INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 9, Number 4, Pages 876–891

ON ERROR ESTIMATES OF THE PENALTY METHOD FOR THE UNSTEADY CONDUCTION-CONVECTION PROBLEM I: TIME DISCRETIZATION

HAIYAN SUN, YINNIAN HE, AND XINLONG FENG

Abstract. In this paper, the penalty method is proposed and discussed for the unsteady conduction-convection problem in two dimensions. In addition, we analyze its time discretization which is based on the backward Euler implicit scheme. Finally, the main results of this paper that optimal error estimates are obtained for the penalty system and the time discretization under reasonable assumptions on the physical data.

Key words. Unsteady conduction-convection problem, Penalty method, Time discretization, Optimal error estimates.

1. Introduction

In this paper, let Ω be a bounded domain in $\mathbb{R}^2$ with C^2 boundary $\partial\Omega$ or a convex polygon. Now we consider the following unsteady conduction-convection problem (cf. [3, 5]).

Problem (I) : Find u, p and T such that for $t_N > 0$,

$$
(1)\qquad \begin{cases} u_t - \nu\Delta u + (u\cdot\nabla)u + \nabla p = \lambda jT, & (x,t)\in\Omega\times(0,t_N),\\ \operatorname{div} u = 0, & (x,t)\in\Omega\times(0,t_N),\\ T_t - \lambda^{-1}\Delta T + u\cdot\nabla T = 0, & (x,t)\in\Omega\times(0,t_N),\\ u(x,t)=0,\ T(x,t)=0, & (x,t)\in\partial\Omega\times(0,t_N),\\ u(x,0)=0,\ T(x,0)=\varphi(x), & x\in\Omega, \end{cases}
$$

where $u=(u_1(x,t),u_2(x,t))$ represents velocity vector, $p(x,t)$ the pressure, $T(x,t)$ the temperature, $\nu>0$ the viscosity, $\lambda^{-1}>0$ the thermal diffusivity, $j=(0,1)$ the two-dimensional unit vector, $\varphi(x,y)$ is the given function, t_N is the final time.

The unsteady conduction-convection Problem (I) is an important dissipative nonlinear system in atmospheric dynamics. It is the coupled equations governing viscous incompressible flow and heat transfer process [6, 22], where the incompressible flow is the Boussinesq approximation to the unsteady Navier-Stokes equations. There are many numerical methods have been studied on the conduction-convection problem (see [2, 5]) and many literatures (see [12, 13, 14, 15, 18]) are put into the construction, analysis and implementation for conduction-convection problem. Shen [19] firstly analyzed the existence uniqueness of approximation solution for steady conduction-convection equations with the Bernadi-Raugel element. Luo and his coworkers gave an optimizing reduced PLSMFE in [14] and a least squares Galerkin/Petrov mixed finite element method in [15]. Shi provided nonconforming

Received by the editors July 15, 2010 and, in revised form September 27, 2011.
2000 *Mathematics Subject Classification.* 65N15, 65Z05.
This work is in parts supported by the National Natural Science Foundation of China (No.10971166, No.10901131), the National High Technology Research and Development Program of China (863 Program, No.2009AA01A135), the China Postdoctoral Science Foundation (No.200801448, No.20070421155), and the Natural Science Foundation of Xinjiang Province (No.2010211B04).

mixed finite element method in [18]. An analysis of conduction natural convection conjugate heat transfer in the gap between concentric cylinders under solar irradiation was studied in [11], etc.

As we known, the velocity u, the pressure p and the temperature T are coupled together by the impressibility constraint "div$u = 0$" and two dissipative nonlinear equation, which make the system is difficult to solve by using the numerical methods. In order to overcome coupled problem, the penalty method as a popular pseudo-compressibility strategy which initially proposed by Courant [4] is popular used (see [7, 16, 17]). Temam [22] firstly applied it to the Navier-Stokes equations. Then, many works appeared on this subject. Shen [16] derived the optimal error estimates for the unsteady Navier-Stokes equations as follows:

$$\tau^{\frac{1}{2}}(t_n)\|u(t_n) - u_\varepsilon(t_n)\|_{L^2} + \tau(t_n)\|u(t_n) - u_\varepsilon(t_n)\|_{H^1} \leq C\varepsilon,$$

for $t_n \in [0, t_N]$, where $\tau(t_n) = \min\{1, t_n\}$, C is a general positive constant and $u(t_n), u_\varepsilon(t_n)$ are the solution of the Navier-Stokes equations and its penalty system, respectively. Recently, He [7] extended it to the finite element method. For the viscoelastic Oldroyd flow problem, Wang et al derived the optimal error estimates for the penalty system [23] and extended it to the fully discrete schemes [24]. This motivates our interest in solving more complicated problem by this method and we have investigated the unsteady conduction-convection problem. For the unsteady conduction-convection problem, the penalty method for Problem (I) is as follows.

Problem (II): Find $u_\varepsilon = (u_{1\varepsilon}, u_{2\varepsilon})$, p_ε and T_ε such that for $t_N > 0$,

$$(2) \qquad \begin{cases} u_{\varepsilon t} - \nu\Delta u_\varepsilon + \tilde{B}(u_\varepsilon, u_\varepsilon) + \nabla p_\varepsilon = \lambda j T_\varepsilon, & (x,t) \in \Omega \times (0, t_N), \\ \text{div } u_\varepsilon + \frac{\varepsilon}{\nu} p_\varepsilon = 0, & (x,t) \in \Omega \times (0, t_N), \\ T_{\varepsilon t} - \lambda^{-1}\Delta T_\varepsilon + \tilde{B}(u_\varepsilon, T_\varepsilon) = 0, & (x,t) \in \Omega \times (0, t_N), \\ u_\varepsilon(x,t) = 0, \ T_\varepsilon(x,t) = 0, & (x,t) \in \partial\Omega \times (0, t_N), \\ u_\varepsilon(x,0) = 0, \ T_\varepsilon(x,0) = \varphi(x), & x \in \Omega, \end{cases}$$

where $0 < \varepsilon < 1$ is a penalty parameter,

$$\tilde{B}(u_\varepsilon, v_\varepsilon) = (u_\varepsilon \cdot \nabla)v_\varepsilon + \frac{1}{2}(\text{div}u_\varepsilon)v_\varepsilon \quad \text{and} \quad \tilde{B}(u_\varepsilon, T_\varepsilon) = u_\varepsilon \cdot \nabla T_\varepsilon + \frac{1}{2}(\text{div}u_\varepsilon)T_\varepsilon$$

