ ii) & a_0(v,v) \geq \delta_1 ||v||_V^2 & \forall v \in K \\ iii) & L : V \to W & \text{surjective and continuous} \\ iv) & \inf\limits_{\mu_h \in W_h} \sup\limits_{v_h \in V_h} \dfrac{(Lv_h, \mu_h)_H}{||v_h||_V ||\mu_h||_{W'}} \geq \delta_2 > 0 & \end{array} \tag{3.2}$$

We recall that assumptions $i)$, $ii)$ and $iii)$ imply the existence of a positive constant δ_3 such that

$$a_0(v,v) + ||Lv||_W^2 \geq \delta_3 ||v||_V^2 \quad \forall v \in V \tag{3.3}$$

holds. Moreover, it is well-known that from $iv)$ and (3.3) it follows that problem (3.1) is uniquely solvable.

In the error estimates, we shall use the following lemma (cf. Brezzi-Fortin (1991))

Lemma 3.1 *Under hypothesis $iv)$ of (3.2), we have*

$$\inf_{v_h \in K_h} ||u - v_h||_V \leq C \inf_{w_h \in V_h} ||u - w_h||_V, \tag{3.4}$$

where $u \in K$. In particular (3.4) holds if u is the first component of the solution of the continuous problem (2.1).

We are now ready to perform our error analysis. Let $v_h \in K_h$ be given. From the triangle inequality we have

$$||u - u_h||_V \leq ||u - v_h||_V + ||u_h - v_h||_V \,. \tag{3.5}$$

Let us estimate $||u_h - v_h||_V$. From assumption $i)$ of (3.2), (3.3) and the fact that (u, λ) and (u_h, λ_h) solve problems (2.1) and (3.1), respectively, we can deduce that

$$
\begin{aligned}
\delta_3||u_h - v_h||_V^2 + h^{-\alpha}||L(u_h - v_h)||_W^2 \\
&\le 2\Big(a_0(u_h - v_h, u_h - v_h) + h^{-\alpha}||L(u_h - v_h)||_W^2\Big) \\
&= 2\Big(a_0(u_h - u, u_h - v_h) + h^{-\alpha}(L(u_h - u), L(u_h - v_h))_W \\
&\quad + a_0(u - v_h, u_h - v_h) + h^{-\alpha}(L(u - v_h), L(u_h - v_h))_W\Big) \\
&= 2\Big(< L(u_h - v_h), \lambda - \lambda_h >_{W\times W'} \\
&\quad + a_0(u - v_h, u_h - v_h) + h^{-\alpha}(L(u - v_h), L(u_h - v_h))_W\Big) \\
&= T_1 + T_2 + T_3 \, .
\end{aligned} \tag{3.6}
$$

We now estimate the three terms T_1, T_2 and T_3. We first observe that from the definition of K_h, we have

$$< L(u_h - v_h), \lambda - \lambda_h >_{W\times W'} = < L(u_h - v_h), \lambda - \mu_h >_{W\times W'}$$

for any $\mu_h \in W_h$; so we obtain

$$T_1 \le \varepsilon h^{-\alpha}||L(u_h - v_h)||_W^2 + \frac{h^\alpha}{\varepsilon}||\lambda - \mu_h||_{W'}^2 \tag{3.7}$$

with ε positive constant to be chosen.

Due to the continuity of $a_0(\cdot,\cdot)$ and of the linear operator L, we have the following estimates for T_2 and T_3

$$
\begin{aligned}
T_2 &\le C\Big(\sigma||u_h - v_h||_V^2 + \frac{1}{\sigma}||u - v_h||_V^2\Big) \\
T_3 &\le C\Big(\tau h^{-\alpha}||L(u_h - v_h)||_W^2 + \frac{h^{-\alpha}}{\tau}||u - v_h||_V^2\Big) \, ,
\end{aligned} \tag{3.8}
$$

with σ and τ positive constants to be chosen.

