J. Fac. Sci. Univ. Tokyo
Sect. IA, Math.
32 (1985), 499-538.

An abstract analysis of parameter dependent problems and its applications to mixed finite element methods

By Fumio KIKUCHI

Summary. This paper is concerned with numerical analysis of a class of parameter dependent problems. Such problems arise, for example, from the analysis of nearly incompressible media and from the penalization of some constrained minimization problems. We first consider the solvability and the uniqueness of an abstract problem under certain hypotheses, and then analyze the dependence of the solutions on the parameter. We also propose an iteration scheme, by which we can obtain solutions for fairly wide range of parameter values. We introduce a mixed finite element approximation to the present problem, and discuss its approximation properties. In particular, we consider the locking phenomenon, which often arises from numerical approximation of the present type of problems. We also give some simple numerical results to see the dependence of numerical solutions on the parameter. In Appendices, we present some sufficient conditions for the hypotheses employed to assure the solvability of the present problem and its approximations.

Keywords. Mixed finite element method, Parameter dependent problem, Penalty approach, Iteration scheme, Locking

Contents

1. Introduction

In the analysis of an incompressible or nearly incompressible continuum occupying a bounded domain Ω of $\boldsymbol{R}^3$ with boundary $\partial\Omega$, we have a parameter dependent problem of the form: given a vector function $f=\{f_1, f_2, f_3\}$, find another vector function $u=\{u_1, u_2, u_3\}$ and a scalar function p such that

$$\left.\begin{aligned} &-\Delta u-\frac{1}{3}\operatorname{grad}(\operatorname{div} u)+\operatorname{grad} p=f \\ &-\operatorname{div} u-\varepsilon p=0 \end{aligned}\right\} \quad \text{in } \Omega, \quad u=0 \text{ on } \partial\Omega,$$

where Δ, div, and grad are the usual differential operators of $\boldsymbol{R}^3$, and ε is a physical or artificial parameter with non-negative values. The functions f and p are defined in Ω, while u should be considered over $\bar{\Omega}$ (=closure of Ω). Physically, u stands for the velocity or the displacement of the continuum, p means the pressure, and f is the applied body force. The value of ε is almost equal to zero for nearly incompressible continua, and is equal to zero for completely incompressible ones. When the continuum is fluid, the first of the above partial differential equations is called the Stokes equation.

Parameter dependent problems similar to the above appear in various fields, and are usually solved by numerical methods such as the finite element method. However, such analysis is often accompanied with various kinds of troubles especially when ε is close to zero. For example, it is not necessarily easy to find numerical models with nice approximation properties: inappropriate models may yield highly oscillating or unbelievably small numerical solutions. We also have difficulty in numerically solving the obtained approximate equations for small ε due to unavoidable round-off errors in computers. Thus there have been performed many studies concerning the present type of problems from both theoretical and practical standpoints. In what follows, we will give a short review of some theoretical works.

Historically, analysis was first made in the case of $\varepsilon=0$. Then a natural approach to finite element approximation of the present type of problems exists, and is known as the mixed method or the hybrid one, see Atluri, Gallagher, and Zienkiewicz [2]. Babuska [3], Brezzi [5], and Kikuchi [12] gave sufficient conditions to assure the stability and the convergence of such an approach. When this approach is applied to the incompressible continuum equation, both u and p are dealt with as independent unknown quantities. For the two-dimensional Stokes equation with the incompressibility condition, Crouzeix and Raviart [8] presented a variety of mixed finite element models that satisfy both stability and convergence conditions. The mixed method is also available for $\varepsilon>0$. In fact, Bercovier [4] presented a penalized version of the incompressible continuum equation, which is essentially of the form stated above. In his approach, ε is nothing but the penalty parameter and is rather an artificial number. The only interesting value of ε is zero, but, by using ε with a (small) positive value, we have $p=-\varepsilon^{-1}\operatorname{div} u$ and hence we can obtain an equation in u only:

$$-\Delta u-\left(\frac{1}{3}+\frac{1}{\varepsilon}\right)\operatorname{grad}(\operatorname{div} u)=f \text{ in } \Omega, \quad u=0 \text{ on } \partial\Omega.$$

He also discussed some finite element analogs of the penalized problem. Arnold [1] considered a model equation related to the Timoshenko beam problem, which has essentially the same properties as the one above. In this case, ε is a physical parameter, and it is practically meaningful to obtain solutions for ε other than zero. He stressed the importance of the uniformity or the robustness of approximations with respect to the parameter. Kikuchi [13] also considered an equation related to circular arch problems, and generalized the results of Arnold by means of the asymptotic expansion. The last three works also dealt with the so-called reduced integration, which is a technique of obtaining better approximations by lowering the order of the integration rule used in the evaluation of some terms in the approximate equations based on the finite element method: see also [15] and [16].

This work is a continuation of the one by the present author [13], and improves the results obtained there. First, we consider a parameter dependent problem in a product Hilbert space as a generalization of the above problem. Then we derive some analytical results for the solutions of the problem. We show the uniqueness and existence of the solution for each $\varepsilon\geqq 0$ under essentially the same hypotheses as are employed for $\varepsilon=0$. Moreover, we can estimate the norms of the solutions uniformly in $\varepsilon \geqq 0$ (Lemma 1 in section 2), and prove that the solutions can be expanded as a power series with respect to ε (Theorem 1 in section 3). An iterative version of the expansion is also presented with another theorem on the convergence of the iteration process (Theorem 2 in section 4). Its relation to the augmented Lagrangian method is also noted: see Fortin and Glowinski [10]. Then the mixed finite element method is introduced and error analysis is performed. It turns out that the numerical solutions based on such a method can approximate well the exact solutions under certain hypotheses that are commonly employed for $\varepsilon=0$ (Theorems 3 and 4 of section 6). The error analysis is made also in the case where some of the hypotheses are lacking. Especially, we consider the locking phenomenon, where numerical solutions completely fail in approximating certain solutions of the original problem when ε is close to zero. We also give some numerical examples for the Timoshenko beam problem to see the validity of the present analysis. In particular, we observe the dependence of the numerical solutions on the parameter. The effectiveness of the proposed iteration scheme is also discussed with these examples. We will elsewhere report numerical experiment of this approach for the Stokes equation. In Appendices, we summarize some results for stability conditions and approximate properties of the finite element method, although some of them may be already used in particular cases. See also Fortin [9] for related results.

This work is in part supported by the Grant-in-Aid from the Ministry of Education.

2. Preliminaries: Statement of an abstract problem with some basic results

The norm of a Banach space X is denoted by $\|\cdot\|_X$, and, when X is a Hilbert space, its inner product is designated by $(\cdot,\cdot)_X$. For two Banach spaces X and Y, $\mathcal{L}(X, Y)$ implies the Banach space of all linear bounded operators from X into Y. The norm of $\mathcal{L}(X, Y)$ should be conventionally written by $\|\cdot\|_{\mathcal{L}(X,Y)}$, but is often abbreviated as $\|\cdot\|$ if there exists no fear of confusion. The null space and the range of an operator $T\in\mathcal{L}(X, Y)$ are denoted by $N(T)$ and $R(T)$, respectively. For $T\in\mathcal{L}(X, Y)$ and a subset Z of X, $T|Z$ implies the restriction of T to Z.

Let V and W be real Hilbert spaces. We consider two bounded bilinear forms:

$$a(\cdot,\cdot): V\times V\to \boldsymbol{R}^1, \quad b(\cdot,\cdot): V\times W\to \boldsymbol{R}^1,$$

where $\boldsymbol{R}^1$ is the set of all real numbers. We assume that $a(\cdot,\cdot)$ is non-negative in the sense that

$$a(u, u)\geqq 0 \qquad \text{for all } u\in V. \tag{1}$$

Due to the Riesz representation theorem [18], we can define $A\in\mathcal{L}(V, V)$, $B\in\mathcal{L}(V, W)$, and $B^*\in\mathcal{L}(W, V)$ by

$$(Au, v)_V = a(u, v) \qquad \text{for all } u, v\in V, \tag{2}$$

$$(Bu, \lambda)_W = (u, B^*\lambda)_V = b(u, \lambda) \qquad \text{for all } u\in V,\ \lambda\in W. \tag{3}$$

Let ε_0 be a (small) positive constant, and consider the problem: given $\{f, g\}\in V\times W$ and $\varepsilon\in[0, \varepsilon_0]$, find $\{u_\varepsilon, \lambda_\varepsilon\}\in V\times W$ such that

$$\begin{cases} a(u_\varepsilon, v)+b(v, \lambda_\varepsilon)=(f, v)_V\,;\ \forall v\in V, \\ b(u_\varepsilon, \mu)-\varepsilon(\lambda_\varepsilon, \mu)_W=(g, \mu)_W\,;\ \forall\mu\in W, \end{cases} \tag{4}$$

or, equivalently,

$$Au_\varepsilon + B^*\lambda_\varepsilon = f, \quad Bu_\varepsilon - \varepsilon\lambda_\varepsilon = g. \tag{5}$$

We will regard the above as a parameter dependent problem with the parameter ε chosen from $[0, \varepsilon_0]$. In order that the present problem is valid for $\varepsilon=0$, g must belong to $R(B)$ since the second equation of (5) gives $g=Bu_\varepsilon\in R(B)$ for $\varepsilon=0$. When $g\in R(B)$, λ_ε is necessarily in $R(B)$ for $\varepsilon\neq 0$ due to the relation $\lambda_\varepsilon=\varepsilon^{-1}(Bu_\varepsilon-g)\in R(B)$ derived from (5). For $\varepsilon=0$, λ_ε is indefinite in its component in $N(B^*)$. If $R(B)$ is closed, then $R(B)$ is the orthogonal complement of $N(B^*)$ in W, and hence it is sufficient to look for λ_ε in $R(B)$ even for $\varepsilon=0$. We will only consider the case where

$$B\neq 0, \quad N(B)\neq\{0\}. \tag{6}$$

Hereafter, we will analyze the above problem. For its solvability conditions, the following two are wellknown; see Bercovier [4] and Arnold [1].

[H1] *There exists a positive constant* k_1 *such that*

$$a(u, u)\geqq k_1\|u\|_V^2 \quad \textit{for all } u\in N(B). \tag{7}$$

[H2] *There exists a positive constant* k_2 *such that*

$$\|B^*\lambda\|_V \geqq k_2\|\lambda\|_W \quad \textit{for all } \lambda \in R(B). \tag{8}$$

When [H2] holds, it is clear that $R(B^*)$ is closed. Then, by the closed range theorem [18], $R(B)$ is also closed, and [H2] is equivalent to the following condition.

[H2]* *There exists a positive constant* k_2 *such that*

$$\|Bu\|_W \geqq k_2\|u\|_V \quad \textit{for all } u \in R(B^*). \tag{9}$$

Notice here that the positive constant k_2 in [H2]* is identical to that in [H2].

Moreover, [H1] is equivalent to the following when [H2] and (1) hold, as is proved in *Appendix 1.*

[H1]* *There exists a positive constant* k_3 *such that*

$$a(u, u) + \|Bu\|_W^2 \geqq k_3\|u\|_V^2 \quad \textit{for all } u \in V. \tag{10}$$

Note that k_3 is expressed by a continuous function of k_1, k_2, and $\|A\|_{\mathcal{L}(V,V)}$. Sometimes, [H1]* is easier to deal with than [H1].

For later use in error estimation, let us consider a problem slightly more general than (5): given $\{f, g, g^*\} \in V \times R(B) \times W$ and $\varepsilon \in [0, \varepsilon_0]$, find $\{u_\varepsilon, \lambda_\varepsilon\} \in V \times W$ such that

$$Au_\varepsilon + B^*\lambda_\varepsilon = f, \quad Bu_\varepsilon - \varepsilon\lambda_\varepsilon = g + \varepsilon g^*. \tag{11}$$

Clearly, (11) reduces to (5) when $g^* = 0$. Under two hypotheses [H1] and [H2], we can obtain the following lemma for (11).

LEMMA 1. *Assume that* [H1] *and* [H2] *hold, and consider problem* (11) *with* $\{f, g, g^*\} \in V \times R(B) \times W$ *and* $\varepsilon \in [0, \varepsilon_0]$ *given. For* $\varepsilon = 0$, *there exists a solution* $\{u_\varepsilon, \lambda_\varepsilon\}$ *which is unique in* $V \times R(B)$ *(but may not be so when it is considered in* $V \times W$*). Similarly, for* $\varepsilon \in]0, \varepsilon_0]$, *there exists a unique solution in* $V \times W$, *again denoted by* $\{u_\varepsilon, \lambda_\varepsilon\}$, *where* λ_ε *belongs to* $R(B)$ *if* $g^* = 0$. *In both cases,* $\{u_\varepsilon, \lambda_\varepsilon\}$ *satisfies*

$$\|u_\varepsilon\|_V + \|\lambda_\varepsilon\|_W \leqq C_1(\|f\|_V + \|g\|_W + \|g^*\|_W), \tag{12}$$

where C_1 *is a positive number dependent continuously on* k_1, k_2, ε_0, $\|A\|_{\mathcal{L}(V,V)}$ *and* $\|B\|_{\mathcal{L}(V,W)} = \|B^*\|_{\mathcal{L}(W,V)}$ *only, and hence is independent of* $\varepsilon \in [0, \varepsilon_0]$. *Of course, the term* $\|g^*\|_W$ *in* (12) *can be omitted for* $\varepsilon = 0$.

REMARK. In Bercovier [4], and Girault and Raviart [11], [H1]* is employed as an additional hypothesis to prove the present lemma for $g^*=0$. It also plays an essential role in Lions [14]. As is proved in Appendix, it can be derived from [H1] and [H2]. This is mainly due to the non-negativeness of $a(\cdot, \cdot)$ stated in (1). In fact, Arnold [1] proved Lemma 1 for $g^*=0$ by the use of (1), [H1], and [H2] only. If we are interested only in the case where $\varepsilon=0$, we need neither (1) nor [H1]*: see Brezzi [5].

PROOF. See Brezzi [5] for $\varepsilon=0$. For $\varepsilon>0$ with $g^*=0$, refer to Arnold [1], Bercovier [4], and Girault-Raviart [11]. We will only sketch the proof for $\varepsilon>0$ with $g^*\neq 0$.

We have from (11) that

$$\varepsilon A u_\varepsilon + B^* B u_\varepsilon = \varepsilon f + B^* g + \varepsilon B^* g^*, \tag{a}$$

$$\varepsilon \lambda_\varepsilon = B u_\varepsilon - g - \varepsilon g^*. \tag{b}$$

Due to [H1]*, we can apply the Lax-Milgram lemma [18] to (a) to show the existence and the uniqueness of $u_\varepsilon \in V$ for $\varepsilon>0$. Then, by using (b), we can also prove the existence and the uniqueness of $\lambda_\varepsilon \in W$ for $\varepsilon>0$.