is the modified bilinear term, $(\text{div}u_\varepsilon)v_\varepsilon$ and $(\text{div}u_\varepsilon)T_\varepsilon$ are introduced to ensure the dissipativity of Problem (II) as $(\text{div}u)v$ is introduced in the Navier-Stokes equations by Temam [21] to ensure the dissipativity of the Navier-Stokes equations. In this way, p_ε can be eliminated to obtain a penalty system that only contains $u_\varepsilon, T_\varepsilon$, which is much easier to solve than the original equations. Zhang and He have analyzed the penalty finite element for the stationary conduction convection problems [25] and the non-stationary conduction convection problems [26], they have given that, for all $t_n \in [0, t_N]$,

$$\begin{aligned} &\|u(t_n) - u_\varepsilon(t_n)\|_{L^2} + \left(\int_0^{t_n} \|u(t) - u_\varepsilon(t)\|_{H^1}^2 dt\right)^{\frac{1}{2}} + \|T(t_n) - T_\varepsilon(t_n)\|_{L^2} \\ (3) \qquad &+ \left(\int_0^{t_n} \|T(t) - T_\varepsilon(t)\|_{H^1}^2 dt\right)^{\frac{1}{2}} \leq C\sqrt{\varepsilon}, \end{aligned}$$

under the assumptions that the exact solutions are sufficiently smooth. When we consider the discrete problem for the penalty system (2), the estimate (3) is misleading. For instance, if the backward Euler scheme is applied to the penalized system (2), the estimate (3) would lead to

$$\|u(t_n) - u_\varepsilon^n\|_{L^2} + (k\sum_{m=1}^{n}\|u(t_m) - u_\varepsilon^m\|_{H^1}^2)^{\frac{1}{2}} + \|T(t_n) - T_\varepsilon^n\|_{L^2}$$

$$+ (k\sum_{m=1}^{n}\|T(t_m) - T_\varepsilon^m\|_{H^1}^2)^{\frac{1}{2}} \le C(k + \sqrt{\varepsilon}). \tag{4}$$

where $0 < k < 1$ is the time step size, $t_n = nk$ and $(u_\varepsilon^n, T_\varepsilon^n)$ is a penalty approximation of (u, T) at the time t_n. (4) suggests the choice $\varepsilon = \Delta t^2$, which would result in a very ill conditioned system when we make a further spatial discretization (see [16]).

The main focus of this paper is to apply the techniques in [16] to unsteady conduction-convection problem and derive the optimal error estimates for the penalty system and its time discretization. Under some realistic assumptions of the initial value $(u_0,\ \varphi(x))$, we have the following error estimates

$$\tau^{\frac{1}{2}}(t_n)\|u(t_n) - u_\varepsilon(t_n)\|_{L^2} + \tau(t_n)\|u(t_n) - u_\varepsilon(t_n)\|_{H^1} + \tau^{\frac{1}{2}}(t_n)\|T(t_n) - T_\varepsilon(t_n)\|_{L^2}$$
$$+ \tau(t_n)\|T(t_n) - T_\varepsilon(t_n)\|_{H^1} \le C\varepsilon,$$
$$\tau^{\frac{1}{2}}(t_n)\|u(t_n) - u_\varepsilon^n\|_{L^2} + \tau(t_n)\|u(t_n) - u_\varepsilon^n\|_{H^1} + \tau^{\frac{1}{2}}(t_n)\|T(t_n) - T_\varepsilon^n\|_{L^2}$$
$$+ \tau(t_n)\|T(t_n) - T_\varepsilon^n\|_{H^1} \le C(k + \varepsilon),$$

for sufficiently small ε and k, which substantially improve the previous results (3) and (4) and lead to the proper choices of ε for time discretizations of the penalized system.

The remainder of this paper is organized as follows. We firstly introduce some notations and preliminary results for Problem (I) in the next section. Then we provide error behavior for the linearized penalty system in Section 3 and for the nonlinear penalty system in Section 4. In Section 5, we analyze the backward Euler time discretization scheme for the nonlinear penalty system. Finally, conclusions are given in Section 6.

2. Preliminaries

In this section, we describe some of the notations and results which will be frequently used in this paper. For the mathematical setting of conduction-convection Problem (I) and the penalty conduction-convection Problem (II), we introduce the following Hilbert spaces

$$X = H_0^1(\Omega)^2,\ W = H_0^1(\Omega),\ M = \{q \in L^2(\Omega); \int_\Omega q dx = 0\}.$$

The norm corresponding to $H^i(\Omega)^2$ or $H^i(\Omega)$ will be denote $||\cdot||_i$ for i=1, 2. In particular, we use $(\cdot,\cdot)$ and $\|\cdot\|_0$ to denote the inner product and norm in $L^2(\Omega)^2$ or $L^2(\Omega)$. The spaces X and W are equipped with their usual scalar product and norm

$$((u, v)) = (\nabla u, \nabla v),\ \|\nabla u\|_0 = ((u, u))^{1/2}.$$

We also introduce the Hilbert space H and V defined by

$$H = \{v \in L^2(\Omega)^2;\ \text{div}\ v = 0,\ v \cdot n|_{\partial\Omega} = 0\},\ V = \{v \in X; \text{div} v = 0\},$$

Define $Au = -\Delta u$ and $A_\varepsilon u = -\Delta u - \nabla \text{div} u/\varepsilon$, which are the operators associated with the conduction-convection problem and the penalty conduction-convection problem, respectively. They are the positive self-adjoint operators from $D(A) = H^2(\Omega)^2 \cap X$ (or $H^2(\Omega) \cap W$) onto $L^2(\Omega)^2$(or $L^2(\Omega)$). It is valid that

$$(\widetilde{A}u, v) = (\widetilde{A}^{1/2}u, \widetilde{A}^{1/2}v), \quad \forall u \in D(A), v \in X \text{ (or } W),$$

where $\widetilde{A} = A$ or A_ε. In particular, there holds

$$(A^{1/2}u, A^{1/2}v) = (\nabla u, \nabla v), \ \forall u, v \in X \text{ (or } W). \tag{5}$$

$$(A_\varepsilon^{1/2}u, A_\varepsilon^{1/2}v) = (A^{1/2}u, A^{1/2}v) + \frac{1}{\varepsilon}(\text{div}u, \text{div}v), \ \forall u, v \in X. \tag{6}$$

It is known that (see [1, 9]):

$$\|u\|_{L^4} \leq \gamma_0 \|u\|_0^{\frac{1}{2}} \|A^{\frac{1}{2}}u\|_0^{\frac{1}{2}}, \ \|u\|_0 \leq \gamma_0 \|A^{\frac{1}{2}}u\|_0, \qquad \forall\ u \in X \text{ (or } W), \tag{7}$$

$$\|A^{\frac{1}{2}}u\|_{L^4} \leq \gamma_0 \|A^{\frac{1}{2}}u\|_0^{\frac{1}{2}} \|Au\|_0^{\frac{1}{2}}, \ \|A^{\frac{1}{2}}u\|_0 \leq \gamma_0 \|Au\|_0, \qquad \forall u \in D(A), \tag{8}$$

where γ_0 is a positive constant depending only on Ω, which may stand for different values at its different occurrences. Furthermore, we recall the following lemma given in [16].

