View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by CiteSeerX

ELSEVIER Applied Mathematics and Computation 115 (2000) 63–75
APPLIED MATHEMATICS AND COMPUTATION
www.elsevier.com/locate/amc

Least-squares finite element approximations to the Timoshenko beam problem

Jang Jou [a,b,*], Suh-Yuh Yang [c,1]

[a] *Department of Applied Mathematics, National Chiao Tung University, 1001, Ta Hsueh Road, Hsinchu 30050, Taiwan, ROC*
[b] *Department of Statistics, Ming Chuan University, Taipei 11120, Taiwan, ROC*
[c] *Department of Applied Mathematics, I-Shou University, Ta-Hsu, Kaohsiung 84008, Taiwan, ROC*

Abstract

In this paper a least-squares finite element method for the Timoshenko beam problem is proposed and analyzed. The method is shown to be convergent and stable without requiring extra smoothness of the exact solutions. For sufficiently regular exact solutions, the method achieves optimal order of convergence in the H^1-norm for all the unknowns (displacement, rotation, shear, moment), uniformly in the small parameter which is generally proportional to the ratio of thickness to length. Thus the locking phenomenon disappears as the parameter tends to zero. A sharp a posteriori error estimator which is exact in the energy norm and equivalent in the H^1-norm is also briefly discussed. © 2000 Published by Elsevier Science Inc. All rights reserved.

AMS classification: 65N15; 65N30
Keywords: Timoshenko beam problem; Least-squares; Finite element method; Locking phenomenon; A posteriori error estimator

1. Introduction

The locking phenomenon is mainly concerned in the finite element analysis for parameter-dependent problems [2,5]. In this paper we shall propose and analyze a finite element method based on the least-squares principles for

* Corresponding author. E-mail: jangjou@math.nctu.edu.tw
1 The work of this author was supported in part by NSC-grant 87-2811-M-007-0017, Taiwan, ROC.

0096-3003/00/$ - see front matter © 2000 Published by Elsevier Science Inc. All rights reserved.
PII: S0096-3003(99)00139-3

approximating the solution of the Timoshenko beam problem in its first-order system formulation. The method avoids the locking problem as the thickness of the beam tends to zero, and achieves optimal order error estimates for all the unknowns (displacement, rotation, shear, moment) in the H^1-norm.

Finite element analysis of the Timoshenko beam problem has been frequently used as a starting point for a better understanding of the much more complex problem of constructing accurate finite element approximations to the Reissner–Mindlin plate problem. It is well-known that some bad behaviors may occur such as the locking phenomenon when we solve these problems with standard Galerkin finite element methods [2,19,26,27]. That is, the convergence results of the standard finite element approximations deteriorate as the small parameters (which depend on the thickness of the beam and the plate for the Timoshenko beam model and the Reissner–Mindlin plate model, respectively) tend to zero. The most widely used effective approach to overcome the difficulty is based on the reduced integration technique. A complete analysis for the Timoshenko beam problem by using this technique has been addressed in [2]. Recently, a modified reduced integration method with linear finite elements is proposed and analyzed by Cheng et al. [19]. The method presented in [19] uses the reduced integration technique to compute the term involving the small parameter and adds a bubble function space to the rotation to increase the solution accuracy. This method can also be applied to solve the circular arch problem and the Reissner–Mindlin plate problem. Another way to eliminate the locking phenomenon is to use the p and $h-p$ versions of the finite element method for which optimal error estimates are established in [25].

In the present investigation we provide an alternate way to avoid the difficulty by exploiting the least-squares principles on a first-order system formulation of the Timoshenko beam problem. We first introduce a quadratic least-squares functional $\mathscr{Q}$ over a function space $\mathscr{V}$ consisting of functions which satisfy the boundary conditions of the problem. The functional is defined to be the sum of the squared L^2-norms of the residuals in the differential equations. Then the exact solution must be the unique zero minimizer of the functional $\mathscr{Q}$ over $\mathscr{V}$. Therefore, the least-squares finite element approximate solution is defined to be the minimizer of $\mathscr{Q}$ over a finite-dimensional subspace $\mathscr{V}_h^p$ of $\mathscr{V}$. Mathematical analyses show this approach can eliminate the locking phenomenon.

Over the past decade, the use of least-squares principles in connection with finite element techniques has been extensively applied to the approximations in many different fields such as fluid dynamics [9,12,13,16–18], elasticity [14,29,30], electromagnetism [15,24], and semiconductor device physics [8]. The approach offers certain advantages, especially for large-scale computations. It leads to minimization problems rather than saddle point problems led by the mixed finite element approach. A single continuous piecewise polynomial space can be used for the approximation of all unknowns, and accurate approximations of all unknowns can be simultaneously obtained. The resulting algebraic system is symmetric and positive definite. In addition, the value of the least-squares functional of the approximate solution provides a practical and sharp a posteriori error estimator at no additional cost [21,23,28]. This feature is very important in adaptive computations.

The layout of the remainder of the paper is as follows. In Section 2, we introduce the first-order system formulation for the Timoshenko beam problem and we then derive important a priori estimates for the system. The least-squares finite element method is given in Section 3. Then the method is proved to be convergent and stable in Section 4, where optimal error estimates in the H^1-norm are also established. Finally, in Section 5, we briefly discuss the sharp a posteriori error estimator for the least-squares approach.