Then, if one takes ε, σ and τ sufficiently small, Equations (3.6)–(3.8) imply

$$
\begin{aligned}
||u_h - v_h||_V^2 + &h^{-\alpha}||L(u_h - v_h)||_W^2 \\
&\le C\Big(h^\alpha||\lambda - \mu_h||_{W'}^2 + ||u - v_h||_V^2 + h^{-\alpha}||u - v_h||_V^2\Big)
\end{aligned} \tag{3.9}
$$

for any $v_h \in K_h$ and any $\mu_h \in W_h$.

Thanks to Lemma 3.1 and (3.5), estimate (3.9) leads to

$$||u - u_h||_V \le C\Big((1 + h^{-\alpha/2}) \inf_{v_h\in V_h} ||u - v_h||_V + h^{\alpha/2} \inf_{\mu_h\in W_h} ||\lambda - \mu_h||_{W'}\Big) \, . \tag{3.10}$$

Let us now perform an error analysis for the multiplier λ. Again, we use the triangle inequality:

$$||\lambda - \lambda_h||_{W'} \le ||\lambda - \mu_h||_{W'} + ||\lambda_h - \mu_h||_{W'} \, . \tag{3.11}$$

From the inf-sup condition iv) of (3.2) we have

$$
\begin{aligned}
&\delta_2||\lambda_h - \mu_h||_{W'} \le \sup_{z_h \in V_h} \frac{(Lz_h, \lambda_h - \mu_h)_H}{||z_h||_V} \\
&= \sup_{z_h \in V_h} \frac{1}{||z_h||_V} \Big(< Lz_h, \lambda_h - \lambda >_{W \times W'} + < Lz_h, \lambda - \mu_h >_{W \times W'} \Big) \\
&= \sup_{z_h \in V_h} \frac{1}{||z_h||_V} \Big(a_0(u - u_h, z_h) + h^{-\alpha}(L(u - u_h), Lz_h)_W \\
&\quad + < Lz_h, \lambda - \mu_h >_{W \times W'} \Big) .
\end{aligned} \tag{3.12}
$$

In the last equality we have again used the fact that (u, λ) and (u_h, λ_h) solve the problems (2.1) and (3.1), respectively.

Since $a_0(\cdot, \cdot)$ and the operator L are continuous, (3.12) implies

$$
||\lambda_h - \mu_h||_{W'} \le C\Big(||u - u_h||_V + h^{-\alpha}||L(u - u_h)||_W + ||\lambda - \mu_h||_{W'}\Big) . \tag{3.13}
$$

We fix now our attention on the second term in the right-hand side of (3.13). We have

$$
h^{-\alpha}||L(u - u_h)||_W \le h^{-\alpha}||L(u - v_h)||_W + h^{-\alpha}||L(u_h - v_h)||_W \tag{3.14}
$$

for any choice of $v_h \in K_h$.

Recalling (3.9) and Lemma 3.1, we deduce from (3.14) that

$$
\begin{aligned}
&h^{-\alpha}||L(u - u_h)||_W \\
&\le Ch^{-\alpha/2}\Big((1 + h^{-\alpha/2}) \inf_{v_h \in V_h} ||u - v_h||_V + h^{\alpha/2} \inf_{\mu_h \in W_h} ||\lambda - \mu_h||_{W'}\Big) .
\end{aligned} \tag{3.15}
$$

We are now in the position to get the final estimate (cf. Equations (3.11), (3.13) and (3.15))

$$
\begin{aligned}
&||\lambda - \lambda_h||_{W'} \\
&\le C\Big((1 + h^{-\alpha/2} + h^{-\alpha}) \inf_{v_h \in V_h} ||u - v_h||_V + \inf_{\mu_h \in W_h} ||\lambda - \mu_h||_{W'}\Big) .
\end{aligned} \tag{3.16}
$$

What we have done so far can be summarized in the following

Theorem 3.2 *Assume hypotheses (3.2). Let (u, λ) be the solution of problem (2.1) and let (u_h, λ_h) be the solution of the discrete problem (3.1). Then the following error estimates hold*

$$
\begin{aligned}
&||u - u_h||_V \\
&\le C\Big((1 + h^{-\alpha/2}) \inf_{v_h \in V_h} ||u - v_h||_V + h^{\alpha/2} \inf_{\mu_h \in W_h} ||\lambda - \mu_h||_{W'}\Big) \\
&||\lambda - \lambda_h||_{W'} \\
&\le C\Big((1 + h^{-\alpha/2} + h^{-\alpha}) \inf_{v_h \in V_h} ||u - v_h||_V + \inf_{\mu_h \in W_h} ||\lambda - \mu_h||_{W'}\Big) .
\end{aligned} \tag{3.17}
$$

We conclude the analysis of this scheme by remarking that in some cases Theorem 3.2 provides an improvement of error estimate (2.6).