Decompose λ_ε and g^* as

$$\lambda_\varepsilon = \lambda_{\varepsilon 1} + \lambda_{\varepsilon 2}\,;\ \ \lambda_{\varepsilon 1} \in N(B^*),\ \ \lambda_{\varepsilon 2} \in R(B),$$
$$g^* = g_1^* + g_2^*\,;\ \ g_1^* \in N(B^*),\ \ g_2^* \in R(B).$$

The above decompositions are unique due to the orthogonality between $N(B^*)$ and $R(B)$. Substituting these into (11), we have

$$A u_\varepsilon = f - B^* \lambda_{\varepsilon 2},\ \ B u_\varepsilon = \varepsilon \lambda_{\varepsilon 2} + g + \varepsilon g_2^*,\ \ -\lambda_{\varepsilon 1} = g_1^*.$$

The last relation of the above gives

$$\|\lambda_{\varepsilon 1}\|_W = \|g_1^*\|_W, \tag{c}$$

while the first two and (1) yield

$$\begin{aligned}
&(A u_\varepsilon, u_\varepsilon)_V + \|B u_\varepsilon\|_W^2 \\
&\quad = ((1+\varepsilon) f + B^* g + \varepsilon B^* g_2^*, u_\varepsilon)_V - (\lambda_{\varepsilon 2}, g + \varepsilon g_2^*)_W - \varepsilon (A u_\varepsilon, u_\varepsilon)_V - \varepsilon \|\lambda_{\varepsilon 2}\|_W^2 \\
&\quad \leqq ((1+\varepsilon) f + B^* g + \varepsilon B^* g_2^*, u_\varepsilon)_V - (\lambda_{\varepsilon 2}, g + \varepsilon g_2^*)_W.
\end{aligned} \tag{d}$$

Applying [H2] to the relation $B^* \lambda_{\varepsilon 2} = f - A u_\varepsilon \in R(B^*)$, we find

$$\|\lambda_{\varepsilon 2}\|_W \leqq k_2^{-1} (\|f\|_V + \|A\|_{\mathcal{L}(V,V)} \|u_\varepsilon\|_V). \tag{e}$$

From (c), (d), (e), and [H1]*, we can evaluate $\|u_\varepsilon\|_V$ and $\|\lambda_\varepsilon\|_W$ in terms of $\|f\|_V$, $\|g\|_W$, $\|g^*\|_W$, k_2, k_3, ε_0, $\|A\|_{\mathcal{L}(V,V)}$, and $\|B\|_{\mathcal{L}(V,W)}=\|B^*\|_{\mathcal{L}(W,V)}$, if we notice the orthogonality between $N(B^*)$ and $R(B)$. It is now straightforward to obtain (12), and the proof is complete.

When $\varepsilon\neq 0$, the second relation in (5) gives $\lambda_\varepsilon=\varepsilon^{-1}(Bu_\varepsilon-g)$. Substituting this into the first relation in (5), we have a single equation with respect to u_ε:

$$Au_\varepsilon+\varepsilon^{-1}B^*Bu_\varepsilon=f+\varepsilon^{-1}B^*g. \tag{13}$$

The above is equivalently written by

$$a(u_\varepsilon, v)+\varepsilon^{-1}(Bu_\varepsilon, Bv)_W=(f, v)_V+\varepsilon^{-1}(g, Bv)_W\,;\ \forall v\in V. \tag{14}$$

These expressions often appear in the description of actual physical problems; see Lions [14], Arnold [1], and Kikuchi [13]. They are also used frequently in the *penalization* of (4) or (5) for $\varepsilon=0$; see e.g. Bercovier [4] and Girault-Raviart [11]. One of the merits of such an approach is that λ_ε can be dealt with as a subsidiary quantity. Furthermore, λ_ε obtained by the penalty method with $\varepsilon>0$ automatically belongs to $R(B)$, and there exists no need to handle the indefiniteness of λ_ε for $\varepsilon=0$.

3. Dependence of the solution on the parameter

This section is devoted to the analysis of the dependence on ε of the solution $\{u_\varepsilon, \lambda_\varepsilon\}\in V\times R(B)$ to (5). Fix $\varepsilon^*\in[0, \varepsilon_0]$ and try to expand $\{u_\varepsilon, \lambda_\varepsilon\}$ into a (formal) power series

$$u_\varepsilon=\sum_{i=0}^{\infty}(\varepsilon-\varepsilon^*)^i u^{(i)},\quad \lambda_\varepsilon=\sum_{i=0}^{\infty}(\varepsilon-\varepsilon^*)^i\lambda^{(i)}. \tag{15}$$

Substituting (formally) the above into (5), we can easily see that the coefficients $\{u^{(i)}, \lambda^{(i)}\}\in V\times R(B)$ for $i=0, 1, 2, \cdots$ must satisfy

$$\begin{aligned}&Au^{(0)}+B^*\lambda^{(0)}=f,\quad Bu^{(0)}-\varepsilon^*\lambda^{(0)}=g,\\&Au^{(i)}+B^*\lambda^{(i)}=0,\quad Bu^{(i)}-\varepsilon^*\lambda^{(i)}=\lambda^{(i-1)}\quad (i\geqq 1).\end{aligned} \tag{16}$$

The above is a recursive relation, and Lemma 1 assures the uniqueness and the existence of $\{\{u^{(i)}, \lambda^{(i)}\}\}_{i=0}^{\infty}$. Similar type of expansions at $\varepsilon^*=0$ are analyzed by Lions [14], Temam [17], and Kikuchi [13].

We have the following results for the above expansion.

THEOREM 1. *Assume that* [H1] *and* [H2] *hold. Then, for any* $\{f, g\}\in V\times R(B)$ *and any* $\varepsilon^*\in[0, \varepsilon_0]$, *each of the coefficients* $\{\{u^{(i)}, \lambda^{(i)}\}\}_{i=0}^{\infty}$ *in* (15) *is determined from* (16) *uniquely in* $V\times R(B)$ *and satisfies*

$$\|u^{(i)}\|_V+\|\lambda^{(i)}\|_W\leqq C_1^{i+1}(\|f\|_V+\|g\|_W), \tag{17}$$

where C_1 *is the positive number appearing in* (12). *The series* (15) *converges to the solution* $\{u_\varepsilon, \lambda_\varepsilon\}\in V\times R(B)$ *of* (5) *if* $\varepsilon\in[0, \varepsilon_0]$ *is subject to*

$$|\varepsilon-\varepsilon^*|<1/C_1. \tag{18}$$

Furthermore, for each $\varepsilon\in[0, \varepsilon_0]$, *the finite series*

$$u_\varepsilon^{(i)}=\sum_{j=0}^{i}(\varepsilon-\varepsilon^*)^j u^{(j)}, \quad \lambda_\varepsilon^{(i)}=\sum_{j=0}^{i}(\varepsilon-\varepsilon^*)^j\lambda^{(j)} \quad (i\geqq 0) \tag{19}$$

satisfies the estimation

$$\|u_\varepsilon^{(i)}-u_\varepsilon\|_V+\|\lambda_\varepsilon^{(i)}-\lambda_\varepsilon\|_W\leqq C_1^{i+2}|\varepsilon-\varepsilon^*|^{i+1}(\|f\|_V+\|g\|_W). \tag{20}$$

REMARK. From (18), the series converges for $\varepsilon=0$ if $\varepsilon^*<1/C_1$, which is one of the most important cases in applications. From (17), the convergence of (15) is assured even for ε outside $[0, \varepsilon_0]$ if (18) holds. In such a case, the limit is of course a unique solution of (5), and (20) should be replaced with

$$\|u_\varepsilon^{(i)}-u_\varepsilon\|_V+\|\lambda_\varepsilon^{(i)}-\lambda_\varepsilon\|_W\leqq\frac{C_1^{i+2}|\varepsilon-\varepsilon^*|^{i+1}}{1-C_1|\varepsilon-\varepsilon^*|}(\|f\|_V+\|g\|_W). \tag{21}$$

PROOF. Applying Lemma 1 to (16) with ε equated to ε^*, we can assure the existence and the uniqueness of each $\{u^{(i)}, \lambda^{(i)}\}$ $(i\geqq 0)$ in $V\times R(B)$ with the estimations

$$\|u^{(0)}\|_V+\|\lambda^{(0)}\|_W\leqq C_1(\|f\|_V+\|g\|_W),$$
$$\|u^{(i)}\|_V+\|\lambda^{(i)}\|_W\leqq C_1\|\lambda^{(i-1)}\|_W \quad (i\geqq 1).$$

The first relation of the above is exactly (17) for $i=0$. For $i\geqq 1$, repeated use of the second relation in the above gives

$$\|u^{(i)}\|_V+\|\lambda^{(i)}\|_W\leqq C_1\|\lambda^{(i-1)}\|_W\leqq C_1^2\|\lambda^{(i-2)}\|_W\leqq\cdots\cdots\leqq C_1^i\|\lambda^{(0)}\|_W.$$

With (17) for $i=0$ substituted into the above, we obtain (17) for $i\geqq 1$. It is now clear that the series (15) converges if (18) holds and that the limit gives the unique solution of (5).

From (16), we find that the finite sum (19) satisfies

$$Au_\varepsilon^{(i)}+B^*\lambda_\varepsilon^{(i)}=f,\ \ Bu_\varepsilon^{(i)}-\varepsilon^*\lambda_\varepsilon^{(i)}=g+(\varepsilon-\varepsilon^*)\lambda_\varepsilon^{(i)}-(\varepsilon-\varepsilon^*)^{i+1}\lambda^{(i)}.$$

Subtracting (5) from the above, we have for each $i\geqq 0$ that

$$\begin{aligned}&A(u_\varepsilon^{(i)}-u_\varepsilon)+B^*(\lambda_\varepsilon^{(i)}-\lambda_\varepsilon)=0,\\&B(u_\varepsilon^{(i)}-u_\varepsilon)-\varepsilon(\lambda_\varepsilon^{(i)}-\lambda_\varepsilon)=-(\varepsilon-\varepsilon^*)^{i+1}\lambda^{(i)}.\end{aligned}$$

Applying Lemma 1 to the above with (17) taken into account, we can conclude (20), and the proof is complete.

From the above, we can see the dependence of $\{\lambda_\varepsilon, u_\varepsilon\}$ on ε. Especially, it is infinitely differentiable in $\varepsilon\in[0,\varepsilon_0]$. It is also to be noted that (16) for $\varepsilon^*\neq 0$ can be rewritten by

$$\begin{aligned}&Au^{(0)}+\frac{1}{\varepsilon^*}B^*Bu^{(0)}=f+\frac{1}{\varepsilon^*}B^*g\ ;\ \lambda^{(0)}=\frac{1}{\varepsilon^*}(Bu^{(0)}-g),\\&Au^{(i)}+\frac{1}{\varepsilon^*}B^*Bu^{(i)}=\frac{1}{\varepsilon^*}B^*\lambda^{(i-1)}\ ;\ \lambda^{(i)}=\frac{1}{\varepsilon^*}(Bu^{(i)}-\lambda^{(i-1)}),\end{aligned}\tag{22}$$

where $i=1,2,3,\cdots$. Thus $\{\lambda^{(i)}\}_{i=0}^\infty$ can be dealt with as subsidiary quantities, and we can in a sense take advantage of the merits of the penalty approach. Notice also that there is no need for ε^* to be too close to zero even when we want to obtain solutions with ε almost equal to zero, since the series is convergent so long as (18) is satisfied. The latter fact is practically important since we are likely to have numerical instability when we use too small ε^*.

4. An iteration scheme and its convergence

Fix $\varepsilon^*\in[0,\varepsilon_0]$ and rewrite (5) as

$$Au_\varepsilon+B^*\lambda_\varepsilon=f,\quad Bu_\varepsilon-\varepsilon^*\lambda_\varepsilon=(\varepsilon-\varepsilon^*)\lambda_\varepsilon+g,\tag{23}$$

to which we can consider the following simple iteration scheme:

(i) $\lambda_\varepsilon^{(-1)}=0\in R(B)$, (24)

(ii) for $i=0,1,2,\cdots$, decide $\{u_\varepsilon^{(i)},\lambda_\varepsilon^{(i)}\}\in V\times R(B)$ recursively by

$$Au_\varepsilon^{(i)}+B^*\lambda_\varepsilon^{(i)}=f,\quad Bu_\varepsilon^{(i)}-\varepsilon^*\lambda_\varepsilon^{(i)}=(\varepsilon-\varepsilon^*)\lambda_\varepsilon^{(i-1)}+g.\tag{25}$$

It is easy to see that the above scheme is identical to the expansion (15) with (16) in the sense that the i-th iteration $\{u_\varepsilon^{(i)}, \lambda_\varepsilon^{(i)}\}$ coincides with the finite sum (19) for any $i \geqq 0$. If $\varepsilon^* > 0$, (25) may be rewritten by

$$\begin{aligned} &Au_\varepsilon^{(i)} + \frac{1}{\varepsilon^*} B^* B u_\varepsilon^{(i)} = f + \frac{1}{\varepsilon^*} B^* g + \frac{\varepsilon - \varepsilon^*}{\varepsilon^*} B^* \lambda_\varepsilon^{(i-1)}, \\ &\lambda_\varepsilon^{(i)} = \frac{1}{\varepsilon^*}(Bu_\varepsilon^{(i)} - g) + \frac{\varepsilon^* - \varepsilon}{\varepsilon^*} \lambda_\varepsilon^{(i-1)}. \end{aligned} \tag{26}$$

The left-hand side of the first equation of (26) is exactly of the same form as that of (13) employed in the penalty approach. Thus we can again deal with $\{\lambda_\varepsilon^{(i)}\}_{i=0}^\infty$ as subsidiary quantities, and hence we can take full advantage of such an approach. It is also clear that $\{\lambda_\varepsilon^{(0)}, u_\varepsilon^{(0)}\}$ coincides with the approximate solution obtained by the penalty approach with $\varepsilon = \varepsilon^*$.

Notice here that the above iteration scheme reduces to a special case of the augmented Lagrangian method when $a(\cdot, \cdot)$ is symmetric and $\varepsilon = 0$; see Fortin and Glowinski [10]. Suggested by the convergence proof of such a method, we can obtain another theorem on the convergence of the present iteration scheme.

THEOREM 2. *Assume that* [H1] *and* [H2] *hold. The iteration scheme* (24) *and* (25) *gives a sequence* $\{\{u_\varepsilon^{(i)}, \lambda_\varepsilon^{(i)}\}\}_{i=0}^\infty$ *converging in* $V \times R(B)$ *to the unique solution* $\{u_\varepsilon, \lambda_\varepsilon\}$ *of* (5) *if*

$$0 \leqq \varepsilon \leqq 2\varepsilon^* \qquad (|\varepsilon - \varepsilon^*| \leqq \varepsilon^*). \tag{27}$$

REMARK. Theorem 2 is in a sense complementary to Theorem 1. In fact, the constant C_1 in Theorem 1 is difficult to evaluate in actual problems, while there does not appear such a quantity in Theorem 2: we can a priori make the scheme convergent for the considered ε by choosing ε^* to satisfy (27). On the contrary, when ε^* is smaller than $1/C_1$, Theorem 2 gives less information on convergence than Theorem 1: especially, it yields nothing on convergence rate. It is also to be noted that the starting approximation $\lambda_\varepsilon^{(-1)}$ to λ_ε need not be zero for the convergence of the present iteration scheme, provided that $\lambda_\varepsilon^{(-1)}$ lies in $R(B)$. Furthermore, if $0 < \varepsilon < 2\varepsilon^*$, $\lambda_\varepsilon^{(-1)}$ need not be in $R(B)$, and the corresponding sequence converges to $\{u_\varepsilon, \lambda_\varepsilon\}$ in $V \times W$, as may be seen by carefully checking the following proof.