2. The Timoshenko beam problem

In this paper we shall consider the in-plane bending of a clamped uniform beam of length L, cross-section A, moment of inertia I, Young's modulus E, and shear modulus G, subjected to a distributed load $p(\bar{x})$, and a distributed moment $m(\bar{x})$, with $\bar{x} \in (0, L)$ representing the independent variable. According to the Timoshenko beam theory, this problem is governed by the following system of ordinary differential equations of first-order (cf. [27]):

$$-\frac{\mathrm{d}Q}{\mathrm{d}\bar{x}} = p \quad \text{in } (0, L), \tag{2.1}$$

$$-\frac{\mathrm{d}M}{\mathrm{d}\bar{x}} - Q = m \quad \text{in } (0, L), \tag{2.2}$$

$$-\frac{Q}{\kappa GA} + \frac{\mathrm{d}w}{\mathrm{d}\bar{x}} - \theta = 0 \quad \text{in } (0, L), \tag{2.3}$$

$$-\frac{M}{EI} + \frac{\mathrm{d}\theta}{\mathrm{d}\bar{x}} = 0 \quad \text{in } (0, L), \tag{2.4}$$

supplemented with the boundary conditions:

$$w(0) = w(L) = 0, \tag{2.5}$$

$$\theta(0) = \theta(L) = 0, \tag{2.6}$$

where $Q(\bar{x})$ is the shear force; $M(\bar{x})$ the bending moment; $w(\bar{x})$ the transverse displacement; $\theta(\bar{x})$ the cross-section rotation; and κ the shear correction factor.

To explicate the dependence of this problem on a small parameter,

$$\varepsilon^2 = \frac{EI}{\kappa GAL^2}, \tag{2.7}$$

we introduce the following change of variables:

$$x = \frac{\bar{x}}{L}, \tag{2.8}$$

$$u_1 = \frac{w}{L}, \quad u_2 = \theta, \tag{2.9}$$

$$\sigma_1 = \frac{QL^2}{EI}, \quad \sigma_2 = \frac{ML}{EI}, \tag{2.10}$$

$$f_1 = \frac{pL^3}{EI}, \quad f_2 = \frac{mL^2}{EI}, \tag{2.11}$$

which reduces the original problem to finding $\mathbf{u}(x) = (u_1(x), u_2(x))$ and $\boldsymbol{\sigma}(x) = (\sigma_1(x), \sigma_2(x))$, $x \in (0,1)$ satisfying

$$-\sigma_1' = f_1 \quad \text{in } (0,1), \tag{2.12}$$

$$-\sigma_2' - \sigma_1 = f_2 \quad \text{in } (0,1), \tag{2.13}$$

$$-\varepsilon^2\sigma_1 + u_1' - u_2 = 0 \quad \text{in } (0,1), \tag{2.14}$$

$$-\sigma_2 + u_2' = 0 \quad \text{in } (0,1), \tag{2.15}$$

with the boundary conditions

$$u_1(0) = u_1(1) = 0, \tag{2.16}$$

$$u_2(0) = u_2(1) = 0, \tag{2.17}$$

where the prime superscript denotes differentiation with respect to the nondimensional variable x. We observe that the nondimensional problem (2.12)–(2.17) depends explicitly on a parameter ε. In general, ε is a small parameter proportional to the ratio of the thickness to length. For instance, in the rectangular cross-section case [26],

$$\varepsilon^2 = \frac{E}{12\kappa G}\left(\frac{T}{L}\right)^2,$$

where T represents the thickness of the beam. Thus in most realistic applications, e.g., for thin beams, $0 < \varepsilon \ll 1$, and the construction of accurate finite element approximations is delicate.

We need some function spaces throughout this paper. The classical Sobolev spaces $H^r(0,1)$ with their associated norms $\|\cdot\|_r$ are employed [11,20]. As usual, we denote by $(\cdot,\cdot)_0$ and $\|\cdot\|_0$ the conventional inner product and norm on the Hilbert space $L^2(0,1)$ of square-integrable functions. The space $H^1(0,1)$ of functions which together with their first derivatives are square-integrable, is a Hilbert space with the inner product $(\cdot,\cdot)_1$, where

$$(u,v)_1 = \int_0^1 (uv + u'v')\,\mathrm{d}x, \quad u,v \in H^1(0,1).$$

The norm generated by this inner product is denoted by $\|\cdot\|_1$; that is, $\|v\|_1 = (v,v)_1^{1/2}$. We denote by $H_0^1(0,1)$ the subspace of $H^1(0,1)$ consisting of functions which vanish at the ends of the interval. For the product space $[H^1(0,1)]^4$, the corresponding inner product and norm are also denoted by $(\cdot,\cdot)_1$ and $\|\cdot\|_1$, respectively, when there is no chance of confusion.

We shall use C with or without subscripts in this paper to denote a generic positive constant, possibly different at different occurrences, that is independent of the parameter ε and the mesh parameter h introduced in the next section.

The following regularity result plays an important role in the theoretical analysis later.