Suppose, in fact, to choose the spaces V_h and W_h in such a way that

$$\inf_{v_h\in V_h} ||u-v_h||_V = O(h^{k+\beta}), \qquad \inf_{\mu_h\in W_h} ||\lambda-\mu_h||_{W'} = O(h^k) \ , \tag{3.18}$$

with $k>0$ and $\beta\geq 0$.

In this case estimate (2.6) implies that the rate of convergence of the standard scheme (2.3) is of order k for both u_h and λ_h. On the other hand estimates (3.17) for the new scheme (3.1) read as follows

$$\begin{aligned} ||u-u_h||_V &= O(h^{\min\{k+\beta-\alpha/2,k+\alpha/2\}}) \\ ||\lambda-\lambda_h||_{W'} &= O(h^{\min\{k+\beta-\alpha,k\}}) \ . \end{aligned} \tag{3.19}$$

Looking at (3.19), it is suddenly seen that the optimal choice is $\alpha=\beta$. We have obtained the following

Corollary 3.3 *Suppose that the choice of V_h and W_h is such that (3.18) holds, then the optimal rate of convergence is met for $\alpha=\beta$. In this case we obtain*

$$\begin{aligned} ||u-u_h||_V &= O(h^{k+\beta/2}) \\ ||\lambda-\lambda_h||_{W'} &= O(h^k) \ . \end{aligned} \tag{3.20}$$

Note that the choice (3.18) with $\beta>0$ leads to a suboptimal convergence estimate for the variable u_h (cf. (2.6)). The new augmented Lagrangian method with the choice $\alpha=\beta$ allows to reduce the suboptimality from β to $\beta/2$.

Second formulation

The second scheme is obtained from the variational formulation (3.1) just changing the term $h^{-\alpha}(Lu_h,Lv_h)_W$ into $h^{-\alpha}(Lu_h,Lv_h)_H$. This because it might happen that computing the inner product in W is more complicated than computing the inner product in H (which, in most applications, turns out to be the space L^2).

The new problem to solve reads as follows

$$\begin{aligned} &\text{find } (u_h,\lambda_h)\in V_h\times W_h \text{ such that} \\ &\begin{cases} a_0(u_h,v_h)+h^{-\alpha}(Lu_h,Lv_h)_H + < Lv_h,\lambda_h >_{W\times W'} \\ \qquad\qquad\qquad\qquad = < f,v_h >_{V'\times V} \quad \forall v_h\in V_h \\ < Lu_h,\mu_h >_{W\times W'} = 0 \qquad\qquad\qquad \forall \mu_h\in W_h \ . \end{cases} \end{aligned} \tag{3.21}$$

For formulation (3.21) we will suppose true the following coercivity assumption, which is indeed met in several relevant applications (cf. section 4):

$$\exists\delta>0: \quad a_0(v,v)+\gamma||Lv||_H^2 \geq \delta||v||_V^2 \quad \forall v\in V \ , \tag{3.22}$$

with $\gamma>0$.

Remark 3.4 Note that, since the norm in H can be weaker than the one in W, assumption (3.22) is not implied by the well-posedness of the problem (cf. (3.3)) and it has to be checked case by case.

