Finite element approximation to nonlinear coupled thermal problem

Yanzhen Chang\textsuperscript{a,}, Danping Yang\textsuperscript{b}, Jiang Zhu\textsuperscript{c}

\textsuperscript{a} School of Mathematical Sciences, Shandong University, Jinan, 250100, China
\textsuperscript{b} Department of Mathematics, East China Normal University, Shanghai, 200026, China
\textsuperscript{c} Laboratório Nacional de Computação Científica, MCT, Avenida Getúlio Vargas 333, 25651-075 Petrópolis, RJ, Brazil

\section{Abstract}

A nonlinear coupled elliptic system modelling a large class of engineering problems was discussed in \cite{Loula Zhu 2001, Loula Zhu 2006}. The convergence analysis of iterative finite element approximation to the solution was done under an assumption of ‘small’ solution or source data which guarantees the uniqueness of the nonlinear coupled system. Generally, a nonlinear system may have multiple solutions. In this work, the regularity of the weak solutions is further studied. Then, the nonlinear finite element approximations to the nonsingular solutions are then proposed and analyzed. Finally, the optimal order error estimates in $H^1$-norm and $L^2$-norm as well as in $W^{1,p}$-norm and $L^p$-norm are obtained.

\section{Introduction}

In modelling engineering problems with thermal effects due to Joule or viscous heating, for instance \cite{Saloustros 1995, Saloustros 1997, Gurtin 2000} we come across the following nonlinear coupled model problem:

\begin{equation}
\begin{cases}
(\text{a}) \quad -\nabla \cdot (\mu(\theta) \nabla u) = f, & \text{in } \Omega, \\
(\text{b}) \quad -\Delta \theta = \mu(\theta)|\nabla u|^2, & \text{in } \Omega, \\
(\text{c}) \quad u = 0, & \text{on } \Gamma^r, \\
(\text{d}) \quad \theta = 0, & \text{on } \Gamma^s,
\end{cases}
\end{equation}

where $\Omega$ is a regular bounded open subset of $\mathbb{R}^n (1 \leq n \leq 3)$; $\Gamma^r$ is its boundary; $u : \Omega \to R$ is the potential; $\theta : \Omega \to R$ is the temperature; $\mu \in C(R)$ is the coupling function satisfying $0 < K_1 \leq \mu(s) \leq K_2 < \infty$, $\forall s \in R$, $f$ is a source datum.

This type of problem has received special attention recently in practical engineering application and mathematical theory. Mathematical analysis of the problem can be found in, for example, \cite{Gurtin 2000, Lin 2006, Loula Zhu 2001, Loula Zhu 2006}. The time-dependent problem can also be found such as in \cite{Lefebvre 2000}. In \cite{Loula Zhu 2006}, the authors studied the problem (1.1), for ‘small’ $f$, which guarantees the uniqueness of the solution, and gave a convergence analysis of iterative finite element approximation to the solution. However, without any specific assumption, the solution is generally not unique. For the nonlinear problems of this kind, the most difficult part is how to find or approximate multiple solutions including both singular and nonsingular forms. The development of effective numerical schemes in capturing such a structure including both singular and nonsingular solutions plays a critical role in these kinds of problems. We are interested in the approximation accuracy of nonsingular solutions. Numerical analysis of
such a stationary problem presents additional difficulties concerning convergence and error estimate with the presence of Joule or viscous heating. Brezzi, Rappaz and Raviart [5] developed a general theory for the finite-dimensional approximations of nonsingular solutions of nonlinear problems. However, it seems that BRR theory cannot be applied directly to a strongly nonlinear problem (1.1).

The purpose of this article is to establish a framework for convergence analysis of finite element approximations to this kind of strongly nonlinear problem in general cases. By using $L^p$ theory, we give a convergence analysis of the standard finite element approximations in a general case where $f$ is assumed only to be bounded and is not necessarily small, and obtain optimal a priori error estimates in both $W^{1,p}$-norm and $L^p$-norm.

In what follows, we firstly establish existence and regularity of the weak solution of problem (1.1) in Section 2. Then, we propose its finite element approximations in Section 3. In Section 4, we give a convergence analysis and show that the nonlinear finite element solutions strongly converge to different solutions of the system. In Sections 5 and 6, we obtain optimal order error estimates for the finite element approximations to the nonsingular solutions.

2. Existence and regularity

Let $W^{m,s}(\Omega)$ denote the Sobolev space with its norm $\| \cdot \| = \| \cdot \|_{W^{m,s}}$ for $m \geq 0$ and $1 \leq s \leq \infty$. We write $H^m(\Omega) = W^{m,2}(\Omega)$ when $s = 2$, with the norm $\| \cdot \|_H$, and $L^s(\Omega) = W^{0,s}(\Omega)$ when $m = 0$ with the norm $\| \cdot \|_{L^s}$. $W^{m,1}(\Omega)$ is the closure of the space $C^0(\Omega)$ for the norm $W^{m,1}(\Omega)$.

**Definition 1.** We denote by $\mathcal{R}_s$ for $1 < s < \infty$ the class of regular subsets $G$ in $\mathbb{R}^d$ for which the Laplacian operator maps $W^{1,s}_0(G)$ onto $W^{-1,s}(G)$.

**Remark 1.** A bounded $C^1$ domain, for example, is of class $\mathcal{R}_s$ for every $s \in (2, \infty)$, see Theorem 4.6 in [17].

From now on, we assume that $\Omega$ is of class $\mathcal{R}_p$ for some $\hat{p} > 2$. For $1 \leq s \leq \hat{p}$, we define $M_s \geq 1$ by

$$
\inf_{u \in W^{1,s}_0(\Omega)} \sup_{v \in L^\infty(\Omega)} \frac{d(u, v)}{\| \nabla u \| \| \nabla v \|_{L^\infty}} = \frac{1}{M_s},
$$

where

$$
d(u, v) = (\nabla u, \nabla v),
$$

(\cdot, \cdot) denotes the duality between $L^s(\Omega)^n$ and $L^s(\Omega)^n$, $s'$ is the dual number of $s$ such that $1/s + 1/s' = 1$. It is easy to see that $M_2 = 1$ and $M_\infty = M_1$.