Theorem 2.1. *Let* $(\mathbf{v},\boldsymbol{\omega}) = (v_1, v_2, \omega_1, \omega_2) \in [H_0^1(0,1)]^2 \times [H^1(0,1)]^2$. *Then there exist two positive constants* C_1 *and* C_2 *both independent of the parameter* ε *such that*

$$C_1\left(\|v_1\|_1^2 + \|v_2\|_1^2 + \|\omega_1\|_1^2 + \|\omega_2\|_1^2\right) \leqslant \mathcal{Q}(\mathbf{v},\boldsymbol{\omega};\mathbf{0}) \tag{2.18}$$

and

$$\mathcal{Q}(\mathbf{v},\boldsymbol{\omega};\mathbf{0}) \leqslant C_2\left(\|v_1\|_1^2 + \|v_2\|_1^2 + \|\omega_1\|_1^2 + \|\omega_2\|_1^2\right), \tag{2.19}$$

where

$$\mathcal{Q}(\mathbf{v},\boldsymbol{\omega};\mathbf{0}) := \|\omega_1'\|_0^2 + \|\omega_2' + \omega_1\|_0^2 + \|\varepsilon^2\omega_1 - v_1' + v_2\|_0^2 + \|\omega_2 - v_2'\|_0^2.$$

Proof. Upper bound (2.19) is straightforward from the triangle and Cauchy–Schwarz inequalities. To prove the lower bound (2.18), we first define

$$g_1 := -\omega_1', \tag{2.20}$$

$$g_2 := -\omega_2' - \omega_1, \tag{2.21}$$

$$g_3 := -\varepsilon^2\omega_1 + v_1' - v_2, \tag{2.22}$$

$$g_4 := -\omega_2 + v_2', \tag{2.23}$$

in $[0,1]$. Then from the first Eq. (2.20), we find

$$\omega_1(x) = -G_1(x) + c_1,$$

where

$$G_1(x) := \int_0^x g_1(t)\,\mathrm{d}t, \quad x \in [0,1], \quad c_1 := \omega_1(0).$$

Integrating the second Eq. (2.21), we have

$$\omega_2(x) = -G_2(x) + \int_0^x G_1(t)\,\mathrm{d}t - c_1 x + c_2,$$

where

$$G_2(x) := \int_0^x g_2(t)\,\mathrm{d}t, \quad x \in [0,1], \quad c_2 := \omega_2(0).$$

Solving the last Eq. (2.23) together with the boundary condition $v_2(0) = 0$, we obtain

$$v_2(x) = -\int_0^x G_2(s)\,\mathrm{d}s + \int_0^x \int_0^s G_1(t)\,\mathrm{d}t\,\mathrm{d}s - \frac{1}{2}c_1x^2 + c_2x + G_4(x),$$

where

$$G_4(x) := \int_0^x g_4(t)\,\mathrm{d}t, \quad x, s \in [0, 1].$$

Finally, from the third Eq. (2.22) and the boundary condition $v_1(0) = 0$, we find for $x, y, s \in [0, 1]$

$$\begin{aligned} v_1(x) = &-\varepsilon^2 \int_0^x G_1(t)\,\mathrm{d}t + \varepsilon^2 c_1 x + G_3(x) - \int_0^x \int_0^y G_2(s)\,\mathrm{d}s\,\mathrm{d}y \\ &+ \int_0^x \int_0^y \int_0^s G_1(t)\,\mathrm{d}t\,\mathrm{d}s\,\mathrm{d}y - \frac{1}{6}c_1x^3 + \frac{1}{2}c_2x^2 + \int_0^x G_4(y)\,\mathrm{d}y, \end{aligned}$$

where

$$G_3(x) := \int_0^x g_3(t)\,\mathrm{d}t, \quad x \in [0, 1].$$

Utilizing the boundary conditions $v_2(1) = 0$ and $v_1(1) = 0$, we get the following 2×2 linear system of equations in the unknowns c_1 and c_2:

$$\begin{aligned} -\frac{1}{2}c_1 + c_2 &= \int_0^1 G_2(s)\,\mathrm{d}s - \int_0^1 \int_0^s G_1(t)\,\mathrm{d}t\,\mathrm{d}s - G_4(1) \\ (\varepsilon^2 - \frac{1}{6})c_1 + \frac{1}{2}c_2 &= \varepsilon^2 \int_0^1 G_1(t)\,\mathrm{d}t - G_3(1) + \int_0^1 \int_0^y G_2(s)\,\mathrm{d}s\,\mathrm{d}y \\ &\quad - \int_0^1 \int_0^y \int_0^s G_1(t)\,\mathrm{d}t\,\mathrm{d}s\,\mathrm{d}y - \int_0^1 G_4(y)\,\mathrm{d}y. \end{aligned}$$

Solving this linear system, we can find the following estimates for c_1 and c_2, respectively,

$$\begin{aligned} |c_1| &\leqslant C\sum_{i=1}^4 \|g_i\|_0, \\ |c_2| &\leqslant C\sum_{i=1}^4 \|g_i\|_0 \end{aligned}$$

for some constant C independent of the parameter ε. Now it is easy to see that

$$\begin{aligned} \|v_i\|_1^2 &\leqslant C\mathscr{Q}(\mathbf{v}, \boldsymbol{\omega}; \mathbf{0}), \\ \|\omega_i\|_1^2 &\leqslant C\mathscr{Q}(\mathbf{v}, \boldsymbol{\omega}; \mathbf{0}), \end{aligned}$$

for $i = 1, 2$. Thus the assertion (2.18) follows immediately and this completes the proof. □