It is straightforward to show that, under hypotheses (3.22), *i*), *iii*) and *iv*) of (3.2), problem (3.21) admits a unique solution. Furthermore, an error analysis can be developed following similar techniques as those used to prove Theorem 3.2. Nevertheless, some few differences arise. More precisely, inequality (3.6) now becomes

$$\delta||u_h - v_h||_V^2 + h^{-\alpha}||L(u_h - v_h)||_H^2 \leq T_1 + T_2 + T_3 \tag{3.23}$$

where

$$\begin{aligned} T_1 &= 2 < L(u_h - v_h), \lambda - \lambda_h >_{W\times W'} \\ T_2 &= 2a_0(u - v_h, u_h - v_h) \\ T_3 &= 2h^{-\alpha}(L(u - v_h), L(u_h - v_h))_H. \end{aligned}$$

Although T_2 and T_3 can be treated in the same way as in (3.8), leading to

$$\begin{aligned} T_2 \leq &C\Big(\sigma||u_h - v_h||_V^2 + \frac{1}{\sigma}||u - v_h||_V^2\Big) \\ T_3 \leq &C\Big(\tau h^{-\alpha}||L(u_h - v_h)||_H^2 + \frac{h^{-\alpha}}{\tau}||u - v_h||_V^2\Big)\ , \end{aligned} \tag{3.24}$$

for T_1 one has to be particularly careful. Indeed, we have

$$T_1 \leq \varepsilon h^{-\alpha}||L(u_h - v_h)||_W^2 + \frac{h^\alpha}{\varepsilon}||\lambda - \mu_h||_{W'}^2. \tag{3.25}$$

Unfortunately, since the norm in W may be stronger than the one in H, the term $\varepsilon h^{-\alpha}||L(u_h - v_h)||_W^2$ cannot be absorbed into the left-hand side of inequality (3.23).

A possible way out consists in supposing $\lambda \in H$. This is indeed a reasonable regularity assumption (see Remark 3.8).

In this case we have

$$< L(u_h - v_h), \lambda - \lambda_h >_{W\times W'}= (L(u_h - v_h), \lambda - \mu_h)_H,$$

so that the following estimate holds

$$T_1 \leq \varepsilon h^{-\alpha}||L(u_h - v_h)||_H^2 + \frac{h^\alpha}{\varepsilon}||\lambda - \mu_h||_H^2. \tag{3.26}$$

Consequently, the error estimate (3.10) becomes, in the present case,

$$||u - u_h||_V \leq C\Big((1 + h^{-\alpha/2}) \inf_{v_h\in V_h} ||u - v_h||_V + h^{\alpha/2} \inf_{\mu_h\in W_h} ||\lambda - \mu_h||_H\Big)\ . \tag{3.27}$$

The error analysis concerning the multiplier follows the lines of that for formulation (3.1). We get the analogous of theorem 3.2 for this formulation.

Theorem 3.5 *Assume hypotheses (3.2) and (3.22). Let* (u, λ) *be the solution of problem (2.1) and let* (u_h, λ_h) *be the solution of the discrete problem (3.21). Assume moreover that* $\lambda \in H$*. Then the following error estimates hold*

(3.28)
$$||u-u_h||_V \leq C\Big((1+h^{-\alpha/2}) \inf_{v_h\in V_h} ||u-v_h||_V + h^{\alpha/2} \inf_{\mu_h\in W_h} ||\lambda-\mu_h||_H\Big)$$

$$||\lambda-\lambda_h||_{W'} \leq C\Big((1+h^{-\alpha/2}+h^{-\alpha}) \inf_{v_h\in V_h} ||u-v_h||_V + \inf_{\mu_h\in W_h} ||\lambda-\mu_h||_H\Big) .$$

Furthermore, it is straightforward to obtain, for formulation (3.21), the following

Corollary 3.6 *Suppose that the choice of V_h and W_h satisfies*

$$\inf_{v_h\in V_h} ||u-v_h||_V = O(h^{k+\beta}), \quad \inf_{\mu_h\in W_h} ||\lambda-\mu_h||_H = O(h^k),\ k>0,\ \beta\geq 0 .$$

Then the best rate of convergence is met for $\alpha=\beta$. In this case we obtain

$$\begin{aligned} ||u-u_h||_V &= O(h^{k+\beta/2}) \\ ||\lambda-\lambda_h||_{W'} &= O(h^k) . \end{aligned} \tag{3.29}$$

Remark 3.7 We observe that estimate (3.28) is not optimal, since the W'–norm of $\lambda-\lambda_h$ is controlled by the H–norm of the interpolation error. Nonetheless this result can be useful in some applications, such as for instance the Reissner–Mindlin plate problem and its extension to Naghdi shells (cf. section 4).