**Lemma 2** (cf. [20]). If $p \in (2, \hat{p})$ is such that

$$
\frac{K_2 - K_1}{K_1 + K_2} < 1,
$$

then, for any $\theta$, we have

$$
\inf_{u \in W^{1,p}_0(\Omega)} \sup_{v \in L^{p'}(\Omega)} \frac{a(\theta; u, v)}{\| \nabla u \|_{L^p} \| \nabla v \|_{L^{p'}}} \geq \gamma
$$

where

$$
a(\theta; u, v) = (\mu(\theta) \nabla u, \nabla v),
$$

$$
\gamma = \frac{K_1 + K_2}{2M_p} \left( 1 - \frac{K_2 - K_1}{K_1 + K_2} \right) > 0.
$$

Similarly to [15,20], we have:

**Lemma 3.** Let $f \in W^{-1,p}(\Omega)$, where $p$ is defined in Lemma 2. Then, for any given $\theta$, there exists a unique $u \in W^{1,p}_0(\Omega)$ such that

$$
a(\theta; u, v) = (f, v), \quad \forall \ v \in W^{1,p}_0(\Omega).
$$

Furthermore, there exists a constant $C > 0$ dependent on $\Omega, K_1, K_2$ and $p$ such that

$$
\| \nabla u \|_{L^p} \leq C \| f \|_{W^{-1,p}}.
$$
**Theorem 4** (Existence). Let $p$ be defined in Lemma 2 and
\[ q = \begin{cases} dp/(2n - p), & \text{if } p < 2n, \\ \text{any number in } (2, \infty), & \text{if } p \geq 2n, \end{cases} \tag{2.9} \]
then the variational form of problem (1.1)
\[
\begin{aligned}
\text{Find } (u, \theta) \in W^{1,p}_0(\Omega) \times W^{1,q}_0(\Omega) \text{ such that} \\
(a) & \ a(\theta; u, v) = (f, v), \\
(b) & \ d(\theta, \eta) = (\mu(\theta)|\nabla u|^2, \eta), \\
& \forall \ v \in W^{1,p}_0(\Omega), \\
& \forall \ \eta \in W^{1,q}_0(\Omega)
\end{aligned}
\tag{2.10}
\]
has at least one solution. Furthermore, there exists a constant $C > 0$ only dependent on $\Omega, K_1, K_2$ and $p$ such that
\[
\|\nabla u\|_{L^p} \leq C\|f\|_{W^{-1,p}},
\tag{2.11}
\]
\[
\|\nabla \theta\|_{L^q} \leq C\|f\|_{W^{-1,p}}^2.
\tag{2.12}
\]
**Proof.** Existence of solutions and (2.11) are the direct consequence of Theorem 10 in Section 4 and Lemma 3, respectively. Since
\[
(\mu(\theta)|\nabla u|^2, \eta) \leq C\|\nabla u\|_{L^p}^2 \|\eta\|_{L^{p/(p-2)}} \leq C\|\nabla u\|_{L^p}^2 \|\nabla \eta\|_{L^{np/(np-2n+p)}}
\]
and noticing that $q$ is the dual number of $np/(np-2n+p)$, (2.10)(b) and (2.1), we have
\[
\|\nabla \theta\|_{L^q} \leq C\|\nabla u\|_{L^p}^2.
\tag{2.13}
\]
That is (2.12), then the proof is completed. \Box

**Remark 5.** By the Maximum principle (see Theorem 8.16 in [10]) and similar arguments in [14], we can get existence of problem (1.1) for any $f \in L^s(\Omega)$ ($s > n/2$) without those assumptions in Theorem 4. However, the above framework is generally necessary for further study such as uniqueness and regularity.

Next, we have better results for regularity of the solutions.

**Theorem 6** (Regularity). Let $\mu'(\theta)$ be Lipschitz continuous, and $f \in L^{2+s}(\Omega)$ for $s \geq 0$. If $p$, defined in Lemma 2, satisfies $p > n$, and furthermore we assume that $\Omega$ is of class $C^{1,1}$, then $(u, \theta) \in W^{2,2+s}(\Omega)$ and there exists a defined polynomial $P_m(x)$, whose coefficients are dependent only on $\Omega, L, K_1, K_2, p, S$, Sobolev constants and $m$ (an integer that will be determined later on), such that
\[
\|u\|_{W^{2,2+s}} \leq P_m(\|f\|_{L^{2+s}}),
\tag{2.14}
\]
\[
\|\theta\|_{W^{2,2+s}} \leq C P_m(\|f\|_{L^{2+s}})^2,
\tag{2.15}
\]
where constant $C$ is defined by (2.13). As a consequence, there holds
\[
\|\nabla u\|_{L^\infty} \leq P_m(\|f\|_{L^{2+s}}), \quad \text{if } s > n - 2.
\tag{2.16}
\]
**Proof.** Firstly, we have
\[
- \Delta u = \frac{f}{\mu(\theta)} + \frac{\mu'(\theta)}{\mu(\theta)} \nabla \theta \cdot \nabla u.
\tag{2.17}
\]
For each $r \geq p$, let $t = nr/(2n - r)$. It follows from Theorem 4 that $\theta \in W^{1,t}(\Omega)$ if $u \in W^{1,r}(\Omega)$. Now assume that $u \in W^{1,r}(\Omega)$ and $\theta \in W^{1,t}(\Omega)$, we have
\[
\left( \frac{\mu'(\theta)}{\mu(\theta)} \nabla \theta \cdot \nabla u, v \right) \leq L/K_1\|\nabla \theta\|_{L^1}\|\nabla u\|_{L^{t'}}\|v\|_{L^{t'^*}},
\tag{2.18}
\]
where $1/r + 1/t + 1/k' = 1$, so
\[
k' = \frac{rt}{rt - r - t}.
\tag{2.19}
\]
Thus, from the discussion of regularity in [12,13],
\[
u \in W^{2,\min(2+s,\kappa)}(\Omega), \quad \kappa = \frac{k'}{k' - 1} = \frac{nr}{3n - r}
\tag{2.20}
\]

and
\[ \|u\|_{W^{2,\min}(\Omega)} \leq C(\|f\|_{L^2} + \|f\|_{W^{-1,r}}^3) \leq C(\|f\|_{L^2} + \|f\|_{L^2}^3). \] (2.21)