PROOF. It is clear that $\{u_\varepsilon^{(i)}, \lambda_\varepsilon^{(i)}\}$ exists uniquely in $V\times R(B)$ for each $i\geqq 0$. Define $\{v^{(i)}, \mu^{(i)}\}\in V\times R(B)$ for $i\geqq 0$ by $v^{(i)}=u_\varepsilon^{(i)}-u_\varepsilon$ and $\mu^{(i)}=\lambda_\varepsilon^{(i)}-\lambda_\varepsilon\in R(B)$. Moreover, put $\mu^{(-1)}=-\lambda_\varepsilon\in R(B)$. Then we find for $i\geqq 0$ that

$$Av^{(i)}+B^*\mu^{(i)}=0,\quad Bv^{(i)}-\varepsilon^*\mu^{(i)}=(\varepsilon-\varepsilon^*)\mu^{(i-1)}.$$

For $\varepsilon^*=0$, (27) implies that $\varepsilon=0$, and the conclusion of the theorem is clear. For $\varepsilon^*>0$, we have $\mu^{(i)}=\left(1-\frac{\varepsilon}{\varepsilon^*}\right)\mu^{(i-1)}+\frac{1}{\varepsilon^*}Bv^{(i)}$, and hence

$$\|\mu^{(i)}\|_W^2=\left(1-\frac{\varepsilon}{\varepsilon^*}\right)^2\|\mu^{(i-1)}\|_W^2+\frac{2}{\varepsilon^*}\left(1-\frac{\varepsilon}{\varepsilon^*}\right)(\mu^{(i-1)},\ Bv^{(i)})_W+\left(\frac{1}{\varepsilon^*}\right)^2\|Bv^{(i)}\|_W^2.$$

On the other hand, we find

$$Av^{(i)}+\frac{1}{\varepsilon^*}B^*Bv^{(i)}=-\left(1-\frac{\varepsilon}{\varepsilon^*}\right)B^*\mu^{(i-1)},$$

which gives

$$\left(1-\frac{\varepsilon}{\varepsilon^*}\right)(\mu^{(i-1)},\ Bv^{(i)})_W=-(Av^{(i)},\ v^{(i)})_V-\frac{1}{\varepsilon^*}\|Bv^{(i)}\|_W^2.$$

From the above relations, we obtain

$$\|\mu^{(i)}\|_W^2=\left(1-\frac{\varepsilon}{\varepsilon^*}\right)^2\|\mu^{(i-1)}\|_W^2-\frac{2}{\varepsilon^*}(Av^{(i)},\ v^{(i)})_V-\left(\frac{1}{\varepsilon^*}\right)^2\|Bv^{(i)}\|_W^2. \tag{a}$$

Since $\left(1-\frac{\varepsilon}{\varepsilon^*}\right)^2\leqq 1$ from (27) and $(Av^{(i)}, v^{(i)})_V\geqq 0$ from (1), the above implies that $\|\mu^{(i)}\|_W^2\leqq\|\mu^{(i-1)}\|_W^2$. That is, $\{\|\mu^{(i)}\|_W\}_{i=0}^\infty$ is a non-increasing non-negative sequence and hence has a limit. Then we find from (a) that $(Av^{(i)}, v^{(i)})_V\to 0$ and $\|Bv^{(i)}\|_W\to 0$ $(i\to\infty)$. From [H1]*, this implies that $v^{(i)}=u_\varepsilon^{(i)}-u_\varepsilon\to 0$ in V $(i\to\infty)$. Now the relation $Av^{(i)}+B^*\mu^{(i)}=0$ gives $B^*\mu^{(i)}\to 0$ in V $(i\to\infty)$, which, by [H2], assures that $\mu^{(i)}=\lambda_\varepsilon^{(i)}-\lambda_\varepsilon\to 0$ in $R(B)$ $(i\to\infty)$, since $\mu^{(i)}\in R(B)$ for each $i\geqq 0$ (note that, for $0<\varepsilon<2\varepsilon^*$, (a) directly assures that $\mu^{(i)}\to 0$ in W). Thus we have shown that $\{u_\varepsilon^{(i)}, \lambda_\varepsilon^{(i)}\}\to\{u_\varepsilon, \lambda_\varepsilon\}$ in $V\times R(B)$ $(i\to\infty)$, and the proof is complete.

5. Approximation based on the mixed finite element method

To solve (5) numerically by the finite element method, we prepare a family of spaces $\{V^h\times W^h\}_{h\in\Lambda}$, where the index set Λ is contained in $]0, h_0]$ for a constant $h_0>0$ and has zero as an accumulation point, and, for each $h\in\Lambda$, V^h and W^h are respectively finite-dimensional subspaces of V and W. In actual finite element schemes, h implies the representative element size, and we are usually interested in the asymptotic behaviors of the numerical solutions for $h\downarrow 0$.

For each $h\in\Lambda$, let us introduce two bilinear forms

$$a_h(\cdot,\cdot): V^h\times V^h\to\boldsymbol{R}^1,\quad b_h(\cdot,\cdot): V^h\times W^h\to\boldsymbol{R}^1,$$

which will be used as the approximations to $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$, respectively. They are defined over finite-dimensional spaces, and are necessarily bounded for each $h\in\Lambda$. In the Galerkin method, $a_h(\cdot,\cdot)$ and $b_h(\cdot,\cdot)$ are the restrictions of $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ to $V^h\times V^h$ and $V^h\times W^h$, respectively. However, it is not seldom to adopt more general type of approximations in practical schemes. For example, we can have such general type of $b_h(\cdot,\cdot)$ if we use the selective reduced integration or the partial approximation; see Kikuchi [13], and Oden and Kikuchi [16]. We assume that $a_h(\cdot,\cdot)$ for each $h\in\Lambda$ is non-negative:

$$a_h(u_h, u_h)\geqq 0\qquad\text{for all } u_h\in V^h. \tag{28}$$

It is possible to define $A_h\in\mathcal{L}(V^h, V^h)$, $B_h\in\mathcal{L}(V^h, W^h)$, and $B_h^*\in\mathcal{L}(W^h, V^h)$ by

$$(A_h u_h, v_h)_V=a_h(u_h, v_h)\qquad\text{for all } u_h, v_h\in V^h, \tag{29}$$

$$(B_h u_h, \lambda_h)_W=(u_h, B_h^*\lambda_h)_V=b_h(u_h, \lambda_h)\qquad\text{for all } u_h\in V^h, \lambda_h\in W^h, \tag{30}$$

which will be employed as approximations to A, B, and B^*. We will also use the following orthogonal projection operators:

$$P_h: V\to V^h;\ (P_h u, v_h)_V=(u, v_h)_V\quad\text{for all } u\in V, v_h\in V^h, \tag{31}$$

$$Q_h: W\to W^h;\ (Q_h\lambda, \mu_h)_W=(\lambda, \mu_h)_W\quad\text{for all } \lambda\in W, \mu_h\in W^h. \tag{32}$$

In the Galerkin method, it holds that $A_h=P_hA|V^h$, $B_h=Q_hB|V^h$, and $B_h^*=P_hB^*|W^h$.

Now the determination condition for the approximate solution $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}\in V^h\times W^h$ to the solution $\{u_\varepsilon, \lambda_\varepsilon\}$ of (4) or (5) is

$$\begin{cases} a_h(u_{h\varepsilon}, v_h) + b_h(v_h, \lambda_{h\varepsilon}) = (f, v_h)_V \ ; \ \forall v_h \in V^h, \\ b_h(u_{h\varepsilon}, \mu_h) - \varepsilon(\lambda_{h\varepsilon}, \mu_h)_W = (g, \mu_h)_W \ ; \ \forall \mu_h \in W^h, \end{cases} \tag{33}$$

or, equivalently,

$$A_h u_{h\varepsilon} + B_h^* \lambda_{h\varepsilon} = P_h f, \quad B_h u_{h\varepsilon} - \varepsilon \lambda_{h\varepsilon} = Q_h g, \tag{34}$$

where ε, f, and g are the same as those appearing in (4) or (5). Clearly, $Q_h g$ must belong to $R(B_h)$ for the above to be meaningful for $\varepsilon=0$. Therefore, for the validity of the present problem for all $g \in R(B)$, it is necessary that $Q_h R(B) \subset R(B_h)$.

The present finite element approximation is of mixed or hybrid type since it involves two unknowns $u_{h\varepsilon}$ and $\lambda_{h\varepsilon}$. However, as in the continuous case, we can eliminate $\lambda_{h\varepsilon}$ from (34) for $\varepsilon>0$:

$$A_h u_{h\varepsilon} + \varepsilon^{-1} B_h^* B_h u_{h\varepsilon} = P_h f + \varepsilon^{-1} B_h^* Q_h g, \tag{35}$$

or

$$a_h(u_{h\varepsilon}, v_h) + \varepsilon^{-1}(B_h u_{h\varepsilon}, B_h v_h)_W = (f, v_h)_V + \varepsilon^{-1}(g, B_h v_h)_W \quad \text{for all } v_h \in V^h. \tag{36}$$

Therefore, we can also utilize the penalty approach for solving the discrete problem (34) with $\varepsilon=0$.

6. Analysis of the approximate problem

This section is devoted to establishing the uniqueness, the existence, and the approximation properties of the discrete problem (33) or (34) under certain hypotheses. In order that the approximate problem may be valid, it is natural to require the following approximation condition for $\{V^h \times W^h\}_{h \in \Lambda}$:

$[\mathrm{H0}]_h$ *For each* $\{u, \lambda\} \in V \times W$, *it holds that*

$$\lim_{h \downarrow 0} E_h(u, \lambda) = 0, \tag{37}$$

where, for each $h \in \Lambda$,

$$E_h(u, \lambda) = \inf_{u_h \in V^h} (\|u_h - u\|_V + \|A_h u_h - Au\|_V + \|B_h u_h - Bu\|_W) + \inf_{\lambda_h \in W^h} (\|\lambda_h - \lambda\|_W + \|B_h^* \lambda_h - B^* \lambda\|_V). \tag{38}$$

When $A_h = P_h A | V^h$, $B_h = Q_h B | V^h$, and $B_h^* = P_h B^* | W^h$, it is clear that (37) may be replaced with

$$\lim_{h \downarrow 0} \left(\inf_{u_h \in V^h} \|u_h - u\|_V + \inf_{\lambda_h \in W^h} \|\lambda_h - \lambda\|_W \right) = 0.$$

It is also natural to assume the discrete analogs of [H1] and [H2] as stability conditions.

$[\mathrm{H1}]_h$ *There exists a positive constant* k_1^* *independent of* $h \in \Lambda$ *such that*

$$a_h(u_h, u_h) \geqq k_1^* \|u_h\|_V^2 \qquad \text{for all } u_h \in N(B_h). \tag{39}$$

$[\mathrm{H2}]_h$ *There exists a positive constant* k_2^* *independent of* $h \in \Lambda$ *such that*

$$\|B_h^* \lambda_h\|_V \geqq k_2^* \|\lambda_h\|_W \qquad \text{for all } \lambda_h \in R(B_h). \tag{40}$$

As in the continuous case, the following is equivalent to $[\mathrm{H2}]_h$.

$[\mathrm{H2}]_h^*$ *There exists a positive constant* k_2^* *independent of* $h \in \Lambda$ *such that*

$$\|B_h u_h\|_W \geqq k_2^* \|u_h\|_V \qquad \text{for all } u_h \in R(B_h^*). \tag{41}$$

Clearly, k_2^* in $[\mathrm{H2}]_h^*$ is equal to that in $[\mathrm{H2}]_h$.

Based on the consideration in the preceding section, we also require that

$[\mathrm{H3}]_h$
$$Q_h R(B) \subset R(B_h) \qquad \text{for all } h \in \Lambda. \tag{42}$$

Notice that $Q_h R(B)$ and $R(B_h)$ are both closed since they are subspaces of a finite-dimensional space W^h. Thus the orthogonal complements of $Q_h R(B)$ and $R(B_h)$ in W^h are shown to be $N(B^*) \cap W^h$ and $N(B_h^*)$, respectively, and hence $[\mathrm{H3}]_h$ is equivalent to:

$[\mathrm{H3}]_h^*$
$$N(B_h^*) \subset N(B^*) \qquad \text{for any } h \in \Lambda. \tag{43}$$

In *Appendix 2*, some results are summarized for $[\mathrm{H2}]_h$ and $[\mathrm{H3}]_h$.

The following is an additional stability condition.

$[\mathrm{H4}]_h$ *There exists a positive constant* k_4 *independent of* $h \in \Lambda$ *such that*

$$\|A_h\|_{\mathcal{L}(V^h, V^h)} \leqq k_4, \quad \|B_h\|_{\mathcal{L}(V^h, W^h)} \leqq k_4. \tag{44}$$

When $A_h=P_hA|V^h$ and $B_h=Q_hB|V^h$, the above holds automatically with $k_4=\max\{\|A\|_{\mathcal{L}(V,V)}, \|B\|_{\mathcal{L}(V,W)}\}$.

As in the continuous case, we can show that $[\mathrm{H1}]_h$ is equivalent to the following when (28), $[\mathrm{H2}]_h$, and $[\mathrm{H4}]_h$ hold.

$[\mathrm{H1}]_h^*$ *There exists a positive constant* k_3^* *independent of* $h\in\Lambda$ *such that*

$$a_h(u_h, u_h)+\|B_hu_h\|_W^2\geqq k_3^*\|u_h\|_V^2 \qquad \textit{for all } u_h\in V^h. \tag{45}$$

We can express k_3^* by a continuous function of k_1^*, k_2^*, and k_4. Some results for $[\mathrm{H1}]_h^*$ are given in *Appendix 3*.

Let us consider a discrete analog of (11): given $\{f, g, g^*\}\in V\times R(B)\times W$ and $\varepsilon\in[0, \varepsilon_0]$, find $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}\in V^h\times W^h$ such that

$$A_hu_{h\varepsilon}+B_h^*\lambda_{h\varepsilon}=P_hf, \quad B_hu_{h\varepsilon}-\varepsilon\lambda_{h\varepsilon}=Q_hg+\varepsilon Q_hg^*. \tag{46}$$

Clearly, the present problem reduces to (34) when $Q_hg^*=0$. We obtain the following lemma from Lemma 1, if we notice the relations $\|P_h\|_{\mathcal{L}(V,V^h)}\leqq 1$ and $\|Q_h\|_{\mathcal{L}(W,W^h)}\leqq 1$.

LEMMA 2. *Assume that* $[\mathrm{H1}]_h$, $[\mathrm{H2}]_h$, $[\mathrm{H3}]_h$, *and* $[\mathrm{H4}]_h$ *hold, and consider problem* (46) *with* $\{f, g, g^*\}\in V\times R(B)\times W$ *and* $\varepsilon\in[0, \varepsilon_0]$ *given. For* $\varepsilon=0$, *there exists a solution* $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}$ *which is unique in* $V^h\times R(B_h)$ *(but may not be so in* $V^h\times W^h$*). Similarly, for* $\varepsilon\in\,]0, \varepsilon_0]$, *there exists a unique solution in* $V^h\times W^h$, *again denoted by* $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}$, *where* $\lambda_{h\varepsilon}$ *belongs to* $R(B_h)$ *if* $Q_hg^*=0$. *In both cases,* $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}$ *satisfies*

$$\|u_{h\varepsilon}\|_V+\|\lambda_{h\varepsilon}\|_W\leqq C_1^*(\|f\|_V+\|g\|_W+\|g^*\|_W), \tag{47}$$

where C_1^* *is a positive number dependent continuously on* k_1^*, k_2^*, k_4, and ε_0 *only, and hence independent of* $h\in\Lambda$ *and* $\varepsilon\in[0, \varepsilon_0]$.