3. The least-squares finite element method

In this section we introduce the least-squares finite element method for problem (2.12)–(2.17). Define a function space $\mathscr{V}$ for our problem by

$$\mathscr{V} = H_0^1(0,1) \times H_0^1(0,1) \times H^1(0,1) \times H^1(0,1) \tag{3.1}$$

and then define a quadratic least-squares energy functional $\mathscr{Q}$: $\mathscr{V} \to \mathbf{R}$ by

$$\mathscr{Q}(\mathbf{v}, \boldsymbol{\omega}; \mathbf{f}) = \| -\omega_1' - f_1\|_0^2 + \| -\omega_2' - \omega_1 - f_2\|_0^2 + \| -\varepsilon^2\omega_1 + v_1' - v_2\|_0^2 + \| -\omega_2 + v_2'\|_0^2, \tag{3.2}$$

where $\mathbf{v} = (v_1, v_2)$, $\boldsymbol{\omega} = (\omega_1, \omega_2)$, and $\mathbf{f} = (f_1, f_2)$. Note that the quadratic energy functional $\mathscr{Q}(\mathbf{v}, \boldsymbol{\omega}; \mathbf{f})$ is defined to be the sum of the squared L^2-norms of the residuals in the differential equations. Obviously, the exact solution $(\mathbf{u}, \boldsymbol{\sigma}) \in \mathscr{V}$ of problem (2.12)–(2.17) is the unique zero minimizer of the functional $\mathscr{Q}$ on $\mathscr{V}$, that is,

$$\mathscr{Q}(\mathbf{u}, \boldsymbol{\sigma}; \mathbf{f}) = 0 = \min\{\mathscr{Q}(\mathbf{v}, \boldsymbol{\omega}; \mathbf{f}):\ (\mathbf{v}, \boldsymbol{\omega}) \in \mathscr{V}\}. \tag{3.3}$$

Applying the variational techniques, we can find that (3.3) is equivalent to

$$\mathscr{B}((\mathbf{u}, \boldsymbol{\sigma}), (\mathbf{v}, \boldsymbol{\omega})) = \mathscr{F}((\mathbf{v}, \boldsymbol{\omega})) \quad \forall\ (\mathbf{v}, \boldsymbol{\omega}) \in \mathscr{V}, \tag{3.4}$$

where the bilinear form $\mathscr{B}(\cdot,\cdot)$ and the linear form $\mathscr{F}(\cdot)$ are defined, respectively, by

$$\begin{aligned}\mathscr{B}((\mathbf{v}, \boldsymbol{\omega}), (\mathbf{z}, \boldsymbol{\varrho})) = \int_0^1 & (\omega_1'\varrho_1' + (\omega_2' + \omega_1)(\varrho_2' + \varrho_1) \\ & + (-\varepsilon^2\omega_1 + v_1' - v_2)(-\varepsilon^2\varrho_1 + z_1' - z_2) \\ & + (-\omega_2 + v_2')(-\varrho_2 + z_2'))\,\mathrm{d}x, \end{aligned} \tag{3.5}$$

$$\mathscr{F}((\mathbf{v}, \boldsymbol{\omega})) = \int_0^1 \big(-\omega_1' f_1 + (-\omega_2' - \omega_1) f_2\big)\,\mathrm{d}x \tag{3.6}$$

for all $(\mathbf{v}, \boldsymbol{\omega})$, $(\mathbf{z}, \boldsymbol{\varrho}) \in \mathscr{V}$.

It is evidently that $\mathscr{B}(\cdot,\cdot)$ is symmetric and continuous (bounded) on $\mathscr{V} \times \mathscr{V}$ and, for each given $\mathbf{f} \in [L^2(0,1)]^2$, $\mathscr{F}(\cdot)$ is also continuous (bounded) on $\mathscr{V}$. Furthermore, by (2.18) and (2.19), we have

$$\begin{aligned} C_1\left(\|v_1\|_1^2+\|v_2\|_1^2+\|\omega_1\|_1^2+\|\omega_2\|_1^2\right) &\leqslant \mathscr{B}((\mathbf{v},\boldsymbol{\omega}),(\mathbf{v},\boldsymbol{\omega})) \\ &= \mathscr{Q}(\mathbf{v},\boldsymbol{\omega};\mathbf{0}) \\ &\leqslant C_2\left(\|v_1\|_1^2+\|v_2\|_1^2+\|\omega_1\|_1^2+\|\omega_2\|_1^2\right) \end{aligned} \tag{3.7}$$

for all $(\mathbf{v},\boldsymbol{\omega}) \in \mathscr{V}$. Thus $\mathscr{B}(\cdot,\cdot)$ is coercive on $\mathscr{V}\times\mathscr{V}$ and

$$\|(\mathbf{v},\boldsymbol{\omega})\|_{\mathscr{B}} = (\mathscr{B}((\mathbf{v},\boldsymbol{\omega}),(\mathbf{v},\boldsymbol{\omega})))^{1/2} \quad \forall\ (\mathbf{v},\boldsymbol{\omega}) \in \mathscr{V} \tag{3.8}$$

define a new norm on $\mathscr{V}$ which is equivalent to the H^1-norm.

For the purpose of discretization we shall use finite element spaces defined with reference to partitions of $[0,1]$. Let $\mathscr{T}_h$ be a partition of $[0,1]$ such that the interval $[0,1]$ is partitioned into subintervals $I_i=(x_{i-1},x_i)$, $i=1,\ldots,N$, with $0=x_0<x_1<\cdots<x_N=1$. The mesh parameter h is defined by $h=\max\{|x_i-x_{i-1}|:\ i=1,\ldots,N\}$. We denote by $\mathscr{P}_h^p$, $p\geqslant 1$ integer, the space of continuous functions on $[0,1]$ whose restrictions to I_i, $i=1,\cdots,N$, are polynomials of degree p, that is,

$$\mathscr{P}_h^p=\left\{v\in C^0(0,1):\ v|_{I_i}\text{is a polynomial of degree } p\right\}.$$