The Sobolev inequality implies that
\[ \|\nabla u\|_{L^\infty} \leq C\|u\|_{W^{2,\min}(\Omega)} \leq C(\|f\|_{L^2} + \|f\|_{L^2}^3), \] (2.22)
where
\[ r_{nw} = \frac{n\kappa}{n - \kappa} = \frac{nr}{3n - 2}. \] (2.23)

Since \( \sigma = \frac{n}{3n - 2} \geq \frac{n}{3n - 2p} > 1 \) for \( p > n \), \( r_{nw} = \sigma r \) and \( t_{nw} = \frac{m_{nw}}{2n - r_{nw}} > \sigma t \), and noticing that \( \sigma \) is monotonic increasing respect to \( r \), then repeating the above procedure finite \( m \) times, we can get the a priori estimate (2.14). This ends the proof. \( \square \)

3. Finite element scheme

Let \( k^h \) be a quasi-uniform triangulation of the domain \( \Omega \) and nodes \( x_1, \ldots, x_m \). We suppose that \( \tilde{\Omega} \) is the union of the elements of \( k^h \) so that element edges lying on the boundary may be curved. Introduce
\[ s^h = \{ v_h \in C(\tilde{\Omega}) : v_h|_{k^h} \in P_i(K^h), r \geq 1 \}. \]
Let us define the space of finite elements \( \mathcal{V}^h \subset s^h \) with the Lagrange basis \( \{ v_i : i = 1, \ldots, m \} \) and appropriate modification for boundary elements.

A standard finite element approximation to (2.10) is thus defined by:
\[
\begin{align*}
\text{Find } (u^h, \phi^h) & \in \mathcal{V}^h \times \mathcal{V}^h \text{ such that } \\
(a) & \quad a(\phi^h; u^h, v^h) = (f, v^h), \quad \forall v^h \in \mathcal{V}^h, \\
(b) & \quad d(\phi^h, \mu^h) = (\mu(\phi^h)\|\nabla u^h\|^2, \eta^h), \quad \forall \eta^h \in \mathcal{V}^h.
\end{align*}
\] (3.1)

**Theorem 7.** The finite element scheme (3.1) has at least one solution.

To prove this theorem, we give some basic properties of the finite element system.

**Lemma 8** (cf. [18]). For \( p \) defined in Lemma 2, and any \( 0 < \epsilon \ll 1 \), there exist \( h_0 > 0 \) and \( p_\epsilon \in (2, p) \) such that, for any \( h \in (0, h_0) \) and any \( p \in (2, p_\epsilon) \)
\[
\inf_{u^h \in \mathcal{V}^h, p \in \mathcal{V}^h} \sup_{v^h \in \mathcal{V}^h, p \in \mathcal{V}^h} \frac{d(u^h, v^h)}{\|\nabla u^h\|_{L^p} \|\nabla v^h\|_{L^p}} \geq \frac{1}{M_{p_\epsilon}(1 + \epsilon)}. \] (3.2)

**Lemma 9.** For any given \( \theta^h \in \mathcal{V}^h \), if \( (u^h, \phi^h) \) satisfies (3.1), then there exist \( h_0 > 0 \) and \( p^* \in (2, p) \) (\( p \) is defined in Lemma 2), such that, for any \( h \in (0, h_0) \)
\[ \|\nabla u^h\|_{L^{q^*}} \leq C\|f\|_{W^{-1, p^*}}, \] (3.3)
\[ \|\nabla \phi^h\|_{L^{q^*}} \leq C\|f\|^2_{W^{-1, p^*}}, \] (3.4)
where
\[ q^* = \begin{cases} dp^*/(2n - p^*), & \text{if } p^* < 2n, \\ \text{any number in } (2, \infty), & \text{if } p^* \geq 2n. \end{cases} \] (3.5)

**Proof.** For sufficiently small \( \epsilon \), there exists \( p^\epsilon \in (2, p_\epsilon) \) such that
\[ M_{p^\epsilon}(1 + \epsilon) \frac{K_2 - K_1}{K_1 + K_2} < 1. \] (3.6)
Thus, by Lemma 8, we have, for any \( \theta^h \),
\[
a(\theta^h; u^h, v^h) = (\mu(\theta^h)\nabla u^h, \nabla v^h) = \frac{K_1 + K_2}{2} (\nabla u^h, \nabla v^h) + \left( \mu(\theta^h) - \frac{K_1 + K_2}{2} \right) (\nabla u^h, \nabla v^h) \geq \frac{K_1 + K_2}{2M_{p^*}(1 + \epsilon)} \left( 1 - M_{p^*}(1 + \epsilon) \frac{K_2 - K_1}{K_2 + K_1} \right) \|\nabla u^h\|_{L^{q^*}} \|\nabla v^h\|_{L^{q^*}},
\]
The proof is divided into four steps. Firstly, we show that there exist subsequences of weak convergence in the system of the finite element procedure. We cannot identify the limits. On the other hand, in the next section we give the error estimate for a subsequence which converges strongly to solutions of the nonlinear system. However, since in general there is no uniqueness we cannot identify the limits. On the other hand, in the next section we give the error estimate for a subsequence which converges strongly to a nonsingular solution.

4. Convergence of finite element solutions

Using compacting and consistency arguments we prove that the family of finite element approximations has subsequences which converge strongly to solutions of the nonlinear system. However, since in general there is no uniqueness we cannot identify the limits. On the other hand, in the next section we give the error estimate for a subsequence which converges strongly to a nonsingular solution.

Theorem 10. Assume that \( \mu(s) \) is uniformly continuous, and \( p^* \) and \( q^* \) are defined as in Lemma 9. The solution sequence \((u^h, \theta^h)\) of the finite element procedure (3.1) can be divided into several subsequences which converge strongly to different solutions \((u, \theta)\) of system (1.1) in \( W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega) \).

Proof. The proof is divided into four steps. Firstly, we show that there exist subsequences of weak convergence in \( W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega) \). Secondly, we prove that the limit of each weakly convergent subsequence is a solution of system (3.1).
Thirdly, we prove that for each weakly convergent subsequence, there exists one sub-subsequence of strong convergence in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$. Fourthly, we prove that each weakly convergent subsequence also is strongly convergent.