Lemma 2.1. *There exists a constant $C > 0$, depending only on Ω and such that for ε sufficiently small, we have*

$$\begin{aligned}
&\|Au\|_0 \leq C\|A_\varepsilon u\|_0, \ \forall u \in D(A), \\
&\|A^{\frac{1}{2}}u\|_0 \leq C\|A_\varepsilon^{\frac{1}{2}}u\|_0, \ \forall u \in X, \\
&\|A_\varepsilon^{-1}u\|_0 \leq C\|u\|_{-2}, \ \forall\ u \in H^{-2}(\Omega)^2.
\end{aligned}$$

Associated with the conduction-convection Problem (I) and the penalty conduction-convection Problem (II), we define the continuous bilinear forms

$$\begin{aligned}
&a(u, v) = \nu(\nabla u, \nabla v), \qquad a_\varepsilon(u, v) = \nu(A_\varepsilon^{1/2}u, A_\varepsilon^{1/2}v), \ \forall u, v \in X, \\
&\bar{a}(T, \varphi) = \lambda^{-1}(\nabla T, \nabla \phi), \ \forall T, \varphi \in W, \\
&d(v, q) = (q, \text{div}v), \ \forall v \in X, q \in M,
\end{aligned}$$

respectively. We also introduce a continuous trilinear form $\tilde{b}(\cdot, \cdot, \cdot)$ on $X \times X \times X$, $\hat{b}(\cdot, \cdot, \cdot)$ on $X \times W \times W$, respectively.

$$\tilde{b}(u, v, w) = \frac{1}{2} \int_\Omega \sum_{i,k=1}^{2} (u_i \frac{\partial v_k}{\partial x_i} w_k - u_i \frac{\partial w_k}{\partial x_i} v_k) dx, \ \ \forall\ u, v, w \in X,$$

$$\hat{b}(u, T, \psi) = \frac{1}{2} \int_\Omega \sum_{i=1}^{2} (u_i \frac{\partial T}{\partial x_i} \psi - u_i \frac{\partial \psi}{\partial x_i} T) dx, \ \ \forall\ u \in X, \ T, \psi \in W.$$

Some estimates of the trilinear $\tilde{b}(\cdot, \cdot, \cdot)$ can be found in [8, 20, 27].

$$\tilde{b}(u, v, w) = -\tilde{b}(u, w, v), \quad \forall u, v, w \in X, \tag{9}$$

$$|\tilde{b}(u, v, w)| \leq C\|u\|_1 \|v\|_1^{\frac{1}{2}} \|v\|_2^{\frac{1}{2}} \|w\|_0, \quad \forall u, w \in X, v \in D(A), \tag{10}$$

$$\begin{aligned}
&|\tilde{b}(u, v, w)| + |\tilde{b}(v, u, w)| + |\tilde{b}(w, u, v)| \\
&\qquad \leq C\|u\|_2 \|v\|_1 \|w\|_0, \ \forall u \in D(A), \ v, \ w \in X.
\end{aligned} \tag{11}$$

Similarly, it is easy to verify that $\hat{b}(\cdot,\cdot,\cdot)$ satisfy the following important property

$$\hat{b}(u,T,\psi) = -\hat{b}(u,\psi,T), \quad \forall u \in X,\ T,\psi \in W, \tag{12}$$

$$|\hat{b}(u,T,\psi)| \le C\|u\|_1\|T\|_1^{\frac{1}{2}}\|T\|_2^{\frac{1}{2}}\|\psi\|_0, \quad \forall u \in X, T \in D(A), \psi \in W, \tag{13}$$

$$\begin{aligned} &|\hat{b}(u,T,\psi)| + |\hat{b}(T,u,\psi)| + |\hat{b}(\psi,T,u)| \\ &\qquad \le C\|u\|_1\|T\|_2\|\psi\|_0,\ \forall\, T \in D(A),\ u \in X, \psi \in W. \end{aligned} \tag{14}$$

With above notations, the variational formulation of Problem (I) is written as follow.

Problem (III): Find $(u,p,T) \in (X,M,W)$, for all $t \in [0,t_N]$, such that for all $(v,\phi,\psi) \in (X,M,W)$:

$$\left\{ \begin{array}{l} (u_t,v) + a(u,v) + \tilde{b}(u,u,v) - d(v,p) = \lambda(jT,v), \\ d(u,\phi) = 0, \\ (T_t,\psi) + \bar{a}(T,\psi) + \hat{b}(u,T,\psi) = 0, \end{array} \right. \tag{15}$$

and the variational formulation of the penalty system Problem (II) reads as.

Problem (IV): Find $(u_\varepsilon,p_\varepsilon,T_\varepsilon) \in (X,M,W)$, for all $t \in [0,t_N]$, such that for all $(v,\phi,\psi) \in (X,M,W)$:

$$\left\{ \begin{array}{l} (u_{\varepsilon t},v) + a(u_\varepsilon,v) + \tilde{b}(u_\varepsilon,u_\varepsilon,v) - d(v,p_\varepsilon) = \lambda(jT_\varepsilon,v), \\ d(u_\varepsilon,\phi) + \frac{\varepsilon}{\nu}(p_\varepsilon,\phi) = 0, \\ (T_{\varepsilon t},\psi) + \bar{a}(T_\varepsilon,\psi) + \hat{b}(u_\varepsilon,T_\varepsilon,\psi) = 0, \end{array} \right. \tag{16}$$

with $u(x,0) = u_\varepsilon(x,0) = 0,\ T(x,0) = T_\varepsilon(x,0) = \varphi(x)$, respectively.

Theorem 2.2. [13] If $\varphi(x) \in H^2(\Omega)$, Problem (III) has a unique solution $(u,p,T) \in [L^2(0,t_N;X) \cap H^1(0,t_N;V)] \times L^2(0,t_N;M) \times H^1(0,t_N;W)$, satisfies

$$\|\nabla T^{(i)}\|_0 + \|T^{(i)}\|_0 + \|\nabla u^{(i)}\|_0 + \|u^{(i)}\|_0 \le \theta(t), \quad 0 \le i \le 3,$$

where $\theta(t)$ is continuous general positive function about t only depends on the data $\varphi(x)$.