Define

$$\tilde{\mathscr{P}}_h^p=\mathscr{P}_h^p\cap H_0^1(0,1).$$

Then we will seek the least-squares finite element approximations in the following finite-dimensional subspace of $\mathscr{V}$,

$$\mathscr{V}_h^p=\tilde{\mathscr{P}}_h^p\times\tilde{\mathscr{P}}_h^p\times\mathscr{P}_h^p\times\mathscr{P}_h^p. \tag{3.9}$$

By the interpolation theory, the finite element space possesses the following approximation property: for any $(\mathbf{v},\boldsymbol{\omega})\in\mathscr{V}\cap[H^{p+1}(0,1)]^4$, there exists $(\mathbf{v}_h,\boldsymbol{\omega}_h)\in\mathscr{V}_h^p$ such that

$$\|(\mathbf{v},\boldsymbol{\omega})-(\mathbf{v}_h,\boldsymbol{\omega}_h)\|_1\leqslant Ch^p\|(\mathbf{v},\boldsymbol{\omega})\|_{p+1}, \tag{3.10}$$

where the positive constant C is independent of $(\mathbf{v},\boldsymbol{\omega})$ and the mesh parameter h.

The least-squares finite element method for problem (2.12)–(2.17) is then the following.

Find $(\mathbf{u}_h,\boldsymbol{\sigma}_h)\in\mathscr{V}_h^p$ such that

$$\mathscr{B}((\mathbf{u}_h,\boldsymbol{\sigma}_h),(\mathbf{v}_h,\boldsymbol{\omega}_h))=\mathscr{F}((\mathbf{v}_h,\boldsymbol{\omega}_h))\quad\forall\ (\mathbf{v}_h,\boldsymbol{\omega}_h)\in\mathscr{V}_h^p. \tag{3.11}$$

Applying the Lax–Milgram theorem [11,20], we know that the approximation problem (3.11) has a unique solution. Once a basis for the space $\mathscr{V}_h^p$ is chosen, the matrix associated with problem (3.11) can easily be shown to be symmetric and positive definite.

4. Stability, convergence and error estimates

Let $(\mathbf{u},\boldsymbol{\sigma}) \in \mathscr{V}$ be the exact solution of problem (2.12)–(2.17) with the given function $\mathbf{f} \in [L^2(0,1)]^2$. We first prove the stability of the least-squares finite element solution $(\mathbf{u}_h,\boldsymbol{\sigma}_h) \in \mathscr{V}_h^p$.

Theorem 4.1. *The unique solution* $(\mathbf{u}_h,\boldsymbol{\sigma}_h) \in \mathscr{V}_h^p$ *of problem* (3.11) *is stable in the* H^1*-norm in the following sense*:

$$\|(\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_1 \leqslant C\|\mathbf{f}\|_0, \tag{4.1}$$

where the positive constant C *is independent of* ε *and* h.

Proof. By (3.7), (3.8) and (3.11), we have

$$\begin{aligned} C_1\|(\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_1^2 &\leqslant \|(\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_{\mathscr{B}}^2 \\ &= \mathscr{B}((\mathbf{u}_h,\boldsymbol{\sigma}_h),(\mathbf{u}_h,\boldsymbol{\sigma}_h)) \\ &= \mathscr{F}((\mathbf{u}_h,\boldsymbol{\sigma}_h)) \\ &\leqslant \|\mathbf{f}\|_0\|(\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_{\mathscr{B}} \\ &\leqslant C\|\mathbf{f}\|_0\|(\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_1, \end{aligned}$$

which implies (4.1). This completes the proof. □

Estimate (4.1) indicates that the least-squares finite element solution $(\mathbf{u}_h,\boldsymbol{\sigma}_h)$ is stable with respect to the H^1-norm, that is, when we change the given data function $\mathbf{f}$ slightly in the L^2-norm, the least-squares solution $(\mathbf{u}_h,\boldsymbol{\sigma}_h)$ changes only slightly in the H^1-norm.

We now introduce some function spaces which will be used to prove the next theorem. Let $C_0^\infty(0,1)$ denote the linear space of infinitely differentiable functions with compact support in $(0,1)$, and let $C^\infty[0,1]$ denote the restrictions of the functions in $C_0^\infty(\mathbf{R})$ to $[0,1]$. Then it is obvious that the product space

$$\mathscr{S} := [C_0^\infty(0,1)]^2 \times [C^\infty[0,1]]^2 \tag{4.2}$$

is dense in $\mathscr{V}$ with respect to the H^1-norm.

Now utilizing the standard density argument [20], we can obtain the following results for the convergence.

Theorem 4.2. *The least-squares finite element solution* $(\mathbf{u}_h,\boldsymbol{\sigma}_h)$ *is convergent in the* H^1*-norm without requiring any extra regularity assumption on the exact solution* $(\mathbf{u},\boldsymbol{\sigma})$, *that is*,

$$\lim_{h\to 0}\|(\mathbf{u},\boldsymbol{\sigma}) - (\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_1 = 0. \tag{4.3}$$

Moreover, *if the exact solution* $(\mathbf{u},\boldsymbol{\sigma}) \in \mathscr{V} \cap [H^{p+1}(0,1)]^4$, *then we have the following error estimate*:

$$\|(\mathbf{u},\boldsymbol{\sigma})-(\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_1 \leqslant Ch^p\|(\mathbf{u},\boldsymbol{\sigma})\|_{p+1}, \tag{4.4}$$

where C is a positive constant independent of ε and h.