The first step: By Lemma 9, we know that the sequence $(u^h, \theta^h)$ is bounded in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$. So, we can choose a subsequence $(u^h_i, \theta^h_i)$ from $(u^h, \theta^h)$ such that it is weakly convergent in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$ and strongly convergent in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$ for any $0 \leq s < 1$. Let $\lim_{n \to \infty} (u^h_n, \theta^h_n) = (\bar{u}, \bar{\theta})$ weakly in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$.

The second step: Let $(u^h, \theta^h)$ be a weakly convergent sequence in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$. Furthermore, since $(u^h, \theta^h)$ is a strongly convergent sequence in $L^p(\Omega) \times L^q(\Omega)$, one can choose a subsequence $(u^h_i, \theta^h_i)$ from $(u^h, \theta^h)$ such that it is almost uniformly convergent in $\Omega$, i.e., for each $\delta > 0$, there exists a sub-domain $\Omega_\delta \subset \Omega$ such that $\text{meas}(\Omega \setminus \Omega_\delta) \leq \delta$ and $(u^h_i, \theta^h_i)$ converges uniformly to $(\bar{u}, \bar{\theta})$ in $\Omega_\delta$.

We show that $(\bar{u}, \bar{\theta})$ is a solution of system (1.1). For each $v \in C_0^\infty(\Omega)$, we have

$$a(\bar{\theta}; \bar{u}, v) - (f, v) = [a(\bar{\theta}; \bar{u}, v) - a(\bar{\theta}; \bar{u}, h)] + [a(\bar{\theta}; \bar{u}, v) - a(h; \bar{u}, v)] + [a(\bar{\theta}; u^h, v - h; \bar{u}, v)] + \inf_{v \in v^h} \left[ a(\theta^h; u^h, v - \bar{h}; \bar{u}, v) - (f, v - \bar{h}) \right].$$

(4.1)

By the strong convergence in $W^{1,p^*}$ for any $0 \leq s < 1$ and weak convergence in $W^{1,p^*}$, we see

$$[a(\bar{\theta}; \bar{u}, v) - a(\bar{\theta}; u^h, v)] + [a(\bar{\theta}; u^h - \bar{u}, v) - a(h; \bar{u}, v)] \to 0, \quad \text{as } i \to \infty. $$

(4.2)

By approximation properties of finite element spaces, it follows

$$\inf_{v \in v^h} \left[ a(\theta^h; u^h, v - \bar{h}) - (f, v - \bar{h}) \right] \to 0, \quad \text{as } i \to \infty. $$

(4.3)

For each $0 < \epsilon \ll 1$, by continuity of Lebesgue integration, we know that there exists $\delta > 0$ such that, for any $\tilde{g}$ with $\text{meas}(\tilde{g}) \leq \delta$,

$$\|\nabla \bar{u}\|_{L^{p^*}(\tilde{g})} \leq \epsilon. $$

(4.4)

By almost uniform convergence of $\theta^h$, we get that

$$a(\bar{\theta}; \bar{u}, v) - a(\theta^h; \bar{u}, v) = ((\mu(\bar{\theta}) - \mu(\theta^h)))(\nabla \bar{u}, \nabla v)$$

$$\leq \max_{\theta^h} |\mu(\bar{\theta}) - \mu(\theta^h)| \|\nabla \bar{u}\|_{L^{p^*}(\Omega)} \|\nabla v\|_{L^{p^*}(\Omega)}$$

$$\leq C \|\nabla \bar{u}\|_{L^{p^*}(\Omega)}, \quad \text{as } i \to \infty. $$

(4.5)

Summarizing the above results, we obtain

$$|a(\bar{\theta}; \bar{u}, v) - (f, v)| \leq C \|\nabla \bar{u}\|_{L^{p^*}(\Omega)}$$

(4.6)

for any $\epsilon > 0$. This means that

$$a(\bar{\theta}; \bar{u}, v) = (f, v), \quad \forall v \in W^{1,p^*}(\Omega). $$

(4.7)

Similarly, let

$$b(u; \theta, \eta) = (\mu(\theta)|\nabla u|^2, \eta), $$

and we see

$$d(\bar{\theta}, \eta) - b(u; \bar{\theta}, \eta) = d(\bar{\theta}, \eta) - d(\theta^h, \eta) - b(\bar{u}; \bar{\theta}, \eta) + b(\bar{u}; \theta^h, \eta)$$

$$- b(\bar{u}^h; \theta^h, \eta) + b(u^h; \theta^h, \eta) + \inf_{\eta \in \eta^h} \left[ d(\theta^h, \eta - \eta^h) - b(u^h; \theta^h, \eta - \eta^h) \right], $$

(4.8)

and

$$b(\bar{u}; \theta^h, \eta) - b(u^h; \theta^h, \eta) = (\mu(\theta^h)(|\nabla \bar{u}|^2 - |\nabla u^h|^2), \eta)$$

$$= (\mu(\theta^h)|\nabla (\bar{u} - u^h)| \cdot \nabla (\bar{u} + u^h), \eta). $$

(4.9)

We can easily see that the r.h.s of (4.8) tends to zero, then the r.h.s of (4.7) tends to zero. This also means that

$$d(\bar{\theta}, \eta) = b(u; \bar{\theta}, \eta), \quad \forall \eta \in W^{1,q^*}(\Omega). $$

(4.10)

$(\bar{u}, \bar{\theta})$ is thus a solution of system (1.1).
The third step: We will prove that such a weakly convergent sequence also is strongly convergent in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$. Since

$$a(\theta^h; u^h, v^h) = a(\theta^h; u^h - \bar{u}, v^h) + a(\theta^h; \bar{u} - l^h\bar{u}, v^h)$$

$$= ((\mu(\bar{\theta}) - \mu(\theta^h)) \nabla \bar{u}, \nabla v^h) + a(\theta^h; \bar{u} - l^h\bar{u}, v^h), \quad \forall v^h \in \mathcal{V}^h,$$

where $l^h : W^{1,p^*}(\Omega) \rightarrow \mathcal{V}^h$ is the interpolation operator. From (3.7), we have

$$\|\nabla (u^h - \bar{u})\|_{L^{p^*}} \leq \|\nabla (u^h - l^h\bar{u})\|_{L^{p^*}} + \|\nabla (\bar{u} - l^h\bar{u})\|_{L^{p^*}}$$

$$\leq \gamma^{-1}_s \sup_{v^h \in \mathcal{V}^h} a(\theta^h; u^h - l^h\bar{u}, v^h) + \|\nabla (\bar{u} - l^h\bar{u})\|_{L^{p^*}}$$

$$\leq \gamma^{-1}_s \sup_{v^h \in \mathcal{V}^h} (\mu(\theta)) \nabla \bar{u}, \nabla v^h + (1 + K_2)\|\nabla (\bar{u} - l^h\bar{u})\|_{L^{p^*}}. \quad (4.11)$$