More precisely, let us compare $\inf_{W_h} ||\lambda-\mu_h||_H$ and $\inf_{W_h} ||\lambda-\mu_h||_{W'}||$ for some choice of W_h. The spaces involved in the plate problem are

$$\begin{aligned} &W = \mathrm{H}_0(\mathrm{rot};\Omega) = \big\{\mathbf{u}\in \big(\mathrm{L}^2(\Omega)\big)^2 |\ \mathrm{rot}\,\mathbf{u}\in \mathrm{L}^2(\Omega), \mathbf{u}\cdot\mathbf{t}=0 \text{ on } \partial\Omega\big\}, \\ &H = \big(\mathrm{L}^2(\Omega)\big)^2. \end{aligned} \tag{3.30}$$

Hence (cf. Brezzi–Fortin 1991) the dual space of W is

$$W' = \mathrm{H}^{-1}(\mathrm{div};\Omega) = \big\{\boldsymbol{\gamma}\in \big(\mathrm{H}^{-1}(\Omega)\big)^2 |\ \mathrm{div}\,\boldsymbol{\gamma}\in \mathrm{H}^{-1}(\Omega)\big\}. \tag{3.31}$$

Choosing $W_h = \{\text{piecewise constant vector functions}\}$, one can realize that both $\inf_{W_h} ||\lambda-\mu_h||_H$ and $\inf_{W_h} ||\lambda-\mu_h||_{W'}$ are $O(h)$. Similarly the choice $W_h = \{\text{discontinuous piecewise linear vector functions}\}$ provides $O(h^2)$ for both cases. These examples show that in some cases having in the right-hand side of estimate (3.28) the H norm instead of the W' one, does not change the scheme rate of convergence, even though the norms of H and W' are not equivalent.

On the other hand, if one selects, for instance, $W_h = RT_1$ (cf. Brezzi–Fortin 1991), we have $\inf_{W_h} ||\lambda-\mu_h||_H = O(h^2)$, while $\inf_{W_h} ||\lambda-\mu_h||_{W'} = O(h^3)$.

Remark 3.8 It is obvious that if $W=H$ (cf. examples 1 and 2 in section 4), formulation (3.1) is the same as (3.21).

In the first two examples proposed in the following section, it turns out that $W=H$, so that both formulations give the same result. On the other hand, in the third example (cf. section 4), we have that $W=\mathrm{H}_0(\mathrm{rot})$ and $H=\mathrm{L}^2$. In order to obtain the convergence result of theorem 3.5, we have to assume that the multiplier λ is in L^2, instead of in $\mathrm{H}^{-1}(\mathrm{div})$, which is indeed a weak regularity assumption.

In any case, to achieve a good error estimate (in terms of powers of h) one should require, on λ, much more regularity than L^2.

4. Applications

In this section we present some classical problem for which our analysis could be of some interest. For most of these problems, namely the Stokes problem, the Laplace equation, and the Reissner–Mindlin plate, performant and optimal convergent mixed schemes have already been proposed (cf. Bathe-Wilson (1976), Hughes (1987) and Zienkiewicz-Taylor (1989)). However, one could sometimes prefer to employ "not well-balanced" methods because, for instance, one has in mind other more complex applications such as Naghdi shells (cf. example 3). In a sense, the examples described below have to be intended as "simplified models" of other actual and more interesting problems.

In examples 1 and 2, since $W = H$, we will use formulation (3.1) which is indeed equivalent to (3.21) (cf. remark 3.8). On the other hand, in example 3 the space W and H are not the same and we will describe formulation (3.21), altough formulation (3.1) would not be much more expensive.