As in the continuous case, let us expand the solution $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}\in V^h\times R(B_h)$ to (34) at $\varepsilon^*\in[0, \varepsilon_0]$:

$$u_{h\varepsilon}=\sum_{i=0}^{\infty}(\varepsilon-\varepsilon^*)^iu_h^{(i)}, \quad \lambda_{h\varepsilon}=\sum_{i=0}^{\infty}(\varepsilon-\varepsilon^*)^i\lambda_h^{(i)}, \tag{48}$$

where the coefficients $\{u_h^{(i)}, \lambda_h^{(i)}\}\in V^h\times R(B_h)$ for $i=0, 1, 2, \cdots$ must satisfy

$$\begin{aligned} &A_hu_h^{(0)}+B_h^*\lambda_h^{(0)}=P_hf, \quad B_hu_h^{(0)}-\varepsilon^*\lambda_h^{(0)}=Q_hg, \\ &A_hu_h^{(i)}+B_h^*\lambda_h^{(i)}=0, \quad B_hu_h^{(i)}-\varepsilon^*\lambda_h^{(i)}=\lambda_h^{(i-1)} \quad (i\geqq 1). \end{aligned} \tag{49}$$

As before, the above expansion is also realized by the following iteration scheme:

(i) $\lambda_{h\varepsilon}^{(-1)}=0\in R(B_h)$, (50)

(ii) for $i=0,1,2,\cdots$, decide $\{u_{h\varepsilon}^{(i)},\lambda_{h\varepsilon}^{(i)}\}\in V^h\times R(B_h)$ recursively by

$$A_h u_{h\varepsilon}^{(i)}+B_h^*\lambda_{h\varepsilon}^{(i)}=P_h f,\quad B_h u_{h\varepsilon}^{(i)}-\varepsilon^*\lambda_{h\varepsilon}^{(i)}=(\varepsilon-\varepsilon^*)\lambda_{h\varepsilon}^{(i-1)}+Q_h g. \tag{51}$$

For the above expansion or the iteration scheme, we can easily deduce the following theorem from Theorems 1 and 2.

THEOREM 3. *Assume that* $[\mathrm{H1}]_h$, $[\mathrm{H2}]_h$, $[\mathrm{H3}]_h$, *and* $[\mathrm{H4}]_h$ *hold. Then, for any* $\{f,g\}\in V\times R(B)$ *and any* $\varepsilon^*\in[0,\varepsilon_0]$, *each of the coefficients* $\{\{u_h^{(i)},\lambda_h^{(i)}\}\}_{i=0}^{\infty}$ *in* (48) *is determined from* (49) *uniquely in* $V^h\times R(B_h)$ *and satisfies*

$$\|u_h^{(i)}\|_V+\|\lambda_h^{(i)}\|_W\leqq(C_1^*)^{i+1}(\|f\|_V+\|g\|_W), \tag{52}$$

where C_1^* *is the positive number appearing in* (47). *The series* (48) *converges to the solution* $\{u_{h\varepsilon},\lambda_{h\varepsilon}\}\in V^h\times R(B_h)$ *of* (34) *if* $\varepsilon\in[0,\varepsilon_0]$ *is subject to either*

$$|\varepsilon-\varepsilon^*|<1/C_1^*\quad or\quad 0\leqq\varepsilon\leqq 2\varepsilon^*. \tag{53}$$

Furthermore, for each $\varepsilon\in[0,\varepsilon_0]$, *the finite series*

$$u_{h\varepsilon}^{(i)}=\sum_{j=0}^{i}(\varepsilon-\varepsilon^*)^j u_h^{(j)},\quad \lambda_{h\varepsilon}^{(i)}=\sum_{j=0}^{i}(\varepsilon-\varepsilon^*)^j\lambda_h^{(j)}\quad(i\geqq 0), \tag{54}$$

which is equal to the i-th approximation $\{u_{h\varepsilon}^{(i)},\lambda_{h\varepsilon}^{(i)}\}$ *defined by* (50) *and* (51), *satisfies the estimation*

$$\|u_{h\varepsilon}^{(i)}-u_{h\varepsilon}\|_V+\|\lambda_{h\varepsilon}^{(i)}-\lambda_{h\varepsilon}\|_W\leqq(C_1^*)^{i+2}|\varepsilon-\varepsilon^*|^{i+1}(\|f\|_V+\|g\|_W). \tag{55}$$

The above theorem states that the discrete solution $\{u_{h\varepsilon},\lambda_{h\varepsilon}\}$ has essentially the same dependence properties on the parameter ε as the continuous one $\{u_\varepsilon,\lambda_\varepsilon\}$. It is also clear that $\lambda_h^{(i)}$ and $\lambda_{h\varepsilon}^{(i)}$ $(i\geqq 0)$ for $\varepsilon^*>0$ can be dealt with as subsidiary quantities just as in the penalty approach.

Hereafter, we will give error analysis of the approximate problem (34). We first obtain the following results.

THEOREM 4. *Assume that* [H1], [H2], $[H1]_h$, $[H2]_h$, $[H3]_h$, *and* $[H4]_h$ *hold, and consider the solutions* $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\} \in V^h \times R(B_h)$ *of* (34) *with* $h \in \Lambda$ *and* $\{u_\varepsilon, \lambda_\varepsilon\} \in V \times R(B)$ *of* (5) *for common* $\{f, g\} \in V \times R(B)$ *and* $\varepsilon \in [0, \varepsilon_0]$. *Then it holds that*

$$\|u_{h\varepsilon}-u_\varepsilon\|_V+\|A_h u_{h\varepsilon}-Au_\varepsilon\|_V+\|B_h u_{h\varepsilon}-Bu_\varepsilon\|_W+\|\lambda_{h\varepsilon}-\lambda_\varepsilon\|_W \\ +\|B_h^*\lambda_{h\varepsilon}-B^*\lambda_\varepsilon\|_V \leqq C_2 E_h(u_\varepsilon, \lambda_\varepsilon) \tag{56}$$

with

$$C_2=1+C_1^*(1+2k_4), \tag{57}$$

where C_1^* *and* k_4 *are respectively the positive numbers in Lemma* 2 *and* $[H4]_h$, *and* $E_h(\cdot, \cdot)$ *is defined by* (38). *Moreover, if* $[H0]_h$ *also holds, the left-hand side of* (56) *tends to zero as* $h \downarrow 0$.

REMARK. The coefficient C_2 appearing in (56) is independent of ε. Therefore, we have a uniform error estimation with respect to ε, if we can evaluate $E_h(u_\varepsilon, \lambda_\varepsilon)$ uniformly in ε. This is a very important fact in practical applications, and see Arnold [1] and Kikuchi [13] for such discussion and concrete examples.

PROOF. From (34) and (5), it holds for any $\{v_h, \mu_h\} \in V^h \times W^h$ that

$$A_h(u_{h\varepsilon}-v_h)+B_h^*(\lambda_{h\varepsilon}-\mu_h)=P_h(Au_\varepsilon-A_h v_h)+P_h(B^*\lambda_\varepsilon-B_h^*\mu_h),$$
$$B_h(u_{h\varepsilon}-v_h)-\varepsilon(\lambda_{h\varepsilon}-\mu_h)=Q_h(Bu_\varepsilon-B_h v_h)-\varepsilon Q_h(\lambda_\varepsilon-\mu_h).$$

For $\varepsilon>0$, applying Lemma 2 to the above, we have

$$\|u_{h\varepsilon}-v_h\|_V+\|\lambda_{h\varepsilon}-\mu_h\|_W \leqq C_1^*(\|A_h v_h-Au_\varepsilon\|_V+\|B_h^*\mu_h-B^*\lambda_\varepsilon\|_V \\ +\|B_h v_h-Bu_\varepsilon\|_W+\|\mu_h-\lambda_\varepsilon\|_W). \tag{a}$$

We have excluded the case of $\varepsilon=0$ since μ_h is not necessarily in $R(B_h)$. However, by the continuity of $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}$ and $\{u_\varepsilon, \lambda_\varepsilon\}$ in $\varepsilon \in [0, \varepsilon_0]$ assured in Theorems 1 and 3, the above inequality holds also for $\varepsilon=0$. By the triangle inequality and $[H4]_h$,

$$\|u_{h\varepsilon}-u_\varepsilon\|_V+\|\lambda_{h\varepsilon}-\lambda_\varepsilon\|_W \leqq \|u_{h\varepsilon}-v_h\|_V+\|\lambda_{h\varepsilon}-\mu_h\|_W+\|v_h-u_\varepsilon\|_V+\|\mu_h-\lambda_\varepsilon\|_W,$$
$$\|A_h u_{h\varepsilon}-Au_\varepsilon\|_V \leqq \|A_h\| \cdot \|u_{h\varepsilon}-v_h\|_V+\|A_h v_h-Au_\varepsilon\|_V \\ \leqq k_4\|u_{h\varepsilon}-v_h\|_V+\|A_h v_h-Au_\varepsilon\|_V,$$
$$\|B_h u_{h\varepsilon}-Bu_\varepsilon\|_W \leqq k_4\|u_{h\varepsilon}-v_h\|_V+\|B_h v_h-Bu_\varepsilon\|_W,$$
$$\|B_h^*\lambda_{h\varepsilon}-B^*\lambda_\varepsilon\|_V \leqq k_4\|\lambda_{h\varepsilon}-\mu_h\|_W+\|B_h^*\mu_h-B^*\lambda_\varepsilon\|_V.$$

From (a) and the above estimates, we have (56) with (57) due to the arbitrariness of $\{v_h, \mu_h\}$ in $V^h \times W^h$. By $[\mathrm{H0}]_h$, the left-hand side of (56) tends to 0 as $h \downarrow 0$, and the proof is complete.

Based upon the above results, we can also obtain error estimations for the expansion (48).

THEOREM 5. *Assume that* [H1], [H2], $[\mathrm{H1}]_h$, $[\mathrm{H2}]_h$, $[\mathrm{H3}]_h$, *and* $[\mathrm{H4}]_h$ *hold, and consider expansions* (15) *and* (48) $(h \in \Lambda)$ *for* $\{f, g\} \in V \times R(B)$, $\varepsilon^* \in [0, \varepsilon_0]$, *and* $\varepsilon \in [0, \varepsilon_0]$. *Then the coefficients* $\{u_h^{(i)}, \lambda_h^{(i)}\} \in V^h \times R(B_h)$ $(i \geqq 0)$ *defined by* (49) *and the ones* $\{u^{(i)}, \lambda^{(i)}\} \in V \times R(B)$ $(i \geqq 0)$ *defined by* (16) *satisfy*

$$\|u_h^{(i)} - u^{(i)}\|_V + \|\lambda_h^{(i)} - \lambda^{(i)}\|_W \leqq C_2 \sum_{j=0}^{i} (C_1^*)^{i-j} E_h(u^{(j)}, \lambda^{(j)}), \tag{58}$$

where C_1^*, C_2, *and* $E_h(\cdot, \cdot)$ *are introduced in Lemma* 2, *Theorem* 4, *and* (38), *respectively, Furthermore, the difference between the finite series* (54) *and* (19) *satisfies, for* $i = 0, 1, 2, \cdots$,

$$\begin{aligned} &\|u_{h\varepsilon}^{(i)} - u_\varepsilon^{(i)}\|_V + \|\lambda_{h\varepsilon}^{(i)} - \lambda_\varepsilon^{(i)}\|_W \\ &\leqq \min \{C_2 \sum_{j=0}^{i} |\varepsilon - \varepsilon^*|^j \sum_{k=0}^{j} (C_1^*)^{j-k} E_h(u^{(k)}, \lambda^{(k)}), \\ &\qquad C_2 \sum_{j=0}^{i} (C_1^*)^{i-j} |\varepsilon - \varepsilon^*|^{i-j} E_h(u_\varepsilon^{(j)}, \lambda_\varepsilon^{(j)}), \\ &\qquad C_2 E_h(u_\varepsilon^{(i)}, \lambda_\varepsilon^{(i)}) + C_2 |\varepsilon - \varepsilon^*|^{i+1} \sum_{j=0}^{i} (C_1^*)^{i+1-j} E_h(u^{(j)}, \lambda^{(j)})\}. \end{aligned} \tag{59}$$

PROOF. From (16) and (49), we find that, for $i = 0, 1, 2, \cdots$,

$$\begin{aligned} &A_h u_h^{(i)} + B_h^* \lambda_h^{(i)} = P_h A u^{(i)} + P_h B^* \lambda^{(i)}, \\ &B_h u_h^{(i)} - \varepsilon^* \lambda_h^{(i)} = Q_h B u^{(i)} - \varepsilon^* Q_h \lambda^{(i)} + Q_h (\lambda_h{}^{(i-1)} - \lambda^{(i-1)}), \end{aligned}$$

where $\lambda_h^{(-1)} = \lambda^{(-1)} = 0$. Then, as in the proof of Theorem 4, we have

$$\begin{aligned} &\|u_h^{(0)} - u^{(0)}\|_V + \|\lambda_h^{(0)} - \lambda^{(0)}\|_W \leqq C_2 E_h(u^{(0)}, \lambda^{(0)}), \\ &\|u_h^{(i)} - u^{(i)}\|_V + \|\lambda_h^{(i)} - \lambda^{(i)}\|_W \leqq C_2 E_h(u^{(i)}, \lambda^{(i)}) + C_1^* \|\lambda_h^{(i-1)} - \lambda^{(i-1)}\|_W \quad (i \geqq 1). \end{aligned}$$

Solving this recurrence relation, we obtain (58). The first estimation in (59) immediately follows from (58). To obtain the second estimation in (59), we should note the following identity derived from (25) and (51).

$$\begin{aligned} &A_h u_{h\varepsilon}^{(i)} + B_h^* \lambda_{h\varepsilon}^{(i)} = P_h A u_\varepsilon^{(i)} + P_h B^* \lambda_\varepsilon^{(i)}, \\ &B_h u_{h\varepsilon}^{(i)} - \varepsilon^* \lambda_{h\varepsilon}^{(i)} = Q_h B u_\varepsilon^{(i)} - \varepsilon^* Q_h \lambda_\varepsilon^{(i)} + (\varepsilon - \varepsilon^*) Q_h (\lambda_{h\varepsilon}^{(i-1)} - \lambda_\varepsilon^{(i-1)}). \end{aligned}$$

Then we can deduce the desired estimation just as (58). Noting that the above identity is easily rewritten by

$$A_h u_{h\varepsilon}^{(i)}+B_h^*\lambda_{h\varepsilon}^{(i)}=P_h A u_\varepsilon^{(i)}+P_h B^*\lambda_\varepsilon^{(i)},$$
$$B_h u_{h\varepsilon}^{(i)}-\varepsilon\lambda_{h\varepsilon}^{(i)}=Q_h B u_\varepsilon^{(i)}-\varepsilon Q_h\lambda_\varepsilon^{(i)}-(\varepsilon-\varepsilon^*)^{i+1}Q_h(\lambda_h^{(i)}-\lambda^{(i)}),$$

we have similarly the last estimation in (59) by (58). This completes the proof.

7. Error analysis of the approximate solutions under weaker assumptions

We will give some results for error estimation of the discrete problem (34) with $g=0$ under assumptions weaker than those in the preceding section. Then the continuous problem and the discrete one to solve respectively become

$$Au_\varepsilon+B^*\lambda_\varepsilon=f,\quad Bu_\varepsilon-\varepsilon\lambda_\varepsilon=0, \tag{60}$$

$$A_h u_{h\varepsilon}+B_h^*\lambda_{h\varepsilon}=P_h f,\quad B_h u_{h\varepsilon}-\varepsilon\lambda_{h\varepsilon}=0. \tag{61}$$

In particular, for $\varepsilon=0$, u_ε and $u_{h\varepsilon}$ respectively belong to $N(B)$ and $N(B_h)$ and are characterized by

$$a(u_\varepsilon, v)=(f, v)_V \qquad \text{for all } v\in N(B), \tag{62}$$

$$a_h(u_{h\varepsilon}, v_h)=(f, v_h)_V \qquad \text{for all } v_h\in N(B_h). \tag{63}$$

Since $g=0$, $[\mathrm{H3}]_h$ may be omitted at least formally. We do not employ $[\mathrm{H2}]_h$, but still adopt $[\mathrm{H1}]_h^*$ and $[\mathrm{H4}]_h$. Recall that $[\mathrm{H1}]_h^*$ implies $[\mathrm{H1}]_h$ with $k_1^*=k_3^*$, while $[\mathrm{H1}]_h^*$ is not deduced from $[\mathrm{H1}]_h$ without $[\mathrm{H2}]_h$. Although $[\mathrm{H2}]_h$ is not assumed, there still exists a positive number k_{h2}, possibly dependent on $h\in\Lambda$, such that

$$\|B_h^*\lambda_h\|_V\geqq k_{h2}\|\lambda_h\|_W \qquad \text{for all } \lambda_h\in R(B_h). \tag{64}$$

Let us consider the solution $\{u_\varepsilon, \lambda_\varepsilon\}\in V\times R(B)$ of (60) for $f\in V$ and $\varepsilon\in[0, \varepsilon_0]$. By Theorem 1, it can be expressed as

$$u_\varepsilon=u_0+\varepsilon v_\varepsilon,\quad \lambda_\varepsilon=\lambda_0+\varepsilon\mu_\varepsilon, \tag{65}$$

where $\{u_0, \lambda_0\}\in N(B)\times R(B)$ and $\{v_\varepsilon, \mu_\varepsilon\}\in V\times R(B)$ are respectively the unique solutions of

$$Au_0+B^*\lambda_0=f,\quad Bu_0=0, \tag{66}$$

$$Av_\varepsilon + B^*\mu_\varepsilon = 0, \quad Bv_\varepsilon - \varepsilon\mu_\varepsilon = \lambda_0. \tag{67}$$

It is clear that for any $\varepsilon \in [0, \varepsilon_0]$

$$\|u_0\|_V + \|\lambda_0\|_W \leqq C_1\|f\|_V, \quad \|v_\varepsilon\|_V + \|\mu_\varepsilon\|_W \leqq C_1^2\|f\|_V, \tag{68}$$

where C_1 is the positive constant in (12).