Similarly to (4.4) and by the approximation property of the interpolate $l^h$, we can see that

$$\lim_{i \rightarrow \infty} \|\nabla (u^h - \bar{u})\|_{L^{p^*}} = 0. \quad (4.12)$$

Similarly, we can prove that $\theta^h$ is strongly convergent to $\bar{\theta}$ in $W^{1,q^*}(\Omega)$.

The fourth step: In the general case, let $(u^h, \theta^h) \rightarrow (u, \theta)$ weakly in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$ as $i \rightarrow \infty$. We can prove that $(u^h, \theta^h)$ is also strongly convergent in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$. If it is not true then there exist a constant $\epsilon_0 > 0$ and its subsequence $(u^{i_n}, \theta^{i_n})$ such that, for $i = 1, 2, \ldots, \infty$,

$$\|\nabla (u^{i_n} - u)\|_{L^{p^*}(\Omega)} + \|\nabla (\theta^{i_n} - \theta)\|_{L^{q^*}(\Omega)} \geq \epsilon_0. \quad (5.13)$$

On the other hand, since $(u^h, \theta^h)$ is bounded in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$, by the last step, we can find a subsequence of $(u^{i_n}, \theta^{i_n})$ which is strongly convergent in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$. This is contradictory to (4.13). Thus we complete the proof. \(\square\)

5. Error estimate in $L^{p^*} \times L^{q^*}$-norm

In this and the following sections, we will give the error estimates of finite element approximations. Generally, only strong convergence can be obtained (see Theorem 10). However, in the case of nonsingular solutions, we can obtain error estimates of optimal order. Similarly to the definition of nonsingular solutions of nonlinear problems in [5, 19], we say that the solution $(u, \theta)$ of the nonlinear system (1.1) is nonsingular if the following linear adjoint system

$$\begin{align*}
-\nabla \cdot (\mu(\theta) \nabla w) + 2\nabla \cdot (\mu(\theta) \nabla u) &= \varphi, \\
-\Delta T - T \mu(\theta) |\nabla u|^2 + \mu(\theta) \nabla u \cdot \nabla w &= \varphi, \\
w|_{\Gamma} &= 0, \\
T|_{\Gamma} &= 0,
\end{align*} \quad (5.1)$$

has a unique solution for any given $\varphi$ and $\varphi$. By the global theory on linear Banach operators (cf. e.g. Brezzi [4], Babuška [1] and Nečas [16]), we can see that the unique solution $(w, T)$ of (5.1) should satisfy the a priori estimate

$$\|w\|_{W^{1,p^*}(\Omega)} + \|T\|_{W^{1,q^*}(\Omega)} \leq C(\|\varphi\|_{W^{-1,p^*}(\Omega)} + \|\varphi\|_{W^{-1,q^*}(\Omega)}). \quad (5.2)$$

In other words, the first order variation system of nonlinear system (1.1) at the nonsingular system is well posed and stable.

**Theorem 11.** Let $\mu(\theta)$ be Lipschitz continuous and $\Omega$ be of the class $C^{1,1}$, and $(u^h, \theta^h)$ be a solution sequence of the nonlinear finite element system (3.1) strongly convergent to a nonsingular solution $(u, \theta)$ of the nonlinear system (1.1) in $W^{1,p^*}(\Omega) \times W^{1,q^*}(\Omega)$ as $h \rightarrow 0$. Assume that $u \in W^{r+1,p^*}$, $\theta \in W^{r+1,q^*}$ ($r \geq 1$), then the a priori error estimate holds

$$\|u - u^h\|_{L^{p^*}} + \|\theta - \theta^h\|_{L^{q^*}} + h(\|u - u^h\|_{W^{1,p^*}} + \|\theta - \theta^h\|_{W^{1,q^*}}) \leq Ch^{r+1}, \quad (5.3)$$

where the constant $C$ is independent of $h$.

**Proof.** The proof is divided into two steps. Firstly, we prove error estimate in $W^{1,p^*} \times W^{1,q^*}$-norm. Then we prove error estimate in $L^{p^*} \times L^{q^*}$-norm. Here we need to point out that the following proof is given formally when $p^* < 3$ and $n = 3$; but when $p^* \geq 3(n - 3)$ or $n = 2$ it is easier to give the proof similarly and in order to avoid repetition we will not give the details again.
The first step: It follows from (1.1) and (3.1) that

(a) \( a(\theta^h; u^h - u, v^h) = a(\theta; u, v^h) - a(\theta^h; u, v^h), \quad \forall \ u^h \in \mathcal{V}^h, \)

(b) \( d(\theta^h - \theta, \eta^h) = (|\nabla u|^2 (\mu(\theta^h) - \mu(\theta)), \eta^h) + (\mu(\theta^h)(|\nabla u|^2 - |\nabla v|^2), \eta^h), \quad \forall \ \eta^h \in \mathcal{V}^h. \) (5.4)