Proof. Subtracting the equation in (3.11) from the equation in (3.4), we get the following orthogonality relation:

$$\mathscr{B}((\mathbf{u},\boldsymbol{\sigma})-(\mathbf{u}_h,\boldsymbol{\sigma}_h),(\mathbf{v}_h,\boldsymbol{\omega}_h))=0 \quad \forall\ (\mathbf{v}_h,\boldsymbol{\omega}_h)\in\mathscr{V}_h^p. \tag{4.5}$$

Using (3.7), (4.5) and the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}\|(\mathbf{u},\boldsymbol{\sigma})-(\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_1^2 &\leqslant \frac{1}{C_1}\|(\mathbf{u},\boldsymbol{\sigma})-(\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_{\mathscr{B}}^2\\ &= \frac{1}{C_1}\mathscr{B}((\mathbf{u},\boldsymbol{\sigma})-(\mathbf{u}_h,\boldsymbol{\sigma}_h),(\mathbf{u},\boldsymbol{\sigma})-(\mathbf{u}_h,\boldsymbol{\sigma}_h))\\ &= \frac{1}{C_1}\mathscr{B}((\mathbf{u},\boldsymbol{\sigma})-(\mathbf{u}_h,\boldsymbol{\sigma}_h),(\mathbf{u},\boldsymbol{\sigma})-(\mathbf{v}_h,\boldsymbol{\omega}_h))\\ &\leqslant \frac{C}{C_1}\|(\mathbf{u},\boldsymbol{\sigma})-(\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_1\|(\mathbf{u},\boldsymbol{\sigma})-(\mathbf{v}_h,\boldsymbol{\omega}_h)\|_1\end{aligned}$$

for all $(\mathbf{v}_h,\boldsymbol{\omega}_h)\in\mathscr{V}_h^p$. Thus,

$$\|(\mathbf{u},\boldsymbol{\sigma})-(\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_1 \leqslant C\|(\mathbf{u},\boldsymbol{\sigma})-(\mathbf{v}_h,\boldsymbol{\omega}_h)\|_1 \quad \forall\ (\mathbf{v}_h,\boldsymbol{\omega}_h)\in\mathscr{V}_h^p. \tag{4.6}$$

Now, since the subspace $\mathscr{S}\subset\mathscr{V}\cap[H^{p+1}(0,1)]^4$ is dense in $\mathscr{V}$ with respect to the H^1-norm, for any $\delta>0$, there exists $(\tilde{\mathbf{u}},\tilde{\boldsymbol{\sigma}})\in\mathscr{S}$ independent of h such that

$$\|(\mathbf{u},\boldsymbol{\sigma})-(\tilde{\mathbf{u}},\tilde{\boldsymbol{\sigma}})\|_1 < \frac{\delta}{2C}, \tag{4.7}$$

where C is the same constant as in (4.6). For this fixed sufficiently smooth function $(\tilde{\mathbf{u}},\tilde{\boldsymbol{\sigma}})\in\mathscr{S}\subset[H^{p+1}(0,1)]^4$, by the approximation property (3.10), we can find $(\tilde{\mathbf{u}}_h,\tilde{\boldsymbol{\sigma}}_h)\in\mathscr{V}_h^p$ so that,

$$\|(\tilde{\mathbf{u}},\tilde{\boldsymbol{\sigma}})-(\tilde{\mathbf{u}}_h,\tilde{\boldsymbol{\sigma}}_h)\|_1 \leqslant Ch^p\|(\tilde{\mathbf{u}},\tilde{\boldsymbol{\sigma}})\|_{p+1},$$

which implies, for sufficiently small h,

$$\|(\tilde{\mathbf{u}},\tilde{\boldsymbol{\sigma}})-(\tilde{\mathbf{u}}_h,\tilde{\boldsymbol{\sigma}}_h)\|_1 < \frac{\delta}{2C}, \tag{4.8}$$

where C is the same constant as in (4.6). Combining inequalities (4.7) and (4.8) with (4.6), we immediately obtain

$$\begin{aligned}0\leqslant\|(\mathbf{u},\boldsymbol{\sigma})-(\mathbf{u}_h,\boldsymbol{\sigma}_h)\|_1 &\leqslant C\|(\mathbf{u},\boldsymbol{\sigma})-(\tilde{\mathbf{u}}_h,\tilde{\boldsymbol{\sigma}}_h)\|_1\\ &\leqslant C\Big(\|(\mathbf{u},\boldsymbol{\sigma})-(\tilde{\mathbf{u}},\tilde{\boldsymbol{\sigma}})\|_1+\|(\tilde{\mathbf{u}},\tilde{\boldsymbol{\sigma}})-(\tilde{\mathbf{u}}_h,\tilde{\boldsymbol{\sigma}}_h)\|_1\Big)\\ &< \delta,\end{aligned}$$

which implies (4.3).

We now assume that $(\mathbf{u},\boldsymbol{\sigma}) \in \mathscr{V} \cap [H^{p+1}(0,1)]^4$. By (4.6) and the approximation property (3.10) of the finite element space $\mathscr{V}_h^p$, we can obtain (4.4) immediately. This completes the proof. □

Since the energy norm $\|\cdot\|_{\mathscr{B}}$ is equivalent to the H^1-norm, from (4.3), we can conclude the following consequence.