Theorem 2.3. If $\varphi(x) \in H^2(\Omega)$, Problem (IV) has a unique solution $(u_\varepsilon,p_\varepsilon,T_\varepsilon) \in [L^\infty(0,t_N;H^2(\Omega)) \cap L^2(0,t_N;X)] \times L^2(0,t_N;M) \times H^1(0,t_N;W)$.

The proof of Theorem 2.3 and the solution $(u_\varepsilon,p_\varepsilon,T_\varepsilon)$ of Problem (IV) converges to the solution (u,p,T) of Problem (III) uniformly as $\varepsilon \to 0$ are similar to the proof of Theorem 3.1 and Theorem 3.2 of [26] and much easier than that, here we omit it.

By using a similar argument to [21], we have the following properties.

(A1). Assume that the initial velocity $u_0 \in V$, the initial temperature $\varphi(x) \in W$, then $T \in L^\infty(0,T;L^2(\Omega))$, there exists a finite time $t_{M1} < t_N$ such that

$$\begin{aligned} &u \in C([0,t_{M1}];V) \cap L^2([0,t_{M1}];H^2(\Omega)^2), \\ &T \in L^2([0,t_{M1}];H^2(\Omega)),\ p \in L^2([0,t_{M1}];H^1(\Omega)/R). \end{aligned} \tag{17}$$

(A2). Assume that $\varphi(x) \in H^2(\Omega)$, then $tT_t \in L^2([0,t_N];L^2(\Omega))$.

By using the smoothing property of the conduction-convection problem, that

$$tp_t \in L^2([0,t_N];H^1(\Omega)). \tag{18}$$

Similar to the penalty Navier-Stokes equation [16], one can show that assumption (A1), there exists a finite time $t_{M2} < t_N$ and a constant C independent of ε such that

$$\|u_\varepsilon(t)\|_1^2 + \int_0^t \|u_\varepsilon(t)\|_2^2 ds \le C, \quad \forall t \in [0,t_{M2}].$$

In the sequel, we restrict ourselves to the interval $[0, t_M]$ with $t_M = \min\{t_{M1}, t_{M2}\}$.

We also recall two lemmas of Gronwall type which will be frequently used [21].

Lemma 2.4. (Gronwall Lemma). *Let $y(t),\ g(t),\ h(t),\ f(t)$ be nonnegative functions such that* $\int_0^{t_N} g(t)dt \le M$ *and*

$$y(t) + \int_0^t h(s)ds \le y(0) + \int_0^t (g(s)y(s) + f(s))ds,\ \forall\ 0 \le t \le t_N,$$

then

$$y(t) + \int_0^t h(s)ds \le \exp(M)\Big(y(0) + \int_0^t f(s)ds\Big),\ \forall\ 0 \le t \le t_N.$$

Lemma 2.5. (Discrete Gronwall Lemma). *Let C and $a_n,\ b_n,\ d_n$, for integer $1 \le m \le \frac{t_N}{k}$, be nonnegative numbers, such that*

$$a_m + k\sum_{n=0}^{m} b_n \le k\sum_{n=0}^{m} a_n d_n + C,\ \forall\ 1 \le m \le \frac{t_N}{k}. \tag{19}$$

Assume that $kd_n < \frac{1}{2}$ and for all $1 \le n \le \frac{t_N}{k}$ then

$$a_m + k\sum_{n=1}^{m} b_n \le C\exp(2k\sum_{n=1}^{m-1} a_n),\ \forall\ 1 \le m \le \frac{t_N}{k}. \tag{20}$$

3. Error estimates for the linearized problem

In this section, we will consider the following linear problem. The results in this section will be used in the next section as an intermediate step for analyzing the nonlinear conduction-convection problem.

$$\begin{cases} u_t - \nu\Delta u + \nabla p = \lambda jT, \text{div } u = 0, & (x,t) \in \Omega \times (0, t_N), \\ T_t - \lambda^{-1}\Delta T = 0, & (x,t) \in \Omega \times (0, t_N), \\ u(x,t) = T(x,t) = 0, & (x,t) \in \partial\Omega \times (0, t_N). \end{cases} \tag{21}$$

The penalty method of (21) is as follows:

$$\begin{cases} u_{\varepsilon t} - \nu\Delta u_\varepsilon + \nabla p_\varepsilon = \lambda jT_\varepsilon, \text{div } u_\varepsilon + \frac{\varepsilon}{\nu}p_\varepsilon = 0, & (x,t) \in \Omega \times (0, t_N), \\ T_{\varepsilon t} - \lambda^{-1}\Delta T_\varepsilon = 0, & (x,t) \in \Omega \times (0, t_N), \\ u_\varepsilon(x,t) = 0,\ T_\varepsilon(x,t) = 0, & (x,t) \in \partial\Omega \times (0, t_N), \end{cases} \tag{22}$$

with $u(x,0) = u_\varepsilon(x,0) = 0,\ T(x,0) = T_\varepsilon(x,0) = \varphi(x)$, respectively.

Similarly, the variational formulation of the problem (21) is defined as follows: find $(u,p,T) \in (X,M,W)$, for all $t \in [0,t_N]$, such that for all $(v,\phi,\psi) \in (X,M,W)$,

$$\begin{cases} (u_t, v) + a(u,v) - d(v,p) + d(u,\phi) = \lambda(jT, v), \\ (T_t, \psi) + \bar{a}(T,\psi) = 0. \end{cases} \tag{23}$$

The penalty method applied to (23) is that: find $(u_\varepsilon, p_\varepsilon, T_\varepsilon) \in (X,M,W)$, for all $t \in [0,t_N]$, such that for all $(v,\phi,\psi) \in (X,M,W)$,

$$\begin{cases} (u_{\varepsilon t}, v) + a(u_\varepsilon,v) - d(v,p_\varepsilon) + d(u_\varepsilon,\phi) + \frac{\varepsilon}{\nu}(p_\varepsilon,\phi) = \lambda(jT_\varepsilon, v), \\ (T_{\varepsilon t}, \psi) + \bar{a}(T_\varepsilon,\psi) = 0, \end{cases} \tag{24}$$

with $u(x,0) = u_\varepsilon(x,0) = 0,\ T(x,0) = T_\varepsilon(x,0) = \varphi(x)$, respectively.

Setting $e = u - u_\varepsilon,\ q = p - p_\varepsilon,\ \xi = T - T_\varepsilon$, it follows that $e(0) = \xi(0) = 0$. We shall derive a sequence of estimates for the penalty error e, q and ξ.