We should first establish the existence and the uniqueness of the approximate solution as well as some estimates. We will show the results for a problem more general than (61): given $\{f, g^*\} \in V \times W$ and $\varepsilon \in [0, \varepsilon_0]$, find $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\} \in V^h \times W^h$ such that

$$A_h u_{h\varepsilon} + B_h^* \lambda_{h\varepsilon} = P_h f, \quad B_h u_{h\varepsilon} - \varepsilon\lambda_{h\varepsilon} = \varepsilon Q_h g^*. \tag{69}$$

LEMMA 3. *Assume that* $[\mathrm{H1}]_h^*$ *and* $[\mathrm{H4}]_h$ *hold, and consider problem* (69) *with* $\{f, g^*\} \in V \times W$ *and* $\varepsilon \in [0, \varepsilon_0]$ *given. For* $\varepsilon = 0$, *there exists a solution* $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}$ *which is unique in* $V^h \times R(B_h)$ *(but may not be so in* $V^h \times W^h$*). In this case,* u_h *necessarily belongs to* $N(B_h)$. *Similarly, for* $\varepsilon \in]0, \varepsilon_0]$, *there exists a unique solution in* $V^h \times W^h$, *again denoted by* $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}$, *where* $\lambda_{h\varepsilon}$ *belongs to* $R(B_h)$ *for* $g^* = 0$. *In both cases,* $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}$ *satisfies the estimations*

$$\|u_{h\varepsilon}\|_V \leqq (k_3^*)^{-1} \max\{1, \varepsilon\}(\|f\|_V + k_4\|g^*\|_W), \tag{70}$$

$$\|\lambda_{h\varepsilon}\|_W \leqq \|g^*\|_W + k_{h2}^{-1}(\|f\|_V + k_4\|u_{h\varepsilon}\|_V), \tag{71}$$

where k_3^*, k_4, *and* k_{h2} *are the positive numbers in* $[\mathrm{H1}]_h^*$, $[\mathrm{H4}]_h$, and (64), *respectively, and the term* $\|g^*\|_W$ *may be omitted for* $\varepsilon = 0$. *Moreover, for* $\varepsilon > 0$,

$$\|\lambda_{h\varepsilon}\|_W \leqq \|g^*\|_W + \varepsilon^{-1/2}(\|f\|_V + \|u_{h\varepsilon}\|_V). \tag{72}$$

REMARK. From (70), we can see that $\|u_{h\varepsilon}\|_V$ is uniformly bounded with respect to $h \in \Lambda$ and $\varepsilon \in [0, \varepsilon_0]$. Although the estimates for $\lambda_{h\varepsilon}$ do not possess such uniformity, estimate (71) is uniform in ε and estimate (72) is so in h.

PROOF. The existence and the uniqueness follow from Lemma 1 since (64) and $[\mathrm{H1}]_h^*$ hold.

For $\varepsilon = 0$, (69) reduces to (61), and hence $u_{h\varepsilon}$ belongs to $N(B_h)$ and satisfies (63). Then,

$$a_h(u_{h\varepsilon}, u_{h\varepsilon}) + \|B_h u_{h\varepsilon}\|_W^2 = (f, u_{h\varepsilon})_V \leqq \|f\|_V \|u_{h\varepsilon}\|_V,$$

which yields (70) for $\varepsilon=0$ by $[\mathrm{H1}]_h^*$.

For $\varepsilon>0$, we can eliminate $\lambda_{h\varepsilon}$ from (69) to find

$$a_h(u_{h\varepsilon}, u_{h\varepsilon})+\varepsilon^{-1}\|B_h u_{h\varepsilon}\|_W^2=(f+B_h^* Q_h g^*, u_{h\varepsilon})_V.$$

Then we can obtain (70) for $\varepsilon>0$ by $[\mathrm{H4}]_h$ and $[\mathrm{H1}]_h^*$.

To show (71), decompose $\lambda_{h\varepsilon}$ and $Q_h g^*$ as

$$\lambda_{h\varepsilon}=\lambda_{h1}+\lambda_{h2}\,;\ \lambda_{h1}\in N(B_h^*),\ \lambda_{h2}\in R(B_h),$$
$$Q_h g^*=g_{h1}+g_{h2}\,;\ g_{h1}\in N(B_h^*),\ g_{h2}\in R(B_h).$$

Substituting these into (69), we have

$$B_h^*\lambda_{h2}=P_h f-A_h u_{h\varepsilon},\quad -\varepsilon\lambda_{h1}=\varepsilon g_{h1}.$$

Applying (64) and $[\mathrm{H4}]_h$ to the first equation of the above, we find that $\lambda_{h2}\in R(B_h)$ satisfies

$$k_{h2}\|\lambda_{h2}\|_W\leqq\|f\|_V+k_4\|u_{h\varepsilon}\|_V. \tag{a}$$

On the other hand, $\|\lambda_{h1}\|_W$ for $\varepsilon>0$ is evaluated as

$$\|\lambda_{h1}\|_W\leqq\|g_{h1}\|_W\leqq\|g^*\|_W. \tag{b}$$

Now (71) follows from (a) and (b).

Finally, to show (72), we again utilize (69) as follows:

$$\begin{aligned}\varepsilon\|\lambda_{h\varepsilon}\|^2&=(B_h u_{h\varepsilon}-\varepsilon Q_h g^*, \lambda_{h\varepsilon})_W=(u_{h\varepsilon}, B_h^*\lambda_{h\varepsilon})_V-\varepsilon(g^*, \lambda_{h\varepsilon})_W\\&=(u_{h\varepsilon}, P_h f-A_h u_{h\varepsilon})_V-\varepsilon(g^*, \lambda_{h\varepsilon})_W\\&=(f, u_{h\varepsilon})_V-(A_h u_{h\varepsilon}, u_{h\varepsilon})_V-\varepsilon(g^*, \lambda_{h\varepsilon})_W.\end{aligned}$$

Since $(A_h u_{h\varepsilon}, u_{h\varepsilon})_V\geqq 0$ by (28), we can evaluate the above as

$$\varepsilon\|\lambda_{h\varepsilon}\|_W^2\leqq 2^{-1}(\|f\|_V^2+\|u_{h\varepsilon}\|_V^2)+2^{-1}\varepsilon(\|g^*\|_W^2+\|\lambda_{h\varepsilon}\|_W^2),$$

and hence

$$\varepsilon^{1/2}\|\lambda_{h\varepsilon}\|_W\leqq\|f\|_V+\|u_{h\varepsilon}\|_V+\varepsilon^{1/2}\|g^*\|_W.$$

Then (72) is obvious, and the proof is complete.

We will state a theorem for error estimation of the discrete solution. To this end, let us define some quantities for $\{u_h, \lambda_h\}\in V^h\times W^h$ and $\{u, \lambda\}\in V\times W$:

$$\begin{aligned} &e_1(u_h, u)=\|u_h-u\|_V, \quad e_2(u_h, u)=\|A_h u_h-Au\|_V, \\ &e_3(u_h, u)=\|B_h u_h-Bu\|_W, \\ &e_4(\lambda_h, \lambda)=\|\lambda_h-\lambda\|_W, \quad e_5(\lambda_h, \lambda)=\|B_h^*\lambda_h-B^*\lambda\|_V. \end{aligned} \tag{73}$$

THEOREM 6. *Assume that* [H1], [H2], $[\mathrm{H1}]_h^*$, *and* $[\mathrm{H4}]_h$ *hold, and consider the solutions* $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}\in V^h\times R(B_h)$ *of* (61) *with* $h\in\Lambda$ *and* $\{u_\varepsilon, \lambda_\varepsilon\}\in V\times R(B)$ *of* (60) *for common* $f\in V$ *and* $\varepsilon\in[0, \varepsilon_0]$. *Then we have the following results.*
(i) *For* $\varepsilon=0$, *it holds that*

$$\|u_{h\varepsilon}-u_\varepsilon\|_V\leqq \inf_{v_h\in N(B_h)}\{e_1^0+(k_3^*)^{-1}e_2^0\}+(k_3^*)^{-1}\inf_{\mu_h\in W^h}e_5^0, \tag{74}$$

$$\|\lambda_{h\varepsilon}-\lambda_\varepsilon\|_W\leqq k_{h2}^{-1}k_4\|u_{h\varepsilon}-u_\varepsilon\|_V+\inf_{v_h\in N(B_h)}k_{h2}^{-1}(k_4e_1^0+e_2^0)+\inf_{\mu_h\in R(B_h)}(e_4^0+k_{h2}^{-1}e_5^0), \tag{75}$$

where, for $v_h\in N(B_h)$ *and* $\mu_h\in W^h$ (*or* $\mu_h\in R(B_h)$ *in* (75)),

$$e_i^0=e_i(v_h, u_\varepsilon) \quad (i=1, 2, 3), \quad e_i^0=e_i(\mu_h, \lambda_\varepsilon) \quad (i=4, 5). \tag{76}$$

If $B_h=B|V^h$, *the term* $e_5^0=e_5(\mu_h, \lambda_\varepsilon)$ *can be omitted in* (74).
(ii) *For* $\varepsilon>0$, *it holds that*

$$\begin{aligned} \|u_{h\varepsilon}-u_\varepsilon\|_V\leqq &\inf_{v_h\in V^h}[e_1^0+\max\{1, \varepsilon\}(k_3^*)^{-1}(e_2^0+k_4\varepsilon^{-1}e_3^0)] \\ &+\inf_{\mu_h\in W^h}\max\{1, \varepsilon\}(k_3^*)^{-1}(k_4e_4^0+e_5^0), \end{aligned} \tag{77}$$

$$\|\lambda_{h\varepsilon}-\lambda_\varepsilon\|_W\leqq \varepsilon_{h1}\|u_{h\varepsilon}-u_\varepsilon\|_V+\inf_{v_h\in V^h}(\varepsilon_{h1}e_1^0+\varepsilon_{h2}e_2^0+\varepsilon^{-1}e_3^0)+\inf_{\mu_h\in W^h}(2e_4^0+\varepsilon_{h2}e_5^0), \tag{78}$$

where

$$\varepsilon_{h1}=\min\{k_{h2}^{-1}k_4, \varepsilon^{-1/2}\}, \quad \varepsilon_{h2}=\min\{k_{h2}^{-1}, \varepsilon^{-1/2}\}, \tag{79}$$

and $e_1^0, \cdots, e_5^0$ *are formally the same as those in* (76). *Utilizing the expression* $u_\varepsilon=u_0+\varepsilon v_\varepsilon$ *in* (65), *the above estimates may be modified as*

$$\begin{aligned} \|u_{h\varepsilon}-u_\varepsilon\|_V\leqq &\inf_{u_h\in N(B_h)}[e_1^1+\max\{1, \varepsilon\}(k_3^*)^{-1}e_2^1] \\ &+\inf_{w_h\in V^h}[\varepsilon e_1^2+\max\{1, \varepsilon\}(k_3^*)^{-1}(\varepsilon e_2^2+k_4e_3^2)] \\ &+\inf_{\mu_h\in W^h}\max\{1, \varepsilon\}(k_3^*)^{-1}(k_4e_4^0+e_5^0), \end{aligned} \tag{80}$$

$$\begin{aligned} \|\lambda_{h\varepsilon}-\lambda_\varepsilon\|_W\leqq &\varepsilon_{h1}\|u_{h\varepsilon}-u_\varepsilon\|_V+\inf_{u_h\in N(B_h)}(\varepsilon_{h1}e_1^1+\varepsilon_{h2}e_2^1) \\ &+\inf_{w_h\in V^h}(\varepsilon_{h1}\varepsilon e_1^2+\varepsilon_{h2}\varepsilon e_2^2+e_3^2)+\inf_{\mu_h\in W^h}(2e_4^0+\varepsilon_{h2}e_5^0), \end{aligned} \tag{81}$$

where, for $u_h\in N(B_h)$ *and* $w_h\in V^h$,

$$e_i^1=e_i(u_h, u_0) \quad\text{and}\quad e_i^2=e_i(w_h, v_\varepsilon) \quad (i=1, 2, 3). \tag{82}$$

If $a(\cdot,\cdot)$ *is symmetric,* $a_h(\cdot,\cdot)$ *is the restriction of* $a(\cdot,\cdot)$ *to* $V^h\times V^h$, *and* $B_h=B|V^h$, *then*

$$\|u_{h\varepsilon}-u_\varepsilon\|_V\leqq \inf_{v_h\in V^h}\max\{1, \varepsilon^{1/2}\}k_3^{-1/2}(\|A\|^{1/2}+\varepsilon^{-1/2}\|B\|)e_1^0, \tag{83}$$

where k_3 *is the positive number in* (10).

REMARK. Notice that the factor ε^{-1} is absent in (80) and (81) unlike (77) and (78). Moreover, the factor ε^{-1} in (77) is improved up to $\varepsilon^{-1/2}$ in (83). From the above estimates, we can see that the quality of $N(B_h)$ as well as that of $R(B_h)$ are important when $[\mathrm{H2}]_h$ and $[\mathrm{H3}]_h$ do not hold and ε is close to zero. See *Appendix 4* for some elementary results related to $N(B_h)$ and $R(B_h)$. Note also that (i) is a generalization of Theorem 1.1 in Chapter II of Girault-Raviart [11].

PROOF. (i) For $\varepsilon=0$, we find that

$$A_hu_{h\varepsilon}+B_h^*\lambda_{h\varepsilon}=P_hf=P_hAu_\varepsilon+P_hB^*\lambda_\varepsilon, \quad B_hu_{h\varepsilon}=0.$$

Thus, for any $\{v_h, \mu_h\}\in N(B_h)\times R(B_h)$, it holds that

$$\begin{aligned} &A_h(u_{h\varepsilon}-v_h)+B_h^*(\lambda_{h\varepsilon}-\mu_h)=P_h(Au_\varepsilon-A_hv_h)+P_h(B^*\lambda_\varepsilon-B_h^*\mu_h),\\ &B_h(u_{h\varepsilon}-v_h)=0. \end{aligned} \tag{a}$$

Applying Lemm 3 to the above, we have

$$\begin{aligned} &\|u_{h\varepsilon}-v_h\|_V\leqq(k_3^*)^{-1}\{e_2(v_h, u_\varepsilon)+e_5(\mu_h, \lambda_\varepsilon)\},\\ &\|\lambda_{h\varepsilon}-\mu_h\|_W\leqq k_{h2}^{-1}\{e_2(v_h, u_\varepsilon)+e_5(\mu_h, \lambda_\varepsilon)+k_4\|u_{h\varepsilon}-v_h\|_V\}. \end{aligned}$$

By the triangle inequality, we can easily obtain (74) and (75). Note that μ_h in (74) need not lie in $R(B_h)$, since its component in $N(B_h^*)$ has no effect on $e_5(\mu_h, \lambda_\varepsilon)$.