Example 1. The Stokes problem
Let us consider the velocity-pressure formulation of the Stokes problem

$$\text{find } (\mathbf{u}, p) \text{ such that}$$
$$(4.1)\qquad \begin{cases} -\nu \Delta \mathbf{u} + \nabla p = \mathbf{f} & \text{in } \Omega\ , \\ \operatorname{div} \mathbf{u} = 0 & \text{in } \Omega\ , \\ \mathbf{u} = 0 & \text{on } \partial\Omega\ , \end{cases}$$

where Ω is a polygonal domain in $\mathbf{R}^n$ and $\partial\Omega$ its boundary.

It is well-known that problem (3.1) can be stated in the form (2.1) with the following choices:

$$(4.2)\qquad \begin{aligned} & V = \mathrm{H}^1_0(\Omega)^n \\ & W = H = \mathrm{L}^2(\Omega)/\mathbf{R} \\ & a_0(\mathbf{v}, \mathbf{w}) = 2\nu \textstyle\int_\Omega \varepsilon(\mathbf{v}) : \varepsilon(\mathbf{w})\, d\mathbf{x} \quad \mathbf{v}, \mathbf{w} \in V \\ & L\mathbf{v} = \operatorname{div} \mathbf{v}\ . \end{aligned}$$

To fix ideas, let n be equal to 2. Suppose now to approximate problem (4.1) by means of triangular finite elements; let V_h consists of continuous piecewise quadratic polynomials and W_h of piecewise constants. Thanks to the boundary condition, it easily seen that the bilinear form $a_0(\cdot, \cdot)$ is elliptic on the whole space V; on the other hand our choice for the spaces V_h and W_h satisfies the compatibility condition *iv*) of (3.2) (cf. e.g. Brezzi-Fortin (1991)) so that the theory of the previous section can be succesfully applied. The approximating properties of V_h and W_h are the following:

$$\inf_{v_h \in V_h} ||u - v_h||_1 = O(h^2), \quad \inf_{q_h \in W_h} ||p - q_h||_0 = O(h^1) \; . \tag{4.3}$$

It follows from Corollary 3.3 that the optimal choice for α is 1; in this case the new augmented scheme (3.1) reads as follows

$$\text{find } (\mathbf{u}_h, p_h) \in V_h \times W_h \text{ such that}$$
$$\begin{cases} a_0(\mathbf{u}_h, \mathbf{v}_h) + h^{-1}(\operatorname{div} \mathbf{u}_h, \operatorname{div} \mathbf{v}_h)_0 & \\ \qquad\qquad + (\operatorname{div} \mathbf{v}_h, p_h)_0 = (\mathbf{f}, \mathbf{v}_h)_0 & \forall \mathbf{v}_h \in V_h \\ & \\ (\operatorname{div} \mathbf{u}_h, q_h)_0 = 0 & \forall q_h \in W_h \end{cases} \tag{4.4}$$

and the convergence rate for the velocity $\mathbf{u}$ is $3/2$ instead of 1.

On the other hand, if one approximates the pressure space by means of continuous piecewise linear functions (Hood–Taylor method), then the augmented Lagrangian method cannot improve the convergence rate. In this case, in fact, estimate (2.6) shows that $\mathbf{u}$ is already approximated with the optimal order 2.

Example 2. The Laplace equation
It is well-known that the Laplace equation

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ & \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{4.5}$$

admits the following mixed variational formulation

$$\text{find } (\boldsymbol{\sigma}, u) \in V \times H \text{ such that}$$
$$\begin{cases} a_0(\boldsymbol{\sigma}, \boldsymbol{\tau}) + < L\boldsymbol{\tau}, u >_{W \times W'} = 0 & \forall \boldsymbol{\tau} \in V \\ & \\ - < L\boldsymbol{\sigma}, v >_{W \times W'} = (f, v)_{W'} & \forall v \in W' \end{cases} \tag{4.6}$$

with the choices

$$\begin{aligned} & V = \mathrm{H}(\operatorname{div}, \Omega), \\ & W = H = \mathrm{L}^2(\Omega), \\ & a_0(\boldsymbol{\sigma}, \boldsymbol{\tau}) = (\boldsymbol{\sigma}, \boldsymbol{\tau})_0 \\ & L\boldsymbol{\tau} = \operatorname{div} \boldsymbol{\tau} \; . \end{aligned} \tag{4.7}$$