By (3.7),

\[
\|u - u^h\|_{W^{1,p^*}} \leq \inf_{\eta^h \neq 0 \in \mathcal{V}^h} (\|u - u^h\|_{W^{1,p^*}} + \|u - u^h\|_{W^{1,p^*}})
\]

\[
\leq \gamma^1_s \sup_{\eta^h \neq 0 \in \mathcal{V}^h} a(\theta^h; u - u^h, \eta^h) + (1 + K_2) \inf_{\eta^h \neq 0 \in \mathcal{V}^h} \|u - u^h\|_{W^{1,p^*}}
\]

\[
\leq \gamma^1_s \|\theta - \theta^h\|_{L^2} \|\nabla u\|_{L^{2,p^*}} + (1 + K_2) \inf_{\eta^h \neq 0 \in \mathcal{V}^h} \|u - u^h\|_{W^{1,p^*}}
\]

\[
\leq C(\|\theta - \theta^h\|_{L^2} \|\nabla u\|_{L^{2,p^*}} + h^r \|u\|_{W^{r+1,p^*}})
\]

\[
\leq C(\|\theta - \theta^h\|_{L^2} \|u\|_{W^{2,p^*}(\Omega)} + h^r \|u\|_{W^{r+1,p^*}}) \quad \text{(5.5)}
\]

and similarly,

\[
\|\theta - \theta^h\|_{W^{1,q^*}} \leq C \left( \sup_{\eta^h \neq 0 \in \mathcal{V}^h} \frac{d(\theta - \theta^h, \eta^h)}{\|\nabla \eta^h\|_{L^{1,q^*}}} + \inf_{\eta^h \neq 0 \in \mathcal{V}^h} \|\theta - \eta^h\|_{W^{1,q^*}} \right)
\]

\[
\leq C(\|\theta - \theta^h\|_{L^2} \|\nabla u\|_{L^{2,p^*}} + h^r \|u\|_{W^{r+1,p^*}} + \inf_{\eta^h \neq 0 \in \mathcal{V}^h} \|\theta - \eta^h\|_{W^{1,q^*}})
\]

\[
\leq C(\|\theta - \theta^h\|_{L^2} \|u\|_{W^{2,p^*}(\Omega)} + h^r \|\theta\|_{W^{r+1,p^*}}). \quad \text{(5.6)}
\]

We have

\[
\|u - u^h\|_{W^{1,p^*}} + \|\theta - \theta^h\|_{W^{1,q^*}} \leq C(\|\theta - \theta^h\|_{L^2} + h^r (\|u\|_{W^{r+1,p^*}} + \|\theta\|_{W^{r+1,q^*}})). \quad \text{(5.7)}
\]

We need the duality technique to get the estimate of \(\|\theta - \theta^h\|_{L^2}.\) Introduce the auxiliary function \((w, T)\) such that

\[
\begin{align*}
-\nabla \cdot (\mu(\theta)\nabla w) + 2\nabla \cdot (\mu(\theta)T \nabla u) &= \|u^h - u\|^{p^*-1}\text{sign}(u^h - u), \\
-\Delta T + T \mu(\theta)|\nabla u|^2 + \mu(\theta)\nabla u \cdot \nabla w &= |\theta^h - \theta|^2 \text{sign}(\theta^h - \theta), \\
w |_{T^*} &= 0, \\
T |_{T^*} &= 0.
\end{align*}
\]

\[
\|w\|_{W^{2,p^*}} + \|T\|_{W^{2,\frac{2}{2}}} \leq C(\|u - u^h\|_{p^*}^{p^*-1} + \|\theta - \theta^h\|_{L^2}^2). \quad \text{(5.9)}
\]

Using the standard argument, we have

\[
\|u - u^h\|_{p^*} + \|\theta - \theta^h\|_{L^2} = a(\theta; u^h - u, w) + (\mu(\theta)(\theta^h - \theta)\nabla u, \nabla w) + d(\theta^h - \theta, T)
\]

\[
+ ((\mu(\theta) - \mu(\theta^h))|\nabla u|^2, T) + (\mu(\theta)(|\nabla u|^2 - |\nabla u^h|^2), T)
\]

\[
- (\mu(\theta)(\theta^h - \theta))|\nabla u|^2, T) - (\mu(\theta) - \mu(\theta^h))|\nabla u|^2, T)
\]

\[
+ 2(\mu(\theta)|\nabla u - \nabla u^h|, T) - (\mu(\theta)(|\nabla u|^2 - |\nabla u^h|^2), T)
\]

\[
= a(\theta; u^h, w) - a(\theta^h; u^h, w) + (\mu(\theta)(\theta^h - \theta)\nabla u, \nabla w)
\]

\[
+ a(\theta^h; u^h, w) - a(\theta; u, w) + d(\theta^h - \theta, T)
\]

\[
+ ((\mu(\theta) - \mu(\theta^h))|\nabla u|^2, T) + (\mu(\theta)(|\nabla u|^2 - |\nabla u^h|^2), T)
\]

\[
+ ((\mu(\theta) - \mu(\theta^h))|\nabla u|^2, T) - (\mu(\theta^h)(\theta^h - \theta)\nabla u^h|^2, T)
\]

\[
+ 2(\mu(\theta)|\nabla u - \nabla u^h|, T) - (\mu(\theta)(|\nabla u|^2 - |\nabla u^h|^2), T). \quad \text{(5.10)}
\]

We estimate terms on the r.h.s of (5.10) one by one. It is clear that

\[
a(\theta; u^h, w) - a(\theta^h; u^h, w) + (\mu(\theta)(\theta^h - \theta)\nabla u, \nabla w)
\]

\[
= ((\mu(\theta) - \mu(\theta^h) - \mu(\theta)(\theta^h - \theta))\nabla u, \nabla w) + ((\mu(\theta) - \mu(\theta^h))\nabla (u^h - u), \nabla w)
\]
\[
\begin{align*}
\leq C \left( \|\theta - \theta^h\|^2_{\mathcal{L}^{p_0}} \|\nabla u\|_{\mathcal{L}^{p_0}} \|w\|_{\mathcal{L}^{p_0}} \|
abla u\|_{W^{1,p_0}} + \|\theta - \theta^h\|^2_{L^{p_0}} \|u - u_h\|_{W^{1,p_0}} \Big\|w\Big\|_{W^{1,p_0}} \right) \\
\leq C(\|\theta - \theta^h\|_{W^{1,q'}} + \|u - u_h\|_{W^{1,p_0}}) \|w\|_{W^{2,p_r}}, \\
\end{align*}
\]
(5.11)