When $B_h=B|V^h$, we find that $N(B_h)\subset N(B)$ and, from (a),

$$\begin{aligned} &(A_h(u_{h\varepsilon}-v_h), u_{h\varepsilon}-v_h)_V+(B_h^*(\lambda_{h\varepsilon}-\mu_h), u_{h\varepsilon}-v_h)_V\\ &=(Au_\varepsilon-A_hv_h, u_{h\varepsilon}-v_h)_V+(B^*\lambda_\varepsilon-B_h^*\mu_h, u_{h\varepsilon}-v_h)_V. \end{aligned}$$

Noting that $B_h(u_{h\varepsilon}-v_h)=B(u_{h\varepsilon}-v_h)=0$ for $v_h\in N(B_h)$, we obtain

$$a_h(u_{h\varepsilon}-v_h, u_{h\varepsilon}-v_h)=(Au_\varepsilon-A_hv_h, u_{h\varepsilon}-v_h)_V.$$

Then, by $[\mathrm{H1}]_h^*$ and the triangle inequality, we can conclude that (74) holds with the term $e_5(\mu_h, \lambda_\varepsilon)$ omitted.

(ii) For $\varepsilon>0$, we have

$$A_h u_{h\varepsilon}+B_h^*\lambda_{h\varepsilon}=P_h A u_\varepsilon+P_h B^*\lambda_\varepsilon, \quad B_h u_{h\varepsilon}-\varepsilon\lambda_{h\varepsilon}=Q_h B u_\varepsilon-\varepsilon Q_h\lambda_\varepsilon.$$

Thus, for any $\{v_h, \mu_h\}\in V^h\times W^h$,

$$\begin{aligned}&A_h(u_{h\varepsilon}-v_h)+B_h^*(\lambda_{h\varepsilon}-\mu_h)=P_h(Au_\varepsilon-A_h v_h)+P_h(B^*\lambda_\varepsilon-B_h^*\mu_h),\\&B_h(u_{h\varepsilon}-v_h)-\varepsilon(\lambda_{h\varepsilon}-\mu_h)=Q_h(Bu_\varepsilon-B_h v_h)-\varepsilon Q_h(\lambda_\varepsilon-\mu_h).\end{aligned}$$

By Lemma 3, we find

$$\begin{aligned}&\|u_{h\varepsilon}-v_h\|_V\leqq \max\{1,\varepsilon\}(k_3^*)^{-1}\{e_2^0+e_5^0+k_4(\varepsilon^{-1}e_3^0+e_4^0)\},\\&\|\lambda_{h\varepsilon}-\mu_h\|_W\leqq\varepsilon^{-1}e_3^0+e_4^0+\min\{k_{h2}^{-1},\varepsilon^{-1/2}\}(e_2^0+e_5^0)+\min\{k_{h2}^{-1}k_4,\varepsilon^{-1/2}\}\|u_{h\varepsilon}-v_h\|_V,\end{aligned}$$

from which (77) and (78) follow by the triangle inequality.

Decompose u_ε as (65) and v_h in (77) and (78) as

$$v_h=u_h+\varepsilon w_h\,;\quad u_h\in N(B_h),\quad w_h\in V^h.$$

Then $e_i^0\leqq e_i^1+\varepsilon e_i^2$ for $i=1,2,3$ and $e_3^1=\|B_h u_h-Bu_0\|_W=0$ since $u_h\in N(B_h)$ and $u_0\in N(B)$. It is now easy to deduce (80) and (81) from (77) and (78).

When $a(\cdot,\cdot)$ is symmetric, $a_h(\cdot,\cdot)$ is the restriction of $a(\cdot,\cdot)$ to $V^h\times V^h$, and $B_h=B|V^h$, then we find, by the definitions of u_ε and $u_{h\varepsilon}$, that

$$\begin{aligned}&\{a(u_{h\varepsilon}-u_\varepsilon, u_{h\varepsilon}-u_\varepsilon)+\varepsilon^{-1}\|Bu_{h\varepsilon}-Bu_\varepsilon\|^2\}^{1/2}\\&\leqq\{a(v_h-u_\varepsilon, v_h-u_\varepsilon)+\varepsilon^{-1}\|Bv_h-Bu_\varepsilon\|^2\}^{1/2}\qquad\text{for any } v_h\in V^h.\end{aligned}$$

Then, by applying $[\mathrm{H1}]^*$, we have

$$\begin{aligned}[k_3\min\{1,\varepsilon^{-1}\}]^{1/2}\|u_{h\varepsilon}-u_\varepsilon\|_V&\leqq(\|A\|\cdot\|v_h-u_\varepsilon\|_V^2+\varepsilon^{-1}\|B\|^2\|v_h-u_\varepsilon\|_V^2)^{1/2}\\&\leqq(\|A\|^{1/2}+\varepsilon^{-1/2}\|B\|)\|v_h-u_\varepsilon\|_V,\end{aligned}$$

from which (83) follows. This completes the proof.

8. Some observations for locking

As we have seen in the preceding section, the accuracy of the solutions to (61) is expected to be very poor when $N(B_h)$ fails to approximate $N(B)$ and ε is close to 0. Such a phenomenon actually occurs in some finite element schemes and is called *locking*: see Arnold [1], Kikuchi [13], Oden and Kikuchi [16], and many others. In this section, we will analyze asymptotic behaviors of the numerical solutions for ε close to 0.

Let us assume that $[H1]_h^*$ and $[H4]_h$ hold. Then, thanks to (64) and Theorem 1, the discrete solution $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\} \in V^h \times R(B_h)$ to (61) can be expanded with respect to ε. The iteration scheme given in section 4 is also available. Especially, the expansion at $\varepsilon^*=0$ is given by

$$u_{h\varepsilon}=\sum_{i=0}^{\infty} \varepsilon^i u_h^{(i)}, \quad \lambda_{h\varepsilon}=\sum_{i=0}^{\infty} \varepsilon^i \lambda_h^{(i)}, \tag{84}$$

where each of the coefficients $\{\{u_h^{(i)}, \lambda_h^{(i)}\}\}_{i=0}^{\infty}$ belongs to $V^h \times R(B_h)$ and is recursively determined by

$$\begin{aligned} &A_h u_h^{(0)}+B_h^* \lambda_h^{(0)}=P_h f, \quad B_h u_h^{(0)}=0, \\ &A_h u_h^{(i)}+B_h^* \lambda_h^{(i)}=0, \qquad B_h u_h^{(i)}=\lambda_h^{(i-1)} \quad (i \geqq 1). \end{aligned} \tag{85}$$

Moreover, it holds that

$$\|u_h^{(i)}\|_V+\|\lambda_h^{(i)}\|_W \leqq C_{h1}^{i+1}\|f\|_V \quad (i=0,1,2, \cdots), \tag{86}$$

where C_{h1} is a positive number dependent continuously on k_3^*, k_4, and k_{h2} only, and hence may depend on h. Thus the convergence of (84) is assured if

$$|\varepsilon|<1/C_{h1}. \tag{87}$$

As a result, the numerical solution $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}$ is close to $\{u_h^{(0)}, \lambda_h^{(0)}\} \in N(B_h) \times R(B_h)$ when $\varepsilon \in [0, \varepsilon_0]$ is small enough, although the smallness of ε may vary with h. Therefore, the accuracy of $u_{h\varepsilon}$ is in fact poor, if ε is small and $N(B_h)$ cannot approximate $N(B)$ well. In typical examples of locking, $N(B_h)$ is $\{0\}$ for any $h \in \Lambda$ and does not approximate $N(B) \neq \{0\}$ at all. Such an example will be given in the next section. Note here that, for each fixed $\varepsilon>0$, the accuracy of $u_{h\varepsilon}$ may be improved by decreasing h if $[H0]_h$ holds, as may be seen from (77) or (83). However, acceptable values of h for such $\varepsilon>0$ are often outside practical range. That is, the finite element meshes required for such ε are too fine to be used in actual computations. Consequently, it is very important in practice to develop finite element schemes that satisfy all the hypotheses in section 6.

The following proposition, giving a result for k_{h2} in (64), shows that $[H2]_h$ or $[H2]_h^*$ plays an essential role for $N(B_h)$ to approximate $N(B)$ well. We omit its proof since it is only a corollary of Proposition A4 in *Appendix 4*.

PROPOSITION 1. *Assume that there exists $u \in N(B)$ such that*

$$\overline{\lim_{h \downarrow 0}} \inf_{u_h \in N(B_h)} \|u_h - u\|_V > 0. \tag{88}$$

If

$$\lim_{h \downarrow 0} \inf_{v_h \in V^h} (\|v_h - v\|_V + \|B_h v_h - Bv\|_W) = 0 \quad \textit{for each } v \in V, \tag{89}$$

then

$$\underline{\lim_{h \downarrow 0}} \inf_{w_h \in R(B_h^*) - \{0\}} \frac{\|B_h w_h\|_W}{\|w_h\|_V} = \underline{\lim_{h \downarrow 0}} \inf_{\lambda_h \in R(B_h) - \{0\}} \frac{\|B_h^* \lambda_h\|_V}{\|\lambda_h\|_W} = 0. \tag{90}$$

That is, $k_{h2} > 0$ in (64) *can be infinitely close to* 0 *as h tends to* 0, *if V^h and B_h can respectively approximate V and B but $N(B_h)$ fails to approximate $N(B)$.*

9. Numerical results for the Timoshenko beam problem

We will reconsider the Timoshenko beam problem discussed by Arnold [1]. This problem is concerned with the determination of the displacements and stresses of moderately thick beams [19]. In this case,

$$\Omega =]0, 1[\text{ (unit open interval)}, \quad V = H_0^1(\Omega) \times H_0^1(\Omega), \quad W = L_2(\Omega), \tag{91}$$

and the bilinear forms are given by

$$a(u, v) = (Du_2, Dv_2)_{L_2(\Omega)} \text{ for } u = \{u_1, u_2\}, \ v = \{v_1, v_2\} \in V, \tag{92}$$

$$b(u, \lambda) = (Du_1 - u_2, \lambda)_{L_2(\Omega)} \text{ for } u = \{u_1, u_2\} \in V, \ \lambda \in W. \tag{93}$$

In the above, $H_0^1(\Omega)$ and $L_2(\Omega)$ are the usual Sobolev spaces related to Ω, and D denotes d/dx with x being the independent variable of R^1. Later, we will also use the Sobolev spaces $H^1(\Omega)$ and $H^2(\Omega)$. Note that [H1] and [H2] hold with $N(B) \neq \{0\}$ and $R(B) = W$, as was proven by Arnold. The free terms $f \in V$ and $g \in W$ are of special forms: g is chosen 0, while f is determined from each given $p \in L_2(\Omega)$ by

$$(f, u)_V = (p, u_1)_{L_2(\Omega)} \quad \text{for all } u = \{u_1, u_2\} \in V. \tag{94}$$

Clearly, the solution $\{u_\varepsilon, \lambda_\varepsilon\} = \{\{u_{\varepsilon 1}, u_{\varepsilon 2}\}, \lambda_\varepsilon\} \in V \times W$ to (5) must satisfy, in the sense of distribution,

$$-D\lambda_\varepsilon = p, \quad -D^2 u_{\varepsilon 2} - \lambda_\varepsilon = 0, \quad Du_{\varepsilon 1} - u_{\varepsilon 2} - \varepsilon \lambda_\varepsilon = 0, \tag{95}$$

where $D^2=d^2/dx^2$. The boundary conditions are of course

$$u_{\varepsilon 1}=u_{\varepsilon 2}=0 \quad \text{at } x=0, 1. \tag{96}$$

It is also shown by Arnold that $\{u_\varepsilon, \lambda_\varepsilon\}\in V\times W$ actually belongs to $(V\cap\{H^2(\Omega)\}^2)\times H^1(\Omega)$ with the estimation

$$\|u_{\varepsilon 1}\|_2+\|u_{\varepsilon 2}\|_2+\|\lambda_\varepsilon\|_1\leqq C\|p\|_0, \tag{97}$$

where $\|\cdot\|_m$ denotes the norm of $H^m(\Omega)$ for $m=0, 1, 2$ with $H^0(\Omega)=L_2(\Omega)$, and C is a positive constant independent of p and $\varepsilon\in[0, \varepsilon_0]$.

To solve the above problem by the finite element method, we use a uniform partition of Ω with N subintervals (N=positive integer). Then the size of each subinterval is $h=1/N$. For a non-negative integer n, let us define a function space P_n^h by

P_n^h=totality of functions defined over $\bar{\Omega}$ (=closure of Ω) whose restrictions to each subinterval are polynomials of order at most n. (98)

Note that a function in P_n^h may be discontinuous at the endpoints of the subintervals. We use as V^h

$$V^h=\{P_1^h\cap H_0^1(\Omega)\}\times\{P_1^h\cap H_0^1(\Omega)\}, \tag{99}$$

while we adopt as W^h the following two choices:

$$W_0^h=P_0^h, \quad W_1^h=P_1^h. \tag{100}$$

Note that each component of a function $u_h=\{u_{h1}, u_{h2}\}\in V^h$ is continuous over $\bar{\Omega}$: see Ciarlet [7]. As $a_h(\cdot,\cdot)$ and $b_h(\cdot,\cdot)$, we employ the restrictions of $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ to $V^h\times V^h$ and $V^h\times W^h$, respectively.

To sum up, we consider the following two finite element schemes:

(i) **Scheme-1** the Galerkin method based on $V^h\times W_0^h$,

(ii) **Scheme-2** the Galerkin method based on $V^h\times W_1^h$.

Recall that $B_h=B|V^h$ in Scheme-2 since $BV^h\subset W_1^h$, and that Scheme-1 may be derived from Scheme-2 by the use of an appropriate selective reduced integration: see Arnold [1]. It is not so difficult to see that Scheme-1 satisfies $[\mathrm{H1}]_h$, $[\mathrm{H2}]_h$, and $[\mathrm{H3}]_h$ with $R(B_h)=W^h$ for $N\geqq 2$. On the other hand, we can easily find that $N(B_h)=\{0\}$ in Scheme-2, and hence locking actually occurs. Then Proposition 1 in section 8 implies that Scheme-2 does not satisfy $[H2]_h$, although we can see that it satisfies $[H1]_h^*$. In both schemes, $[H4]_h$ holds automatically.

Let us denote the finite element solutions by $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}=\{\{u_{h\varepsilon 1}, u_{h\varepsilon 2}\}, \lambda_{h\varepsilon}\} \in V^h \times W^h$. According to the error analysis by Arnold and estimates (81) and (83) in this paper, Scheme-1 satisfies the estimation

$$\|u_{h\varepsilon}-u_\varepsilon\|_V+h^{-1}\sum_{i=1}^{2}\|u_{h\varepsilon i}-u_{\varepsilon i}\|_0+\|\lambda_{h\varepsilon}-\lambda_\varepsilon\|_0\leqq Ch\|p\|_0, \tag{101}$$

while Scheme-2 for $\varepsilon>0$ satisfies

$$\|u_{h\varepsilon}-u_\varepsilon\|_V+h^{-1}\varepsilon^{1/2}\sum_{i=1}^{2}\|u_{h\varepsilon i}-u_{\varepsilon i}\|_0+\varepsilon^{1/2}\|\lambda_{h\varepsilon}-\lambda_\varepsilon\|_0\leqq C\varepsilon^{-1/2}h\|p\|_0, \tag{102}$$

where C is a positive constant independent of p, ε, and h. The estimates above also show that Scheme-1 is robust against the variation of ε, while Scheme-2 is unfavorable for the present parameter dependent problem since factors such as $\varepsilon^{-1/2}$ exist in (102). As for the details of the error analysis, see Arnold [1].