In this case the bilinear form $a_0(\cdot, \cdot)$ is not coercive on the whole space V and finding finite element spaces meeting conditions (2.4) and (2.5) with α_h and β_h uniformly bounded away from zero can be not an easy matter (cf. Brezzi and Fortin 1991, in which several methods are proposed, such as RT_k and BDM_k). Therefore the theory of the augmented Lagrangian formulation can be useful (cf. Brezzi et al. 1994) in particular if we are interested in approximations with continuous stresses; indeed the form $a_0(\cdot, \cdot) + (L\cdot, L\cdot)_W$ is elliptic in V and we only have to verify the inf-sup condition *iv*) of (3.2).

As in the example of the Stokes problem, if we choose the space V_h consisting of continuous piecewise quadratic polynomials and W_h consisting of piecewise constants, then the best value for α is 1 and the rate of convergence for $\boldsymbol{\sigma}$ is improved from 1 to 3/2.

Example 3. The Reissner-Mindlin plate problem and Naghdi shells
Another example for which our result can be useful is the so-called Reissner-Mindlin plate problem. The Reissner-Mindlin model describes the small deformation of a linear elastic plate in terms of the rotations $\boldsymbol{\theta}$ of the fibers and of deflection w of the middle-surface Ω, which we suppose to be simply connected. An augmented mixed variational formulation, recently proposed in Arnold and Brezzi (1993), in the case of a clamped plate, reads as follows

$$\text{find } \big((\boldsymbol{\theta}, w), \boldsymbol{\gamma}\big) \in V \times H \text{ such that}$$
$$\begin{cases} a_0(\boldsymbol{\theta}, w; \boldsymbol{\eta}, v) + c(L(\boldsymbol{\theta}, w), L(\boldsymbol{\eta}, v))_H \\ \qquad\qquad + (L(\boldsymbol{\eta}, v), \boldsymbol{\gamma})_H = (f, v)_H \quad \forall (\boldsymbol{\eta}, v) \in V \\ (L(\boldsymbol{\theta}, w), \boldsymbol{\xi})_H - \dfrac{t^2}{(1 - ct^2)}(\boldsymbol{\gamma}, \boldsymbol{\xi})_H = 0 \quad \forall \boldsymbol{\xi} \in H \ , \end{cases} \tag{4.8}$$

where $t > 0$ is the thickness of the plate and c is a constant such that $0 < c < t^{-2}$. Furthermore

$$\begin{aligned} &V = \big(\mathrm{H}^1_0(\Omega)\big)^2 \times \mathrm{H}^1_0(\Omega) \\ &W = \mathrm{H}_0(\mathrm{rot}; \Omega) = \big\{\mathbf{u} \in \big(\mathrm{L}^2(\Omega)\big)^2 | \operatorname{rot} \mathbf{u} \in \mathrm{L}^2(\Omega), \mathbf{u} \cdot \mathbf{t} = 0 \text{ on } \partial\Omega\big\} \\ &H = \big(\mathrm{L}^2(\Omega)\big)^2 \\ &a_0(\boldsymbol{\theta}, w; \boldsymbol{\eta}, v) = a(\boldsymbol{\theta}, \boldsymbol{\eta}) \\ &L(\boldsymbol{\eta}, v) = \boldsymbol{\eta} - \nabla v \ , \end{aligned} \tag{4.9}$$

where $a(\cdot, \cdot)$ is an elliptic bilinear form over $\big(\mathrm{H}^1_0(\Omega)\big)^2$. Note that the dual space of W is $\mathrm{H}^{-1}(\mathrm{div}; \Omega) = \big\{\boldsymbol{\gamma} \in \big(\mathrm{H}^{-1}(\Omega)\big)^2 | \operatorname{div} \boldsymbol{\gamma} \in \mathrm{H}^{-1}(\Omega)\big\}$, that the image of L in H is $W = \mathrm{H}_0(\mathrm{rot}; \Omega)$ and that $L : V \to W$ is continuous (cf. Brezzi and Fortin 1991). Moreover, note that, as t tends to zero, formulation (4.8) does not degenerate and it leads to a "limit problem" of the form