Corollary 4.3. *Let $(\mathbf{u}_h,\boldsymbol{\sigma}_h) = (u_{1h}, u_{2h}, \sigma_{1h}, \sigma_{2h})$ be the least-squares finite element solution. Then*

$$\begin{aligned}\lim_{h\to 0}\big(&\|-\sigma'_{1h} - f_1\|_0 + \|-\sigma'_{2h} - \sigma_{1h} - f_2\|_0 + \\ &\|-\varepsilon^2\sigma_{1h} + u'_{1h} - u_{2h}\|_0 + \|-\sigma_{2h} + u'_{2h}\|_0\big) = 0.\end{aligned} \tag{4.9}$$

Remark 4.4. The error estimate (4.4) is optimal in the H^1-norm with respect to the order of approximation of the finite element space $\mathscr{V}_h^p$. Under suitable assumptions (cf. [22]), it is quite possible to derive the optimal error estimates in the L^2-norm for all the unknowns by using the Aubin–Nitsche trick.

5. An equivalent a posteriori error estimator in the H^1-norm

The use of a posteriori error estimators has become an accepted tool for assessing and controlling computational errors in adaptive computations. One of the most important advantageous features of the least-squares finite element approach is that the square root of the value of the least-squares functional of the approximate solution provides a practical and sharp a posteriori error estimator at no additional cost [21,23,28]. This is quite different from the previous error estimators, see [1,3,4,6,7,10,31] for example.

In this section we shall briefly show that the simple estimator is indeed an exact error estimator in the energy norm, $\|\cdot\|_{\mathscr{B}}$. Thus, it is an equivalent error estimator in the H^1-norm. Let $(\mathbf{u}_h,\boldsymbol{\sigma}_h) = (u_{1h}, u_{2h}, \sigma_{1h}, \sigma_{2h})$ be the least-squares finite element solution. For all subintervals $I_i \in \mathscr{T}_h$, let $\mathscr{Q}_{I_i}(\mathbf{u}_h,\boldsymbol{\sigma}_h;\mathbf{f})$ denote the value of quadratic least-squares functional of the approximate solution restricted on I_i, that is,

$$\begin{aligned}\mathscr{Q}_{I_i}(\mathbf{u}_h,\boldsymbol{\sigma}_h;\mathbf{f}) = {}& \int_{I_i}(-\sigma'_{1h} - f_1)^2\,\mathrm{d}x + \int_{I_i}(-\sigma'_{2h} - \sigma_{1h} - f_2)^2\,\mathrm{d}x \\ &+ \int_{I_i}(-\varepsilon^2\sigma_{1h} + u'_{1h} - u_{2h})^2\,\mathrm{d}x + \int_{I_i}(-\sigma_{2h} + u'_{2h})^2\,\mathrm{d}x\end{aligned}$$

for all $I_i \in \mathscr{T}_h$. Then

$$\begin{aligned}\mathscr{Q}_{I_i}(\mathbf{u}_h, \boldsymbol{\sigma}_h; \mathbf{f}) = & \int_{I_i} (-\sigma'_{1h} - (-\sigma'_1))^2 \, \mathrm{d}x + \int_{I_i} (-\sigma'_{2h} - \sigma_{1h} - (-\sigma'_2 - \sigma_1))^2 \, \mathrm{d}x \\ & + \int_{I_i} (-\varepsilon^2 \sigma_{1h} + u'_{1h} - u_{2h} - (-\varepsilon^2 \sigma_1 + u'_1 - u_2))^2 \, \mathrm{d}x \\ & + \int_{I_i} (-\sigma_{2h} + u'_{2h} - (-\sigma_2 + u'_2))^2 \, \mathrm{d}x \\ = & \|(\mathbf{u}, \boldsymbol{\sigma}) - (\mathbf{u}_h, \boldsymbol{\sigma}_h)\|^2_{\mathscr{B}, I_i},\end{aligned}$$

where $\| \cdot \|_{\mathscr{B}, I_i}$ denotes the energy norm restricted on I_i. Therefore, the computable value $(\mathscr{Q}_{I_i}(\mathbf{u}_h, \boldsymbol{\sigma}_h; \mathbf{f}))^{1/2}$ defines an "exact" error indicator of the subinterval I_i which assesses the quality of the approximate solution $(\mathbf{u}_h, \boldsymbol{\sigma}_h)$ in that subinterval I_i and indicates whether the element needs to be refined, derefined or unchanged. Taking the summation,

$$(\mathscr{Q}(\mathbf{u}_h, \boldsymbol{\sigma}_h; \mathbf{f}))^{1/2} = \left(\sum_{I_i \in \mathscr{T}_h} \mathscr{Q}_{I_i}(\mathbf{u}_h, \boldsymbol{\sigma}_h; \mathbf{f}) \right)^{1/2} = \|(\mathbf{u}, \boldsymbol{\sigma}) - (\mathbf{u}_h, \boldsymbol{\sigma}_h)\|_{\mathscr{B}},$$

we get an exact a posteriori error estimator $(\mathscr{Q}(\mathbf{u}_h, \boldsymbol{\sigma}_h; \mathbf{f}))^{1/2}$, which can serve as one of the major stopping criteria for an entire adaptive process. In general, the effectivity index defined by

$$\Theta = \frac{(\mathscr{Q}(\mathbf{u}_h, \boldsymbol{\sigma}_h; \mathbf{f}))^{1/2}}{\|(\mathbf{u}, \boldsymbol{\sigma}) - (\mathbf{u}_h, \boldsymbol{\sigma}_h)\|_{\mathscr{B}}}$$

is then used to quantify the quality of the estimator and hence the quality of the approximate solution $(\mathbf{u}_h, \boldsymbol{\sigma}_h)$. In our case, Θ exactly equals 1.