As a test problem, we consider the case where $p(x)\equiv 1$. Then $\{u_\varepsilon, \lambda_\varepsilon\}$ is explicitly expressed by

$$\begin{aligned} u_{\varepsilon 1}(x)&=x^4/24-x^3/12+x^2/24-\varepsilon(x^2-x)/2,\\ u_{\varepsilon 2}(x)&=x^3/6-x^2/4+x/12, \quad \lambda_\varepsilon(x)=-x+1/2. \end{aligned} \tag{103}$$

Note here that $u_{\varepsilon 2}$ and λ_ε are independent of ε and that $u_{\varepsilon 2}$ is linear in ε. Thus, $\{u_\varepsilon^{(1)}, \lambda_\varepsilon^{(1)}\}$ of the iteration scheme (25) with (24) gives the exact solution for any ε and ε^* in $[0, \varepsilon_0]$. So, the present problem may be inappropriate to test the feasibility of the iteration scheme, but is very convenient to see the dependence of the numerical solutions on ε.

We obtained finite element solutions for various values of ε and N by Schemes-1 and -2. To this end, we used personal computers MB-6885, PC-8001-II, and PC-9801 in double precision arithmetic. Similar results were also given by Arnold, so we here observe mainly the dependence of numerical solutions on the parameter ε. Especially, we will see the locking phenomenon of Scheme-2. We also tested the iteration scheme (51) with (50) under appropriate stopping conditions. It turned out that Scheme-1 has essentially the same dependence properties on ε, and that $\{u_{h\varepsilon}^{(1)}, \lambda_{h\varepsilon}^{(1)}\}$ gives the exact finite element solution, at least when the partition of Ω is uniform. On the other hand, the iteration scheme applied to Scheme-2 is generally slow in convergence unless ε is sufficiently close to ε^*.

Table 1 gives $u_{h\varepsilon 1}(1/2)$ obtained by Scheme-1 for various combinations of ε and N. It may be easily seen that $u_{h\varepsilon 1}(1/2)$ is linear in ε for each N. The convergence of $u_{h\varepsilon 1}(1/2)$ to $u_{\varepsilon 1}(1/2)$ for $h\downarrow 0$ appears to be uniform in ε.

Table 2 gives $u_{h\varepsilon 1}(1/2)$ obtained by Scheme-2. When ε is not small, the results are almost the same as those of Scheme-1. However, as ε tends to 0, the results becomes incredibly worse: $u_{h\varepsilon 1}(1/2)$ converges to 0 as ε tends to 0 for each fixed N. To see the dependence of $u_{h\varepsilon 1}(1/2)$ on ε and $h=1/N$ more clearly, we compute $\varepsilon^{-1}h^2 u_{h\varepsilon 1}(1/2)$ for various combinations of ε and h. The results are given in Table 3, and it is clear that $u_{h\varepsilon 1}(1/2)$ is almost proportional to $\varepsilon/h^2=\varepsilon N^2$ when ε and N are small. In particular, $u_{h\varepsilon 1}(1/2)$ is asymptotically proportional to ε for each fixed

Table 1. $u_{h\varepsilon 1}(1/2)$ obtained by Scheme-1

ε \ N	10	20	30	40	50	$u_{\varepsilon 1}(1/2)$
1	1.2750E-1	1.2758E-1	1.2759E-1	1.2760E-1	1.2760E-1	1.2760E-1
10^{-1}	1.5000E-2	1.5078E-2	1.5093E-2	1.5098E-2	1.5100E-2	1.5104E-2
10^{-2}	3.7500E-3	3.8281E-3	3.8426E-3	3.8477E-3	3.8500E-3	3.8542E-3
10^{-3}	2.6250E-3	2.7031E-3	2.7176E-3	2.7227E-3	2.7250E-3	2.7292E-3
10^{-4}	2.5125E-3	2.5906E-3	2.6051E-3	2.6102E-3	2.6125E-3	2.6167E-3
10^{-5}	2.5013E-3	2.5794E-3	2.5938E-3	2.5989E-3	2.6013E-3	2.6054E-3
0	2.5000E-3	2.5781E-3	2.5926E-3	2.5977E-3	2.6000E-3	2.6042E-3

Table 2. $u_{h\varepsilon 1}(1/2)$ obtained by Scheme-2

ε \ N	10	20	30	40	50	$u_{\varepsilon 1}(1/2)$
1	1.2750E-1	1.2758E-1	1.2759E-1	1.2760E-1	1.2760E-1	1.2760E-1
10^{-1}	1.4979E-2	1.5073E-2	1.5090E-2	1.5096E-2	1.5099E-2	1.5104E-2
10^{-2}	3.5577E-3	3.7755E-3	3.8188E-3	3.8342E-3	3.8414E-3	3.8542E-3
10^{-3}	1.4886E-3	2.2586E-3	2.4979E-3	2.5941E-3	2.6411E-3	2.7292E-3
10^{-4}	2.8036E-4	8.4865E-4	1.3587E-3	1.7205E-3	1.9625E-3	2.6167E-3
10^{-5}	3.0894E-5	1.1933E-4	2.5396E-4	4.1966E-4	6.0125E-4	2.6054E-3
10^{-6}	3.1214E-6	1.2441E-5	2.7826E-5	4.9060E-5	7.5853E-5	2.6043E-3
10^{-7}	3.1246E-7	1.2494E-6	2.8095E-6	4.9904E-6	7.7892E-6	2.6042E-3
0	0	0	0	0	0	2.6042E-3

Table 3. $\varepsilon^{-1}h^2 u_{h\varepsilon 1}(1/2)$ obtained by Scheme-2 $(h=1/N)$

ε \ N	10	20	30	40	50
10^{-4}	0.02804	0.02212	0.01510	0.01075	0.00785
10^{-5}	0.03089	0.02983	0.02822	0.02623	0.02405
10^{-6}	0.03121	0.03110	0.03092	0.03066	0.03034
10^{-7}	0.03125	0.03124	0.03122	0.03119	0.03116

N, as is predicted from expansion (84) with $u_{h\varepsilon}^{(0)}=0$ in the present case. Thus we can observe a typical locking phenomenon in the present problem.

Appendix 1. Equivalence of [H1] to [H1]* under [H2]

It is clear that [H1]* implicates [H1]. We will prove the converse when [H2] and (1) hold.

Since $R(B^*)$ is closed from [H2], each $u\in V$ can be uniquely decomposed as

$$u=u_1+u_2 \quad \text{with} \quad u_1\in N(B) \quad \text{and} \quad u_2\in R(B^*).$$

Then we find

$$\begin{aligned} J &\equiv (Au, u)_V+\|Bu\|_W^2 \\ &=(Au_1, u_1)_V+(Au_1, u_2)_V+(Au_2, u_1)_V+(Au_2, u_2)_V+\|Bu_2\|_W^2. \end{aligned}$$

Due to (1), we can define $\alpha\geqq 0$ and $\beta\geqq 0$ by

$$\alpha=(Au_1, u_1)_V^{1/2}, \quad \beta=(Au_2, u_2)_V^{1/2}.$$

Then, as a generalization of the Schwarz inequality, we have

$$|(Au_1, u_2)_V+(Au_2, u_1)_V|\leqq 2\alpha\beta.$$

From the above results, we obtain

$$J\geqq \alpha^2-2\alpha\beta+\beta^2+\|Bu_2\|_W^2.$$

Since $A\neq 0$ from (6) and (7), it holds that

$$2\alpha\beta\leqq \frac{2\|A\|}{2\|A\|+k_2^2}\alpha^2+\frac{2\|A\|+k_2^2}{2\|A\|}\beta^2,$$

where k_2 is the positive number appearing in (8), and $\|A\|$ implies $\|A\|_{\mathcal{L}(V,V)}$. Now J is further evaluated as

$$J \geqq \frac{k_2^2}{2\|A\|+k_2^2}\alpha^2 - \frac{k_2^2}{2\|A\|}\beta^2 + \|Bu_2\|_W^2.$$

Applying the relations

$$\alpha^2 \geqq k_1\|u_1\|_V^2,\ \beta^2 \leqq \|A\|\cdot\|u_2\|_V^2,\ \|Bu_2\|_W \geqq k_2\|u_2\|_V,$$

we have

$$J \geqq \frac{k_1k_2^2}{2\|A\|+k_2^2}\|u_1\|_V^2 + \frac{k_2^2}{2}\|u_2\|_V^2.$$

Since $\|u\|_V^2 = \|u_1\|_V^2 + \|u_2\|_V^2$ due to the orthogonality between $N(B)$ and $R(B^*)$, the above relation implies (10) with

$$k_3 = \min\left(\frac{k_1k_2^2}{2\|A\|+k_2^2},\ \frac{k_2^2}{2}\right) > 0.$$

REMARK. It is easy to extend the present results to the case where $A=0$. In such a case, $N(B)=\{0\}$ from [H1], and hence $R(B^*)=V$. Then [H2]* implies [H1]* with $k_3 = k_2^2$.

Appendix 2. Some results for $[H2]_h$ and $[H3]_h$

As is well known, $[H2]_h$ plays an essential role in numerical analysis of the mixed or hybrid finite element method, but is not necessarily easy to prove in practical examples. We give here some results for this hypothesis as well as for $[H3]_h$.

PROPOSITION A1. (i) *Assume that for each* $h \in \Lambda$ *there exists a subspace* X^h *of* W^h *such that*

$$\|B_h^*\lambda_h\|_V \geqq k_5\|\lambda_h\|_W \quad \text{for any } \lambda_h \in X^h, \tag{A1}$$

where k_5 *is a positive constant independent of* $h \in \Lambda$. *If* $R(B_h) \subset X^h$, *then* $[H2]_h$ *holds with* $k_2^* = k_5$, *and* $X^h = R(B_h)$.
(ii) *Assume that* $[H2]_h$ *and* $[H3]_h$ *hold. Then*

$$\|B_h^*\lambda_h\|_V \geqq \frac{k_2^*}{\|B\|}\|B^*\lambda_h\|_V \quad \text{for all } \lambda_h \in W^h, \tag{A2}$$

where k_2^* *is the positive constant in* $[\mathrm{H2}]_h$, *and* $\|B\|=\|B\|_{\mathcal{L}(V,W)}=\|B^*\|_{\mathcal{L}(W,V)}>0$ *by* (6).
(iii) *Assume that* [H2] *holds and that there exists a positive constant* k_6, *independent of* $h\in\Lambda$, *such that*

$$\|B_h^*\lambda_h\|_V \geqq k_6\|B^*\lambda_h\|_V \quad \textit{for all } \lambda_h\in W^h. \tag{A3}$$

Then $[\mathrm{H3}]_h$ *and* (A1) *hold with* $X^h=R(B)\cap W^h$ *and* $k_5=k_2k_6$, *where* k_2 *is the positive constant in* [H2]. *Therefore, if* $R(B_h)\subset R(B)$, *then* $[\mathrm{H2}]_h$ *holds with* $k_2^*=k_2k_6$ *and* $R(B_h)=R(B)\cap W^h$ *by* (i).
(iv) *Let* D *be a dense subset of* V. *Assume that there exists a mapping* $\pi_h: D\to V^h$ *for each* $h\in\Lambda$ *such that* $\pi_h\neq 0$,

$$\|\pi_h u\|_V \leqq k_7\|u\|_V \quad \textit{for any } u\in D, \tag{A4}$$

$$(\pi_h u, B_h^*\lambda_h)_V=(u, B^*\lambda_h)_V \quad \textit{for any } u\in D,\ \lambda_h\in W^h, \tag{A5}$$

where k_7 *is a positive constant independent of* $h\in\Lambda$. *Then* (A3) *holds with* $k_6=1/k_7$.

REMARK. (a) Notice that (A3) and (iv) are generalizations of assumption (II) of Kikuchi [12] and Proposition 4.1 of Fortin [9], respectively.
(b) By (ii) and (iii), (A1) with $X^h=R(B)\cap W^h$ follows from $[\mathrm{H2}]_h$ and $[\mathrm{H3}]_h$. This is not strange since $R(B)\cap W^h\subset R(B_h)$ from $[\mathrm{H3}]_h$.
(c) We assume that $\pi_h\neq 0$ in (iv). If $\pi_h=0$, then (A5) gives $W^h\subset N(B^*)$, since $(u, B^*\lambda_h)_V=0$ for any $u\in D$ and $\lambda_h\in W^h$, and D is dense in V. Such W^h cannot be used in practice, and the present assumption is reasonable.

PROOF. (i) It is clear that $[\mathrm{H2}]_h$ holds with $k_2^*=k_5$. To show that $X^h=R(B_h)$, we assume its converse: $X^h\supset R(B_h)$ and $X^h\neq R(B_h)$. Then, noting that $N(B_h^*)$ is the orthogonal complement of $R(B_h)$ in W^h, we can find $\lambda_h\in N(B_h^*)\cap X^h$ such that $\lambda_h\neq 0$. Since (A1) applies to this λ_h, we find that

$$0<k_5\|\lambda_h\|_W\leqq\|B_h^*\lambda_h\|_V=0,$$

which is a contradiction. Thus $X^h=R(B_h)$, and the proof is complete.
(ii) Take any $\lambda_h\in W^h$, which can be uniquely decomposed as

$$\lambda_h=\lambda_{h1}+\lambda_{h2} \text{ with } \lambda_{h1}\in N(B_h^*) \text{ and } \lambda_{h2}\in R(B_h).$$

Then, by $[\mathrm{H2}]_h$,

$$\|B_h^*\lambda_h\|_V = \|B_h^*\lambda_{h2}\|_V \geqq k_2^* \|\lambda_{h2}\|_W.$$

On the other hand, we find

$$\|B^*\lambda_h\|_V = \|B^*\lambda_{h2}\|_V \leqq \|B^*\| \cdot \|\lambda_{h2}\|_W,$$

since $\lambda_{h1} \in N(B_h^*) \subset N(B^*)$ by [H3]$_h^*$. From the above two inequalities, we have (A2).
(iii) It is clear from (A3) that $N(B_h^*) \subset N(B^*)$, which is [H3]$_h^*$. From [H2], $\|B^*\lambda_h\|_V \geqq k_2 \|\lambda_h\|_W$ for any $\lambda_h \in R(B) \cap W^h$. Then (A3) implies (A1) with $k_5 = k_2 k_6$ and $X^h = R(B) \cap W^h$.
(iv) Take any $\lambda_h \in W^h$ and any $u \in D$ such that $\pi_h u \neq 0$. Then, by (A4) and (A5), $u \neq 0$ and

$$\frac{|(\pi_h u, B_h^*\lambda_h)_V|}{\|\pi_h u\|_V} = \frac{|(u, B^*\lambda_h)_V|}{\|\pi_h u\|_V} \geqq \frac{|(u, B^*\lambda_h)_V|}{k_7 \|u\|_V}.$$

Clearly, the left-hand side of the above is bounded from above by $\|B_h^*\lambda_h\|_V$. On the other hand,

$$\sup_{u \in D - \{0\}} \frac{|(u, B^*\lambda_h)_V|}{\|u\|_V} = \sup_{u \in V - \{0\}} \frac{|(u, B^*\lambda_h)_V|}{\|u\|_V} = \|B^*\lambda_h\|_V,$$

since D is dense in V. Notice here that the inclusion of $u \in D$ with $\pi_h u = 0$ has no influence on the supremum, since $(u, B^*\lambda_h)_W = 0$ for such u. From these, we obtain (A3) with $k_6 = 1/k_7$, and the proof is complete.