$$\text{find } \big((\boldsymbol{\theta}, w), \boldsymbol{\gamma}\big) \in V \times W' \text{ such that}$$
$$\begin{cases} a_0(\boldsymbol{\theta}, w; \boldsymbol{\eta}, v) + c(L(\boldsymbol{\theta}, w), L(\boldsymbol{\eta}, v))_H \\ \qquad\qquad + < L(\boldsymbol{\eta}, v), \boldsymbol{\gamma} >_{W \times W'} = (f, v)_H \quad \forall (\boldsymbol{\eta}, v) \in V \\ < L(\boldsymbol{\theta}, w), \boldsymbol{\xi} >_{W \times W'} = 0 \qquad\qquad \forall \boldsymbol{\xi} \in H \ . \end{cases} \tag{4.10}$$

We restrict ourselves to the analysis of the discretization of (4.10), considering (as reasonable) that a performant method for the limit problem (4.10) will lead to a good method for the approximation of problem (4.8) as well.

Note that problem (4.10) is in the framework of formulation (3.21).

As an example of how the theory of the previous section can be applied to this problem, consider the following approximation proposed in Lovadina (1994). Choose V_h consisting of continuous piecewise quadratic polynomials for both rotations and deflections, while set W_h as piecewise constant functions. Such formulation simplifies the original version proposed by Arnold and Brezzi

(1993) in which bubble functions where added to the rotations. It has been shown (cf. Lovadina 1994) that this method is first order convergent; however, the choice $c = h^{-1}$ allows to improve the rate of convergence for rotations and vertical displacements up to 3/2 (cf. Corollary 3.6).

A similar augmented formulation as in (4.8) has been proposed and studied for the case of the Naghdi shells (see Arnold and Brezzi (1994)); we will not detail this formulation, since it would involve the introduction of cumbersome notation. We only recall that the proposed approximation consists of the Lagrange finite elements of second degree augmented by bubble functions of third degree for the primal variables (rotations and displacements) and piecewise constants for the multipliers. Hence Arnold and Brezzi were able to prove that the method is first order convergent; if we apply our modification choosing $\alpha = 1$, the rate of convergence is improved up to 3/2.

We conclude by remarking that, while for the other examples described in this section some optimal method has already been proposed and analyzed, for Naghdi shells this is not the case. Indeed, even if many other schemes for shells have already been presented (cf. Bathe and Wilson (1976), Hughes (1987), Zienkiewicz and Taylor (1989)), for none of them a rigorous mathematical analysis exists yet. That's why, in our opinion, the application to shell problems of the new augmented Lagrangian formulation turns out to be especially interesting.

References

Arnold, D.N., Brezzi, F. (1993): Some new elements for the Reissner-Mindlin plate model. In: Lions, J.L., Baiocchi, C. (eds.) Boundary value problems for partial differential equations and applications

Arnold, D.N., Brezzi, F. (1994): Locking free finite elements for shells. To appear

Bathe, K.-J., Wilson, E. (1976): Numerical methods in finite elements analysis. Englewood Cliffs, NJ

Brezzi, F., Fortin, M. (1991): Mixed and Hybrid Finite Element Methods. Springer-Verlag

Brezzi, F., Fortin, M., Marini, L.D. (1993): Mixed finite element methods with continuous stresses. Math. Mod. Meth. in App. Sc. **3**(2), 275–287

Ciarlet, P.G. (1978): The Finite Element Method for Elliptic Equations. North-Holland

Fortin, M., Glowinski, R. (1983): Augmented Lagrangian Methods. North-Holland

Girault, V., Raviart, P.-A. (1986): Finite Element Methods for the Navier-Stokes Equations. Springer Verlag

Hughes, T.J.R. (1987): The finite element method. Englewood Cliffs, NY

Lovadina, C. (1994): A new class of mixed finite element methods for Reissner-Mindlin plates. SIAM J. Numer. Anal., to appear

Zienkiewicz, O.C., Taylor, R.L. (1989): The finite element methods. McGraw-Hill, New York