Subtracting (24) from (23), we obtain

$$\left\{\begin{array}{l}(e_t,v)+a(e,v)-d(v,q)+d(e,\phi)+\frac{\varepsilon}{\nu}(q,\phi)=\lambda(j\xi,v)+\frac{\varepsilon}{\nu}(p,\phi),\\(\xi_t,\psi)+\bar{a}(\xi,\psi)=0.\end{array}\right. \tag{25}$$

Lemma 3.1. *Suppose (A1) is valid, ε sufficiently small, for all $t\in[0,t_M]$, we have*

$$\|e\|_0^2+\int_0^t\|\nabla e\|_0^2ds+\varepsilon\int_0^t\|q\|_0^2ds\le C\varepsilon, \tag{26}$$

$$\int_0^t\|e\|_0^2ds\le C\varepsilon^2, \tag{27}$$

$$\|\xi\|_0^2+\int_0^t\|\nabla\xi\|_0^2ds=0. \tag{28}$$

Proof. Taking $(v,\phi,\psi)=(e,q,\xi)$ in (25) and using (7), we have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|e\|_0^2+\nu\|\nabla e\|_0^2+\frac{\varepsilon}{\nu}\|q\|_0^2&=\frac{\varepsilon}{\nu}(p,q)+\lambda(j\xi,e)\\&\le\frac{\varepsilon}{2\nu}\|q\|_0^2+\frac{\varepsilon}{2\nu}\|p\|_0^2+\frac{\nu}{2}\|\nabla e\|_0^2+C\|\nabla\xi\|_0^2,\end{aligned} \tag{29}$$

and

$$\frac{1}{2}\frac{d}{dt}\|\xi\|_0^2+\lambda^{-1}\|\nabla\xi\|_0^2=0. \tag{30}$$

Integrating the above two inequalities from 0 to $t\le t_M$, thanks to $e(0)=0$, $\xi(0)=0$ and (17), we have

$$\|e\|_0^2+\int_0^t\|\nabla e\|_0^2ds+\varepsilon\int_0^t\|q\|_0^2ds\le C\varepsilon, \tag{31}$$

$$\|\xi\|_0^2+\int_0^t\|\nabla\xi\|_0^2ds=0. \tag{32}$$

Now we use the standard parabolic duality argument. For any $0<t\le t_M$, we define (w,φ) by

$$\left\{\begin{array}{l}w_s+\nu\Delta w+\nabla\varphi=e(s),\quad\forall\ 0<s\le t,\\\mathrm{div}w=0,\quad w(t)=0.\end{array}\right. \tag{33}$$

There are the following inequality (see [16]),

$$\nu\int_0^t\|w\|_2^2ds+\int_0^t\|\nabla\varphi\|_0^2ds\le C\int_0^t\|e\|_0^2ds. \tag{34}$$

Taking the inner product of (33) with $e(s)$, because of (25) and $\mathrm{div}w=0$, we derive

$$\begin{aligned}\|e\|_0^2&=(w_s,e)+\nu(\Delta w,e)+(\nabla\varphi,e)\\&=\frac{d}{ds}(w,e)-\lambda(j\xi,w)-\frac{\varepsilon}{\nu}(\varphi,p_\varepsilon).\end{aligned}$$

Integrating from 0 to t, using Schwarz inequality, (7), (32) and (34), since $w(t)=e(0)=0$, we have

$$\int_0^t\|e\|_0^2ds\le C\varepsilon^2\int_0^t\|p_\varepsilon\|_0^2ds\le C\varepsilon^2,\quad\forall t\in[0,t_M].$$

Remark: Note that equation (27) implies that ξ is equal to zero, this indicates that equation (25) is the same as equation (3.3)-(3.4) in Shen [16]. Now, we recall some results from the reference [16] .

Lemma 3.2. *Suppose (A1) and (A2) are valid, ε sufficiently small, for all $t \in [0, t_M]$, we have*

$$\int_0^t s^2 \|p_{\varepsilon t}\|_0^2 ds \le C.$$

Theorem 3.3. *Suppose (A1) and (A2) are valid, ε sufficiently small, for all $t \in [0, t_M]$, we have*

$$t\|e\|_0^2 + \int_0^t s\|\nabla e\|_0^2 ds + \varepsilon \int_0^t s\|q\|_0^2 ds \le C\varepsilon^2, \tag{35}$$

$$t^2\|\nabla e\|_0^2 + \int_0^t s^2\|q\|_0^2 ds \le C\varepsilon^2. \tag{36}$$

4. Error estimates for the nonlinear problem

We consider the following intermediate linear problem:

$$\begin{cases} v_t - \nu\Delta v + \nabla\gamma = \lambda jT - \tilde{B}(u,u), \\ \text{div } v + \frac{\varepsilon}{\nu}\gamma = 0, \\ T_t - \lambda^{-1}\Delta T = -\tilde{B}(u,T), \end{cases} \tag{37}$$

with $v(x,0) = 0,\ T(x,0) = \varphi(x)$, where (u,T) is the solution of Problem (I).

Letting $\rho = v - u,\ \sigma = \gamma - p$ and subtracting Problem (I) from (37), we obtain

$$\rho_t - \nu\Delta\rho + \nabla\sigma = 0, \tag{38}$$

$$\text{div}\rho + \frac{\varepsilon}{\nu}\sigma = -\frac{\varepsilon}{\nu}p, \tag{39}$$

with $\rho(x,0) = 0$.

Lemma 4.1. *Suppose (A1) and (A2) are valid, ε sufficiently small, for all $t \in [0, t_M]$, we have*

$$t\|\rho\|_0^2 + t^2\|\nabla\rho\|_0^2 + \int_0^t \|\rho\|_0^2 ds + \int_0^t s\|\nabla\rho\|_0^2 ds + \int_0^t s^2\|\sigma\|_0^2 ds \le C\varepsilon^2.$$

Proof. From Section 3, we note that the assumption $(A1)$ for a linear problem can be replaced by the weaker condition $\lambda jT \in L^2([0,t_M]; L^2(\Omega))$. Thanks to (17), we have $\lambda jT - \tilde{B}(u,u) \in L^2([0,t_M]; L^2(\Omega))$. On the other hand, it can be easily shown (see for instance [9] that $tu_t \in L^2([0,t_M]; L^2(\Omega))$). Hence

$$t\frac{\partial}{\partial t}(\lambda jT - \tilde{B}(u,u)) = t(\lambda jT_t - \tilde{B}(u_t,u) - \tilde{B}(u,u_t) \in L^2([0,t_M]; L^2(\Omega)).$$

Lemma 4.1 is then a direct consequence of Lemma 3.1 and Theorem 3.3 applied to (38)-(39).