\[
a(\theta^h; u^h, w) - a(\theta; u, w) + d(\theta^h - \theta, T) + ((\mu(\theta) - \mu(\theta^h))\|\nabla u\|^2, T) + (\mu(\theta)(\|\nabla u\|^2 - |\nabla u|^2), T) \\
= \inf_{v^h \in \mathcal{V}} (a(\theta^h; u^h, w - v^h) - a(\theta; u, w - v^h)) + \inf_{v^h \in \mathcal{V}} (d(\theta^h - \theta, T - h^h) + ((\mu(\theta) - \mu(\theta^h))\|\nabla u\|^2, T - h^2) \\
+ (\mu(\theta^h)(\|\nabla u\|^2 - |\nabla u|^2), T - h^2)) \leq C((\|\theta - \theta^h\|_{L^2} + \|\nabla u - u_h\|_{L^p}) \inf_{v^h \in \mathcal{V}} \|w - v^h\|_{W^{1,p'}}) \\
+ \|\theta - \theta^h\|_{W^{1,q'}} \inf_{v^h \in \mathcal{V}} \|T - h^2\|_{L^p} + \|\theta - \theta^h\|_{L^2} \inf_{v^h \in \mathcal{V}} \|T - h^2\|_{L^{p'/q'}} \\
+ (\|u^h\|_{L^2} + \|\nabla u\|_{L^p})\|\nabla (u - u_h)\|_{L^{p'}} \inf_{v^h \in \mathcal{V}} \|T - h^2\|_{L^{p'/q'R - 2})} \\
\leq C(\|\theta - \theta^h\|^2_{W^{1,q'}} + \|u - u_h\|^2_{W^{1,p_0}}) \|w\|_{W^{2,p_r}}. \\
\end{align*}
\]
(5.12)

where \(0 < s < 1\).

Let us estimate the last terms:

\[
\begin{align*}
(\mu(\theta^h) - \mu(\theta))\|\nabla u\|^2 + \mu(\theta^h)\|\nabla u\|^2, T \\
&\leq C\|\theta - \theta^h\|^2_{\mathcal{L}^{p_0}} \|\nabla u\|^2 + T \|\theta^h\|_{\mathcal{L}^{p_0}} \|\nabla u\|^2 \\
&\leq C\|\theta - \theta^h\|^2_{W^{1,q'}} \|T\|_{W^{2,1/2}}, \\
\end{align*}
\]
(5.13)

\[
\begin{align*}
2(\mu(\theta)\nabla u \cdot \nabla u, T) - (\mu(\theta)(\|\nabla u\|^2 - |\nabla u|^2), T) \\
= 2(\mu(\theta)\nabla u \cdot (\nabla u - u_h), |\nabla u|, T) - (\mu(\theta)(\nabla u - u_h)(\nabla u - u_h), T) \\
&\leq C((\|\theta - \theta^h\|^2_{W^{1,q'}} + \|u - u_h\|^2_{W^{1,p_0}}) \|T\|_{W^{2,1/2}}^2. \\
\end{align*}
\]
(5.14)

Furthermore, we can have

\[
\begin{align*}
\|u - u_h\|_{L^{p_0}} + \|\theta - \theta^h\|_{L^2} \leq C(h^2((\|\theta - \theta^h\|_{W^{1,q'}} + \|u - u_h\|_{W^{1,p_0}}) + (\|\theta - \theta^h\|^2_{W^{1,q'}} + \|u - u_h\|^2_{W^{1,p_0}})), \\
\|u - u_h\|_{W^{1,p_0}} + \|\theta - \theta^h\|_{W^{1,q'}} \leq C(h^2((\|\theta - \theta^h\|_{W^{1,q'}} + \|u - u_h\|_{W^{1,p_0}}) \\
+ (\|\theta - \theta^h\|^2_{W^{1,q'}} + \|u - u_h\|^2_{W^{1,p_0}}) + h'). \\
\end{align*}
\]
(5.15)
(5.16)

Notice that \((u - u_h)\|_{W^{1,q'}} + \|\theta - \theta^h\|_{W^{1,q'}} \to 0\) as \(h \to 0\). Therefore, we have the error estimate

\[
\|u - u_h\|_{L^{p_0}} + \|\theta - \theta^h\|_{L^{p_0}} \leq C^2, \\
\]
(5.17)

for sufficient small \(h\). Here the constant \(C\) is dependent on the solution \((u, \theta)\) of (1.1) but independent of \(h\).

The second step: Let us estimate \(\|u - u^h\|_{L^{p_0}} + \|\theta - \theta^h\|_{L^{p_0}}\). Introduce the following system

\[
\begin{align*}
-\nabla \cdot (\mu(\theta)\nabla w) + 2\nabla \cdot (T\mu(\theta)\nabla u) &= |u_h - u|^{q-1} \text{sign}(u_h - u), \\
-\Delta T + T\mu'(\theta)|\nabla u|^2 + \mu'(\theta)\nabla u \cdot \nabla w &= |\theta^h - \theta|^{q-1} \text{sign}(\theta^h - \theta), \\
\|w\|_{T} &= 0, \\
T &= 0. \\
\end{align*}
\]
(5.18)

From the auxiliary system (5.18), we similarly get

\[
\begin{align*}
\|u - u_h\|_{L^{p_0}} + \|\theta - \theta^h\|_{L^{p_0}} + \|\theta - \theta^h\|_{L^{p_0}} \\
&= a(\theta; u^h - u, w) + (\mu(\theta)(\theta^h - \theta)\nabla u, \nabla w) + d(\theta^h - \theta, T) \\
&+ (\mu(\theta) - \mu(\theta^h))\|\nabla u\|^2, T) + (\mu(\theta)(\|\nabla u\|^2 - |\nabla u|^2), T) \\
&- (\mu(\theta)\|\nabla u\|^2, T) - (\mu(\theta)(\|\nabla u\|^2 - |\nabla u|^2), T) \\
&+ 2(\mu(\theta)\nabla u \cdot (\nabla u - u_h), |\nabla u|, T) - (\mu(\theta)(\nabla u - u_h)(\nabla u - u_h), T) \\
&= a(\theta; u^h, w) - a(\theta; u^h, w) + (\mu(\theta)(\theta^h - \theta)\nabla u, \nabla w) \\
&+ a(\theta; u^h, w) - a(\theta; u, w) + d(\theta^h - \theta, T) \\
&+ (\mu(\theta) - \mu(\theta^h))\|\nabla u\|^2, T) + (\mu(\theta)(\|\nabla u\|^2 - |\nabla u|^2), T) \\
&+ (\mu(\theta^h)\|\nabla u\|^2, T) - (\mu(\theta)(\theta^h - \theta)\|\nabla u\|^2, T) \\
&+ 2(\mu(\theta)\nabla u \cdot (\nabla u - u_h), |\nabla u|, T) - (\mu(\theta)(\nabla u - u_h)(\nabla u - u_h), T) \\
&\leq C((\|\theta - \theta^h\|^2_{W^{1,q'}} + \|u - u_h\|^2_{W^{1,q'}}) \|w\|_{W^{2,p_r}} + \|T\|_{W^{2,q'}}) \\
\leq C((\|\theta - \theta^h\|^2_{W^{1,q'}} + \|u - u_h\|^2_{W^{1,p_0}}) \|w\|_{W^{2,p_r}} + \|T\|_{W^{2,q'}}), \\
\end{align*}
\]
Proof. Theorem 12. So, we show that
\[ \|u - u_h\|_{\mu^*, \theta} + \|\theta - \theta_h\|_{L^2} \leq C h^{r+1}. \] (5.19)
The proof of Theorem 11 is completed. \(\square\)