From the above lemma, it appears to be meaningful to see when $R(B_h) \subset R(B) \cap W^h$. There is another reason why the relation between $R(B_h)$ and $R(B) \cap W^h$ is important. When $R(B)$ is closed, it may be regarded as a Hilbert space. Then it is natural to use $R(B) \cap W^h$ as approximation of $R(B)$, since $R(B) \cap W^h$ is of course a finite-dimensional subspace of $R(B)$ and may be used in the finite element calculations. But $R(B_h)$ is also a natural approximation of $R(B)$, and hence arises the necessity of clarifying the relation between these two approximations. The following proposition gives an answer to this problem.

PROPOSITION A2. *Assume that $R(B)$ is closed in W. Then*

$$R(B_h) \subset R(B) \tag{A6}$$

if and only if

$$Q_h N(B^*) \subset N(B_h^*). \tag{A7}$$

Moreover,

$$R(B_h) = R(B) \cap W^h \tag{A8}$$

if and only if

$$Q_h N(B^*) = N(B_h^*). \tag{A9}$$

REMARK. If $[H3]_h$ (or $[H3]_h^*$) holds, then (A6) implies that $N(B_h^*) = Q_h N(B^*) = N(B^*) \cap W^h$ and $R(B_h) = Q_h R(B) = R(B) \cap W^h$.

PROOF. The present lemma can be easily proved if we notice that the orthogonal complements of $N(B_h^*)$ and $Q_h N(B^*)$ in W^h are respectively $R(B_h)$ and $R(B) \cap W^h$.

The above lemma may be useful since null spaces are often easier to deal with than ranges. In some practical examples, $N(B^*)$ is finite-dimensional, and is often $\{0\}$ in typical problems of linear elasticity, see Washizu [19]. Then it is possible to use W^h with $N(B^*) \subset W^h$, and the condition $Q_h N(B^*) \subset N(B_h^*)$ becomes simply $N(B^*) \subset N(B_h^*)$. If $b_h(\cdot, \cdot)$ is the restriction of $b(\cdot, \cdot)$ to $V^h \times W^h$, that is, $B_h^* = P_h B^* | W^h$, such a condition holds automatically. Moreover, $N(B_h^*) = N(B^*)$, if $[H3]_h$ also holds.

Appendix 3. Some results for $[H1]_h^*$

We will give some sufficient conditions for $[H1]_h^*$. We only consider the case where $a_h(\cdot, \cdot)$ is the restriction of $a(\cdot, \cdot)$ to $V^h \times V^h$. Of course we assume $[H1]^*$ to hold. Clearly, $[H1]_h$ follows from $[H1]_h^*$. Note also that, in the present case, $[H1]_h$ follows from [H1] if $N(B_h) \subset N(B)$.

A trivial but still practically important case where $[H1]_h^*$ holds is that there exists a positive constant γ such that

$$a(u, u) \geqq \gamma \|u\|_V^2 \quad \text{for any } u \in V. \tag{A10}$$

Under $[H1]^*$, it is also trivial that $[H1]_h^*$ holds for sufficiently small $h \in \Lambda$, if there exists a positive constant γ^* such that

$$\lim_{h \downarrow 0} \inf_{u_h \in V^h - \{0\}} \frac{\|B_h u_h\|_W - \gamma^* \|B u_h\|_W}{\|u_h\|_V} \geqq 0. \tag{A11}$$

The following result is sometimes useful for our purpose.

PROPOSITION A3. *Assume that B and B_h are respectively in the form of*

$$B=B_1+B_2\,;\quad B_1, B_2\in\mathcal{L}(V, W), \tag{A12}$$

$$B_h=B_1|V^h+B_{h2}\,;\quad B_{h2}\in\mathcal{L}(V^h, W^h). \tag{A13}$$

Moreover, let B_2 be compact and the family of operators $\{B_{h2}\}_{h\in\Lambda}$ approximating B_2 satisfy that, for each $u\in V$ and any family $\{u_h\in V^h\}_{h\in\Lambda}$ converging weakly to u in V for $h\downarrow 0$, $B_{h2}u_h$ always converges strongly to B_2u in W for $h\downarrow 0$. Then [H1]$_h^*$ *holds for sufficiently small $h\in\Lambda$.*

REMARK. For (A13) to be meaningful, it is necessary that $B_1V^h\subset W^h$ for any $h\in\Lambda$. The assumed property for $\{B_{h2}\}_{h\in\Lambda}$ is called the *discrete-compact convergence* of $\{B_{h2}\}_{h\in\Lambda}$ to B_2: see Chatelin [6]. Such a property holds, for example, when B_2 is compact and $B_{h2}=Q_hB_2|V^h$, where $Q_h\in\mathcal{L}(W, W^h)$ is the orthogonal projection operator defined by (32). Concrete examples for which the present proposition is available may be found in Arnold [1] and Kikuchi [13].

PROOF. Assume the contrary. Then there is a subset Λ^* of Λ such that it has zero as an accumulation point and a family $\{u_h\in V^h\}_{h\in\Lambda^*}$ exists with the properties

$$\|u_h\|_V=1,\quad a(u_h, u_h)+\|B_hu_h\|_W^2\to 0\quad (h\in\Lambda^*\to 0).$$

Thus we can find a subset Λ^{**} of Λ^* such that it has zero as an accumulation point and u_h converges weakly in V to a certain $u\in V$ as $h\in\Lambda^{**}$ tends to 0. From the assumptions, we have now

$$a(u_h, u_h)\to 0 \text{ in } \boldsymbol{R}^1,\quad B_hu_h\to 0 \text{ strongly in } W,$$

$$B_1u_h\to B_1u \text{ weakly in } W,\quad B_2u_h\to B_2u \text{ strongly in } W,$$

$$B_{h2}u_h\to B_2u \text{ strongly in } W,$$

as $h\in\Lambda^{**}$ tends to 0. Therefore, $B_1u_h=B_hu_h-B_{h2}u_h$ converges strongly in W to $B_1u=-B_2u$, and hence $Bu_h=B_1u_h+B_2u_h$ converges strongly in W to $Bu=0$. This implies that

$$a(u_h, u_h)+\|Bu_h\|_W^2\to 0\quad (h\in\Lambda^{**}\to 0).$$

On the other hand, from [H1]*, we have

$$a(u_h, u_h)+\|Bu_h\|_W^2 \geqq k_3\|u_h\|_V^2=k_3>0 \quad \text{for any } h\in \Lambda^{**}.$$

Now there arises a contradiction, and the proof is complete.

Appendix 4. Approximation of $N(B)$, $N(B^*)$, and $R(B)$

For $\varepsilon=0$, we have considered the solutions $\{u_\varepsilon, \lambda_\varepsilon\}$ of (5) and $\{u_{h\varepsilon}, \lambda_{h\varepsilon}\}$ of (34) respectively in $V\times R(B)$ and $V^h\times R(B_h)$. However, if λ_ε and $\lambda_{h\varepsilon}$ are looked for respectively in W and W^h, they are completely indefinite in their components respectively in $N(B^*)$ and $N(B_h^*)$. Therefore, we should also clarify the approximation properties of $N(B_h^*)$ to $N(B^*)$. Moreover, for $g=0$, the approximation of $N(B)$ by $N(B_h)$ becomes important as well. We will show that $[\mathrm{H0}]_h$ under $[\mathrm{H2}]_h$ assures the desired properties.

PROPOSITION A4. *Assume that* $[\mathrm{H2}]_h$ *holds. If*

$$\lim_{h\downarrow 0} \inf_{\lambda_h\in W^h} (\|\lambda_h-\lambda\|_W+\|B_h^*\lambda_h-B^*\lambda\|_V)=0 \quad \textit{for each } \lambda\in W, \tag{A14}$$

then

$$\lim_{h\downarrow 0} \inf_{\lambda_h\in N(B_h^*)} \|\lambda_h-\lambda\|_W=0 \quad \textit{for each } \lambda\in N(B^*). \tag{A15}$$

Similarly, if

$$\lim_{h\downarrow 0} \inf_{u_h\in V^h} (\|u_h-u\|_V+\|B_hu_h-Bu\|_W)=0 \quad \textit{for each } u\in V, \tag{A16}$$

then

$$\lim_{h\downarrow 0} \inf_{u_h\in N(B_h)} \|u_h-u\|_V=0 \quad \textit{for each } u\in N(B). \tag{A17}$$

PROOF. We will only prove (A15) under (A14): the other may be proven similarly. From (A14), we can find for each $\lambda\in N(B^*)$ a family of elements $\{\lambda_h\in W^h\}_{h\in\Lambda}$ such that

$$\lim_{h\downarrow 0} (\|\lambda_h-\lambda\|_W+\|B_h^*\lambda_h-B^*\lambda\|_V)=0, \tag{a}$$

where each λ_h can be uniquely decomposed into

$$\lambda_h=\lambda_{h1}+\lambda_{h2}\,; \quad \lambda_{h1}\in N(B_h^*), \quad \lambda_{h2}\in R(B_h).$$

Then, by the relation $B_h^*\lambda_{h1}=B^*\lambda=0$ and $[\mathrm{H2}]_h$, we find

$$\|B_h^*\lambda_h - B^*\lambda\|_V = \|B_h^*\lambda_{h2}\|_V \geqq k_2^* \|\lambda_{h2}\|_W$$

with k_2^* being the positive constant in $[\mathrm{H2}]_h$. Thus (a) gives

$$\|\lambda_{h1} + \lambda_{h2} - \lambda\|_W + k_2^* \|\lambda_{h2}\|_W \to 0 \quad (h \downarrow 0),$$

which implies that

$$\|\lambda_{h1} - \lambda\|_W \to 0 \quad (h \downarrow 0).$$

That is, $\lambda_{h1} \in N(B_h^*)$ converges to $\lambda \in N(B^*)$ as h tends to zero. This is the matter to be proven, and the proof is complete.

We give another result for the approximation capability of $\{N(B_h)\}_{h \in \Lambda}$ to $N(B)$.

PROPOSITION A5. *Assume that there exists for each $h \in \Lambda$ a mapping $\pi_h : V \to V^h$ such that*

$$(B_h \pi_h u, \lambda_h)_W = (Bu, \lambda_h)_W \quad \textit{for any } u \in V \textit{ and } \lambda_h \in W^h, \tag{A18}$$

and

$$\lim_{h \downarrow 0} \|\pi_h u - u\|_V = 0 \quad \textit{for each } u \in V. \tag{A19}$$

Then

$$\lim_{h \downarrow 0} \inf_{u_h \in N(B_h)} \|u_h - u\|_V = 0 \quad \textit{for each } u \in N(B). \tag{A20}$$

REMARK. Condition (A18) for the above π_h corresponds to (A5) for π_h in Proposition A1. Thus it is very convenient if we can find an operator from V into V^h that satisfies all the conditions for these two π_h's: see e.g. Crouzeix-Raviart [8] and Fortin [9].

PROOF. From (A18), $B_h \pi_h u = 0$ or $\pi_h u \in N(B_h)$ for any $u \in N(B)$. Thus, putting $u_h = \pi_h u$ for $u \in N(B)$, we have (A20) from (A19), and the proof is complete.

We can also show that $\{R(B_h)\}_{h \in \Lambda}$ can approximate $R(B)$. This may be seen by applying Theorem 4 with $\varepsilon = 0$, $g = 0$, and $f = B^*\lambda$ for $\lambda \in R(B)$. However, we will show this under assumptions weaker than those employed there.

PROPOSITION A6. *Assume that* $[H3]_h$ *holds. If*

$$\lim_{h\downarrow 0} \inf_{\lambda_h \in W^h} \|\lambda_h - \lambda\|_W = 0 \quad \textit{for each } \lambda \in W, \tag{A21}$$

then

$$\lim_{h\downarrow 0} \inf_{\lambda_h \in R(B_h)} \|\lambda_h - \lambda\|_W = 0 \quad \textit{for each } \lambda \in R(B). \tag{A22}$$

PROOF. For each $\lambda \in R(B)$, we can find, by $[H3]_h$, $\lambda_h \in R(B_h)$ such that $\lambda_h = Q_h\lambda$. Then (A21) assures that $\lambda_h - \lambda = Q_h\lambda - \lambda$ converges to zero in W as $h \downarrow 0$, since $Q_h\lambda$ is the best approximation of λ in W^h. This completes the proof.

References

[1] Arnold, D. N., Discretization by finite elements of a model parameter dependent problem, Numer. Math. **37** (1981), 405-421.

[2] Atluri, S. N., Gallagher, R. H., and O. C. Zienkiewicz, Hybrid and Mixed Finite Element Methods, John Wiley & Sons, Chichester-New York-Brisbane-Toronto-Singapore, 1983.

[3] Babuska, I. and A. K. Aziz, Survey lectures on the mathematical foundations of the finite element method, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations (A. K. Aziz, ed.), Academic Press, New York-London, 1972, 3-359.

[4] Bercovier, M., Perturbation of mixed variational problems. Application to mixed finite element methods, RAIRO Anal. Numér. **12** (1978), 211-236.

[5] Brezzi, F., On the existence, uniqueness and approximation of saddle point problems arising from Lagrangian multipliers, RAIRO **8** (1974), 129-151.

[6] Chatelin, F., Spectral Approximation of Linear Operators, Academic Press, New York-London, 1983.

[7] Ciarlet, P. G., The Finite Element Method for Elliptic Problems, North-Holland Publishing Company, Amsterdam-New York-Oxford, 1978.

[8] Crouzeix, M. and P.-A. Raviart, Conforming and nonconforming finite element methods for solving the stationary Stokes equations I, RAIRO **7** (1973), 33-76.

[9] Fortin, M., An analysis of the convergence of mixed finite element methods, RAIRO Anal. Numér. **11** (1977), 341-354.

[10] Fortin, M. and R. Glowinski, Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems, North-Holland Publishing Company, Amsterdam-New York-Oxford, 1983.

[11] Girault, V. and P.-A. Raviart, Finite Element Approximation of the Navier-Stokes Equations, Lecture Notes in Math. No. 749, Springer Verlag, Berlin-Heidelberg-New York, 1979.

[12] Kikuchi, F., Some considerations on the convergence of hybrid stress method, in Theory and Practice in Finite Element Structural Analysis (Y. Yamada and R. H. Gallagher, eds.), University of Tokyo Press, Tokyo, 1973, 25-42.

[13] Kikuchi, F., Accuracy of some finite element models for arch problems, Comput. Methods Appl. Mech. Engrg. **35** (1982), 315-345.
[14] Lions, J. L., Perturbations Singulières dans les Problèmes aux Limites et en Contrôle Optimal, Lecture Notes in Math. No. 323, Springer Verlag, Berlin-Heidelberg-New York, 1973.
[15] Malkus, D. S. and T. J. R. Hughes, Mixed finite element methods—reduced and selective integration techniques: a unification of concepts, Comput. Methods Appl. Mech. Engrg. **15**, (1978), 63-81.
[16] Oden, J. T. and N. Kikuchi, Finite element methods for constrained problems in elasticity, Internat. J. Numer. Methods Engrg. **18** (1982), 701-725.
[17] Temam, R., Navier-Stokes Equations, North-Holland Publishing Company, Amsterdam-New York-Oxford, 1977.
[18] Yosida, K., Functional Analysis, Springer Verlag, Berlin-Heidelberg-New York, 1965.
[19] Washizu, K., Variational Methods in Elasticity and Plasticity, 2nd edition, Pergamon Press, Oxford-New York-Toronto-Sydney, 1974.

(Received February 28, 1985)

Department of Mathematics
College of Arts and Sciences
University of Tokyo
Komaba, Tokyo
153 Japan