References

[1] M. Ainsworth, J.T. Oden, A unified approach to a posteriori error estimation using element residual methods, Numer. Math. 65 (1993) 23–50.

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

[3] I. Babuška, W.C. Rheinboldt, Error estimates for adaptive finite element computations, SIAM J. Numer. Anal. 15 (1978) 736–754.

[4] I. Babuška, W.C. Rheinboldt, A posteriori error analysis of finite element solutions for one-dimensional problems, SIAM J. Numer. Anal. 18 (1981) 565–589.

[5] I. Babuška, M. Suri, On locking and robustness in the finite element method, SIAM J. Numer. Anal. 29 (1992) 1261–1293.

[6] R.E. Bank, R.K. Smith, A posteriori error estimates based on hierarchical bases, SIAM J. Numer. Anal. 30 (1993) 921–935.

[7] R.E. Bank, A. Weiser, Some a posteriori error estimators for elliptic partial differential equations, Math. Comp. 44 (1985) 283–301.

[8] D.M. Bedivan, G.J. Fix, Least squares methods for optimal shape design problems, Comput. Math. Appl. 30 (1995) 17–25.

[9] P.B. Bochev, M.D. Gunzburger, Analysis of least squares finite element methods for the Stokes equations, Math. Comp. 63 (1994) 479–506.
[10] F.A. Bornemann, B. Erdmann, R. Kornhuber, A posteriori error estimates for elliptic problems in two and three space dimensions, SIAM J. Numer. Anal. 33 (1996) 1188–1204.
[11] S.C. Brenner, L.R. Scott, The Mathematical Theory of Finite Element Methods, Springer, New York, 1994.
[12] Z. Cai, T.A. Manteuffel, S.F. McCormick, First-order system least squares for velocity-vorticity-pressure form of the Stokes equations, with application to linear elasticity, Electronic Trans. Numer. Anal. 3 (1995) 150–159.
[13] Z. Cai, T.A. Manteuffel, S.F. McCormick, First-order system least squares for the Stokes equations, with application to linear elasticity, SIAM J. Numer. Anal. 34 (1997) 1727–1741.
[14] Z. Cai, T.A. Manteuffel, S.F. McCormick, S. Parter, First-order system least squares (FOSLS) for planar linear elasticity: pure traction, SIAM J. Numer. Anal. 35 (1998) 320–335.
[15] C.L. Chang, Finite element method for the solution of Maxwell's equations in multiple media, Appl. Math. Comput. 25 (1988) 89–99.
[16] C.L. Chang, B.-N. Jiang, An error analysis of least squares finite element method of velocity-pressure-vorticity formulation for Stokes problem, Comput. Meth. Appl. Mech. Eng. 84 (1990) 247–255.
[17] C.L. Chang, S.-Y. Yang, C.-H. Hsu, A least-squares finite element method for incompressible flow in stress-velocity-pressure version, Comput. Meth. Appl. Mech. Eng. 128 (1995) 1–9.
[18] C.L. Chang, J.J. Nelson, Least-squares finite element method for the Stokes problem with zero residual of mass conservation, SIAM J. Numer. Anal. 34 (1997) 480–489.
[19] X.-L. Cheng, W. Han, H.-C. Huang, Finite element methods for Timoshenko beam, circular arch and Reissner–Mindlin plate problems, J. Comput. Appl. Math. 79 (1997) 215–234.
[20] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.
[21] J.M. Fiard, T.A. Manteuffel, S.F. McCormick, First-order system least squares (FOSLS) for convection–diffusion problems: numerical results, preprint, October 1996.
[22] G.J. Fix, M.E. Rose, A comparative study of finite element and finite difference methods for Cauchy–Riemann type equations, SIAM J. Numer. Anal. 22 (1985) 250–261.
[23] B.-N. Jiang, G.F. Carey, Adaptive refinement for least-squares finite elements with element-by-element conjugate gradient solution, Int. J. Numer. Meth. Eng. 24 (1987) 569–580.
[24] B.-N. Jiang, J. Wu, L.A. Povinelli, The origin of spurious solutions in computational electromagnetics, J. Comput. Phys. 125 (1996) 104–123.
[25] L. Li, Discretization of the Timoshenko beam problem by the p and the $h-p$ versions of the finite element method, Numer. Math. 57 (1990) 413–420.
[26] A.F.D. Loula, T.J.R. Hughes, L.P. Franca, Petrov-Galerkin formulations of the Timoshenko beam problem, Comput. Meth. Appl. Mech. Eng. 63 (1987) 115–132.
[27] A.F.D. Loula, T.J.R. Hughes, L.P. Franca, I. Miranda, Mixed Petrov-Galerkin methods for the Timoshenko beam problem, Comput. Meth. Appl. Mech. Eng. 63 (1987) 133–154.
[28] J.-L. Liu, Exact a posteriori error analysis of the least squares finite element method, Appl. Math. Comput., to appear.
[29] S.-Y. Yang, J.-L. Liu, Least squares finite element methods for the elasticity problem, J. Comput. Appl. Math. 87 (1997) 39–60.
[30] S.-Y. Yang, C.L. Chang, Analysis of a two-stage least squares finite element method for the planar elasticity problem, Math. Meth. Appl. Sci. 22 (1999) 713–732.
[31] O.C. Zienkiewicz, J.Z. Zhu, A simple error estimator and adaptive procedure for practical engineering analysis, Int. J. Numer. Meth. Eng. 24 (1987) 337–357.