6. Error estimate in \(L^2\times L^2\)-norm

In this section, we prove the \(L^2\)-norm error estimation. From the results in Section 5, we derive optimal error estimates for \(u - u_h\) in \(L^2\)-norm and \(H^1\)-norm, since \(\mu^* > 2\). But, in the general case, \(q^*\) is small such that \(q^* < 2\). So error estimates of \(\theta - \theta_h\) in \(L^2\)-norm and \(H^1\)-norm are not derived directly from the results in Section 5. Next we consider the optimal \(L^2\)-error estimate for \(\theta - \theta_h\).

Theorem 12. Let \((u_h, \theta_h)\) be a strongly convergent subsequence of the finite element procedure (3.1) associated with limit \((u, \theta)\), which is a solution of the nonlinear system (1.1). Under the assumptions of Theorem 11, there holds the a priori error estimate
\[ \|u - u_h\|_{L^2} + \|\theta - \theta_h\|_{L^2} \leq C h^{r+1}. \] (6.1)

Proof. Since \(\mu^* > 2\), it is obvious that
\[ \|u - u_h\|_{L^2} + h \|\nabla (u - u_h)\|_{L^2} \leq C h^{r+1}. \] (6.2)
On the other hand, we have
\[ \|\nabla (\theta - \theta_h)\|_{L^2} \leq C (\|\theta - \theta_h\|_{L^2} \|\nabla u\|_{L^2}^2 + \|\nabla (u - u_h)\|_{L^2}^2 + (\|\nabla u_h\|_{L^2} + \|\nabla u\|_{L^2}) \|\theta - \theta_h\|_{L^2}) \leq \inf_{\eta \in \mathcal{V}_h} \|\theta - \eta\|_{H^1}). \] (6.3)
Noticing (6.2) and using the inverse estimate of finite element space, we get \(\|\nabla u - \nabla u_h\|_{L^2} = O(h^s) (s > 0)\), and next we see that \(\|\nabla u_h\|_{L^2}\) is bounded. Now from Theorem 11, then it follows
\[ \|\nabla (\theta - \theta_h)\|_{L^2} \leq C h^r. \] (6.4)
In order to estimate \(\|\theta - \theta_h\|_{L^2}^2\), we use the duality technique similarly. Introduce the following system
\[ \begin{cases} -\nabla \cdot (\mu(\theta)\nabla u) + 2\nabla \cdot (T \mu(\theta)\nabla u) = u_h - u, \\ -\Delta T \mu'(\theta)|\nabla u|^2 + \mu'(\theta)\nabla u \cdot \nabla w = \theta_h - \theta, \\ w|_{r} = 0, \\ T|_{r} = 0, \end{cases} \] (6.5)
The similar regularity to (5.9) can also be obtained, then by the argument of Theorem 11, we know
\[ \|u - u_h\|_{L^2}^2 + \|\theta - \theta_h\|_{L^2}^2 = a(\theta; u_h, w) - a(\theta; u, w) + (\mu'(\theta)\nabla u, \nabla w) \]
\[ + a(\theta; u_h, w) - a(\theta; u, w) + d(\theta_h - \theta, T) \]
\[ + ((\mu(\theta) - \mu(\theta_h))|\nabla u|^2, T) + (\mu(\theta)(|\nabla u|^2 - |\nabla u_h|^2), T) \]
\[ + (2(\mu(\theta)\nabla u \cdot (\nabla u - \nabla u_h)), T) + (\mu(\theta)(|\nabla u|^2 - |\nabla u_h|^2), T) \]
\[ \leq C (\|\theta - \theta_h\|_{L^2}^2 + \|u - u_h\|_{H^1}^2 + \|T\|_{H^2}) \]
\[ + (\|\theta - \theta_h\|_{H^1} + \|\nabla (u - u_h)\|_{L^2}(\|w - u_h\|_{H^1} + \|T - \eta_h\|_{H^1})) \]
\[ + \|\nabla u_h^\prime + |\nabla u|_{L^2}|\nabla (u - u_h)\|_{L^2}(\|T - \eta_h\|_{L^2}) \]
\[ \leq C (h^{2r} + h^{r+1})(\|w\|_{H^2} + \|T\|_{H^2}) \]
\[ \leq C (h^{2r} + h^{r+1})(\|u_h - u\|_{L^2} + \|\theta_h - \theta\|_{L^2}). \] (6.6)
So, we show that
\[ \| u - u^h \|_{L^2} + \| \theta - \theta^h \|_{L^2} \leq Ch^{r+1}. \] (6.7)
The proof is finished. \(\square\)

**Remark 13.** The nonlinear finite element equations (3.1) can be solved generally by iterative methods. An iterative algorithm based on a fixed point method was proposed and analyzed in [15], where some numerical implementations were also presented therein.

**Acknowledgements**

Y. Chang’s work is supported partially by the Research Fund for Doctoral Program of High Education by China State Education Ministry. D. Yang’s work is supported partially by the Research Fund for Doctoral Program of High Education by China State Education Ministry.

**References**