Now, letting $\eta = u_\varepsilon - v,\ \delta = p_\varepsilon - \gamma,\ \varsigma = T_\varepsilon - T$ and subtracting (37) from Problem (II), we obtain

$$\begin{cases} \eta_t - \nu\Delta\eta + \tilde{B}(u_\varepsilon, \eta+\rho) + \tilde{B}(\eta+\rho, u) + \nabla\delta = \lambda j\varsigma, \\ \text{div}\eta + \frac{\varepsilon}{\nu}\delta = 0, \\ \varsigma_t - \lambda^{-1}\Delta\varsigma + \tilde{B}(u_\varepsilon, \varsigma) + \tilde{B}(\eta+\rho, T) = 0, \end{cases} \tag{40}$$

with $\eta(x,0) = 0,\ \varsigma(x,0) = 0$, respectively. The variational formulation of the problem (40) is defined as follows.

Find $(\eta, \delta, \varsigma) \in (X, M, W)$, for all $t \in [0, t_N]$, such that for all $(v, \phi, \psi) \in (X, M, W)$,

$$\left\{ \begin{array}{l} (\eta_t, v) + a(\eta, v) + \tilde{b}(u_\varepsilon, \eta + \rho, v) + \tilde{b}(\eta + \rho, u, v) - d(v, \delta) = \lambda(j\varsigma, v), \\ d(\eta, \phi) + \frac{\varepsilon}{\nu}(\delta, \phi) = 0, \\ (\varsigma_t, \psi) + \bar{a}(\varsigma, \psi) + \hat{b}(u_\varepsilon, \varsigma, \psi) + \hat{b}(\eta + \rho, T, \psi) = 0, \end{array} \right. \tag{41}$$

or

$$\left\{ \begin{array}{l} (\eta_t, v) + a_\varepsilon(\eta, v) + \tilde{b}(u_\varepsilon, \eta + \rho, v) + \tilde{b}(\eta + \rho, u, v) = \lambda(j\varsigma, v), \\ (\varsigma_t, \psi) + \bar{a}(\varsigma, \psi) + \hat{b}(u_\varepsilon, \varsigma, \psi) + \hat{b}(\eta + \rho, T, \psi) = 0, \end{array} \right. \tag{42}$$

Lemma 4.2. *Suppose (A1) and (A2) are valid, ε sufficiently small, for all $t \in [0, t_M]$, we have*

$$t\|\eta\|_0^2 + t^2\|\nabla\eta\|_0^2 + t\|\varsigma\|_0^2 + t^2\|\nabla\varsigma\|_0^2 + \int_0^t s^2\|\delta\|_0^2 ds \le C\varepsilon^2.$$

Proof. Taking $v = A_\varepsilon^{-1}\eta$ in (42), thanks to Lemma 2.1, (11) and Schwarz inequality, we get

$$\frac{1}{2}\frac{d}{dt}\|A_\varepsilon^{-\frac{1}{2}}\eta\|_0^2 + \nu\|\eta\|_0^2 = -\tilde{b}(u_\varepsilon, \eta + \rho, A_\varepsilon^{-1}\eta) - \tilde{b}(\eta + \rho, u, A_\varepsilon^{-1}\eta) + \lambda(j\varsigma, A_\varepsilon^{-1}\eta)$$
$$\le \frac{\nu}{2}\|\eta\|_0^2 + \frac{1}{2}\|\rho\|_0^2 + C(\|u\|_2^2 + \|u_\varepsilon\|_2^2 + 1)\|A_\varepsilon^{-\frac{1}{2}}\eta\|_0^2 + \frac{\lambda^{-1}}{4}\|\varsigma\|_0^2.$$

Taking $\psi = A^{-1}\varsigma$ in (42), thanks to (14) and Schwarz inequality, we get

$$\frac{1}{2}\frac{d}{dt}\|A^{-\frac{1}{2}}\varsigma\|_0^2 + \lambda^{-1}\|\varsigma\|_0^2 = -\hat{b}(u_\varepsilon, \varsigma, A^{-1}\varsigma) - \hat{b}(\eta + \rho, T, A^{-1}\varsigma)$$
$$\le \frac{\lambda^{-1}}{2}\|\varsigma\|_0^2 + C(\|T\|_2^2 + \|u_\varepsilon\|_2^2)\|A^{-\frac{1}{2}}\varsigma\|_0^2 + \frac{\nu}{8}(\|\eta\|_0^2 + \|\rho\|_0^2).$$

Since $\int_0^{t_M}(\|T\|_2^2 + \|u_\varepsilon\|_2^2 + \|u\|_2^2)ds \le C$, we can apply Lemma 2.4 to above two inequalities, use Lemma 4.1, we obtain

$$\|A_\varepsilon^{-\frac{1}{2}}\eta\|_0^2 + \|A^{-\frac{1}{2}}\varsigma\|_0^2 + \nu\int_0^t \|\eta\|_0^2 ds + \lambda^{-1}\int_0^t \|\varsigma\|_0^2 ds \le C\varepsilon^2. \tag{43}$$

Now, taking $(v, \phi) = t(\eta, \delta)$ in (41), summing up the two relations, using (11) and Schwarz inequality, we derive

$$\frac{1}{2}\frac{d}{dt}t\|\eta\|_0^2 + \nu t\|\nabla\eta\|_0^2 + \frac{\varepsilon}{\nu}t\|\delta\|_0^2$$
$$= \frac{1}{2}\|\eta\|_0^2 - t\tilde{b}(u_\varepsilon, \eta + \rho, \eta) - t\tilde{b}(\eta + \rho, u, \eta) + \lambda t(j\varsigma, \eta)$$
$$\le \frac{1}{2}\|\eta\|_0^2 + \frac{\nu t}{2}\|\nabla\eta\|_0^2 + Ct\|\nabla\rho\|_0^2 + Ct(\|u\|_2^2 + \|u_\varepsilon\|_2^2 + 1)\|\eta\|_0^2 + \frac{\lambda^{-1}t}{4}\|\nabla\varsigma\|_0^2.$$

Taking $\psi = t\varsigma$ in (41), using (9), (14) and Schwarz inequality, we derive

$$\frac{1}{2}\frac{d}{dt}t\|\varsigma\|_0^2 + \lambda^{-1}t\|\nabla\varsigma\|_0^2 = \frac{1}{2}\|\varsigma\|_0^2 - t\hat{b}(u_\varepsilon, \varsigma, \varsigma) - t\hat{b}(\eta + \rho, T, \varsigma)$$
$$\le \frac{1}{2}\|\varsigma\|_0^2 + Ct\|T\|_2^2\|\varsigma\|_0^2 + \frac{\nu t}{8}(\|\nabla\eta\|_0^2 + \|\nabla\rho\|_0^2).$$

Integrating over $[0, t]$, we can apply (43), Lemma 4.1 and Lemma 2.4 to above two inequalities, such that

$$t\|\eta\|_0^2 + t\|\varsigma\|_0^2 + \nu\int_0^t s\|\nabla\eta\|_0^2 ds + \lambda^{-1}\int_0^t s\|\nabla\varsigma\|_0^2 ds + \frac{\varepsilon}{\nu}\int_0^t s\|\delta\|_0^2 ds \le C\varepsilon^2. \tag{44}$$

Next, we take the partial derivative with respect to t of the second term of (41) to obtain

$$(\mathrm{div}\eta_t, \phi) + \frac{\varepsilon}{\nu}(\delta_t, \phi) = 0. \tag{45}$$

Now, taking $v = t^2\eta_t$ in (41) and $\phi = t^2\delta$ in (45). Then adding them up, using (11) and Schwarz inequality, we get

$$\begin{aligned}
&t^2\|\eta_t\|_0^2 + \frac{\nu}{2}\frac{d}{dt}t^2\|\nabla\eta\|_0^2 + \frac{\varepsilon}{2\nu}\frac{d}{dt}t^2\|\delta\|_0^2 \\
&= \nu t\|\nabla\eta\|_0^2 + \frac{\varepsilon}{\nu}t\|\delta\|_0^2 - t^2\tilde{b}(u_\varepsilon, \eta + \rho, \eta_t) - t^2\tilde{b}(\eta + \rho, u, \eta_t) + \lambda t^2(j\varsigma, \eta_t) \\
&\le \nu t\|\nabla\eta\|_0^2 + \frac{\varepsilon}{\nu}t\|\delta\|_0^2 + \frac{t^2}{2}\|\eta_t\|_0^2 + C(\|u\|_2^2 + \|u_\varepsilon\|_2^2)(\varepsilon^2 + t^2\|\nabla\eta\|_0^2) + \frac{\lambda^2 t^2}{2}\|\varsigma\|_0^2.
\end{aligned}$$

Taking $\psi = t^2\varsigma$ in (41), using (9), (14) and Schwarz inequality, we derive

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}t^2\|\varsigma\|_0^2 + \lambda^{-1}t^2\|\nabla\varsigma\|_0^2 = t\|\varsigma\|_0^2 - t^2\hat{b}(u_\varepsilon, \varsigma, \varsigma) - t^2\hat{b}(\eta + \rho, T, \varsigma) \\
&\le t\|\varsigma\|_0^2 + Ct^2\|T\|_2^2\|\varsigma\|_0^2 + \delta_3 t^2(\|\nabla\eta\|_0^2 + \|\nabla\rho\|_0^2).
\end{aligned}$$

Integrating over $[0, t]$, using (44), Lemma 2.4 and Lemma 4.1 to above two inequalities and taking δ_3 sufficiently small, we derive

$$\nu t^2\|\nabla\eta\|_0^2 + t^2\|\varsigma\|_0^2 + \frac{\varepsilon}{\nu}t^2\|\delta\|_0^2 + \int_0^t s^2\|\eta_t\|_0^2 ds + \lambda^{-1}\int_0^t s^2\|\nabla\varsigma\|_0^2 ds \le C\varepsilon^2. \tag{46}$$

Taking $\psi = t^2\varsigma_t$ in (41), using (11), (14) and Schwarz inequality, we have

$$\begin{aligned}
&t^2\|\varsigma_t\|_0^2 + \frac{\lambda^{-1}}{2}\frac{d}{dt}t^2\|\nabla\varsigma\|_0^2 \\
&= \lambda^{-1}t\|\nabla\varsigma\|_0^2 - t^2\hat{b}(u_\varepsilon, \varsigma, \varsigma_t) - t^2\hat{b}(\eta + \rho, T, \varsigma_t) \\
&\le \lambda^{-1}t\|\nabla\varsigma\|_0^2 + \frac{t^2}{2}\|\varsigma_t\|_0^2 + Ct^2\|u_\varepsilon\|_2^2\|\nabla\varsigma\|_0^2 + Ct^2\|T\|_2^2\|\nabla(\eta + \rho)\|_0^2.
\end{aligned}$$

Integrating over $[0, t]$, using (44), (46), Lemma 2.4 and Lemma 4.1, we derive

$$\lambda^{-1}t^2\|\nabla\varsigma\|_0^2 + \int_0^t s^2\|\varsigma_t\|_0^2 ds \le C\varepsilon^2.$$

From (40) , there holds

$$\nabla\delta = -\eta_t + \nu\Delta\eta - \tilde{B}(u_\varepsilon, \eta + \rho) - \tilde{B}(\eta + \rho, u) + \lambda j\varsigma.$$

Therefore by using previous estimates on the above equation, we derive

$$\int_0^{t_M} s^2\|\delta\|_0^2 ds \le \int_0^{t_M} s^2\|\nabla\delta\|_{-1}^2 ds \le C\varepsilon^2.$$

By combining Lemma 4.1 with Lemma 4.2, we obtain the following error estimate result.

Theorem 4.3. *Suppose (A1) and (A2) are valid, ε sufficiently small, for all $t \in [0, t_M]$, the following error estimates holds.*

$$\begin{aligned}
&t\|u(t) - u_\varepsilon(t)\|_0^2 + t^2\|\nabla(u(t) - u_\varepsilon(t))\|_0^2 + t\|T(t) - T_\varepsilon(t)\|_0^2 \\
&+t^2\|\nabla(T(t) - T_\varepsilon(t))\|_0^2 + \int_0^t s^2\|p(t) - p_\varepsilon(t)\|_0^2 ds \le C\varepsilon^2.
\end{aligned}$$

5. Time discretizations of the penalized system

In this section, we will analyze the backward Euler time discretization scheme for the nonlinear penalty system. Let $0 < k < 1$ is the time-step size and $t_n = nk$.

Lemma 5.1. *In addition to (A1) and (A2), we assume $u_0 \in H^2(\Omega)^2$, then the solution $(u_\varepsilon,\, T_\varepsilon)$ of Problem (II) satisfies*

$$u_{\varepsilon t} \in L^2([0,t_M];X),\ A_\varepsilon^{-\frac{1}{2}} u_{\varepsilon tt}, \sqrt{t} u_{\varepsilon tt} \in L^2([0,t_M];L^2(\Omega)^2), \tag{47}$$

$$T_{\varepsilon t} \in L^2([0,t_M];W), A^{-\frac{1}{2}} T_{\varepsilon tt}, \sqrt{t} T_{\varepsilon tt} \in L^2([0,t_M];L^2(\Omega)). \tag{48}$$

Proof. By using a similar argument which used by He in [7], we known that $\|u_{\varepsilon t}(0)\|_0$ and $\|T_{\varepsilon t}(0)\|_0$ is bounded. We take the partial derivative with respect to t of problem(IV) to obtain

$$\begin{cases} (u_{\varepsilon tt}, v) + a_\varepsilon(u_{\varepsilon t}, v) + \tilde{b}(u_{\varepsilon t}, u_\varepsilon, v) + \tilde{b}(u_\varepsilon, u_{\varepsilon t}, v) = \lambda(jT_{\varepsilon t}, v), \\ (T_{\varepsilon tt}, \psi) + \bar{a}(T_{\varepsilon t}, \psi) + \hat{b}(u_{\varepsilon t}, T_\varepsilon, \psi) + \hat{b}(u_\varepsilon, T_{\varepsilon t}, \psi) = 0, \end{cases} \tag{49}$$

Now, taking $v = u_{\varepsilon t}$ in (49), using Lemma 2.1, (14) and Schwarz inequality, we obtain

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|u_{\varepsilon t}\|_0^2 + \nu\|A_\varepsilon^{\frac{1}{2}} u_{\varepsilon t}\|_0^2 &= \lambda(jT_{\varepsilon t}, u_{\varepsilon t}) - \tilde{b}(u_{\varepsilon t}, u_\varepsilon, u_{\varepsilon t}) \\ &\le \frac{\nu}{2}\|A_\varepsilon^{\frac{1}{2}} u_{\varepsilon t}\|_0^2 + C\|T_{\varepsilon t}\|_{-1}^2 + C\|u_\varepsilon\|_2^2\|u_{\varepsilon t}\|_0^2. \end{aligned}$$

Taking $\psi = T_{\varepsilon t}$ in Problem (IV), using (14) and Schwarz inequality, we obtain

$$\|T_{\varepsilon t}\|_0^2 + \frac{\lambda^{-1}}{2}\frac{d}{dt}\|\nabla T_\varepsilon\|_0^2 = -\hat{b}(u_\varepsilon, T_\varepsilon, T_{\varepsilon t}) \le \frac{1}{2}\|T_{\varepsilon t}\|_0^2 + C\|u_\varepsilon\|_2^2\|\nabla T_\varepsilon\|_0^2.$$

Integrating over $[0, t]$, using Lemma 2.4 to above two inequalities, we derive

$$\|u_{\varepsilon t}\|_0^2 + \|T_\varepsilon\|_0^2 + \int_0^t \|A_\varepsilon^{\frac{1}{2}} u_{\varepsilon t}\|_0^2 ds + \int_0^t \|T_{\varepsilon t}\|_0^2 ds \le C. \tag{50}$$

Taking $\psi = T_{\varepsilon t}$ in (49), using Lemma 2.4, (14) and Schwarz inequality, we obtain

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|T_{\varepsilon t}\|_0^2 + \lambda^{-1}\|A^{\frac{1}{2}} T_{\varepsilon t}\|_0^2 &= -\hat{b}(u_{\varepsilon t}, T_\varepsilon, T_{\varepsilon t}) \\ &\le \frac{1}{2}\|A_\varepsilon^{\frac{1}{2}} u_{\varepsilon t}\|_0^2 + C\|T_\varepsilon\|_2^2\|T_{\varepsilon t}\|_0^2. \end{aligned}$$

Integrating over $[0, t]$ and using Lemma 2.4, we derive

$$\|T_{\varepsilon t}\|_0^2 + \int_0^t \|A^{\frac{1}{2}} T_{\varepsilon t}\|_0^2 ds \le C. \tag{51}$$

From Problem(II), we known that

$$\lambda^{-1}\Delta T_\varepsilon = T_{\varepsilon t} + \tilde{B}(u_\varepsilon, T_\varepsilon).$$

Taking the inner products of the last relation with ΔT_ε, using Theorem 2.3, and (51), we derive readily that

$$\|T_\varepsilon\|_2 \le C.$$

By using Lemma 2.1 and (50), we have $\int_0^t \|A^{\frac{1}{2}} u_{\varepsilon t}\|_0^2 ds \le C$. Then, we get

$$\|A_\varepsilon^{-\frac{1}{2}} \tilde{B}(u_{\varepsilon t}, u_\varepsilon)\|_0 \le \|\tilde{B}(u_{\varepsilon t}, u_\varepsilon)\|_{-1} \le \sup_{v \in X} \frac{\tilde{b}(u_{\varepsilon t}, u_\varepsilon, v)}{\|v\|_1} \le C\|u_{\varepsilon t}\|_0\|u_\varepsilon\|_2.$$

The same is true for $\tilde{B}(u_\varepsilon, u_{\varepsilon t})$, $\tilde{B}(u_{\varepsilon t}, T_\varepsilon)$ and $\tilde{B}(u_\varepsilon, T_{\varepsilon t})$. Thus

$$A_\varepsilon^{-\frac{1}{2}} u_{\varepsilon tt} = A_\varepsilon^{-\frac{1}{2}} \{\lambda j T_{\varepsilon t} - A_\varepsilon u_{\varepsilon t} - \tilde{B}(u_{\varepsilon t}, u_\varepsilon) - \tilde{B}(u_\varepsilon, u_{\varepsilon t})\} \in L^2([0, t_M], L^2(\Omega)^2).$$
$$A^{-\frac{1}{2}} T_{\varepsilon tt} = A^{-\frac{1}{2}} \{\lambda^{-1} A T_{\varepsilon t} - \tilde{B}(u_{\varepsilon t}, T_\varepsilon) - \tilde{B}(u_\varepsilon, T_{\varepsilon t})\} \in L^2([0, t_M], L^2(\Omega)).$$

Taking $v = t u_{\varepsilon tt}$ in (49), thanks to Lemma 2.1, (11) and Schwarz inequality, we obtain

$$\begin{aligned}
& t\|u_{\varepsilon tt}\|_0^2 + \frac{\nu}{2}\frac{d}{dt} t\|A_\varepsilon^{\frac{1}{2}} u_{\varepsilon t}\|_0^2 \\
&= \frac{\nu}{2}\|A_\varepsilon^{\frac{1}{2}} u_{\varepsilon t}\|_0^2 + t\lambda(jT_{\varepsilon t}, u_{\varepsilon tt}) - t\tilde{b}(u_{\varepsilon t}, u_\varepsilon, u_{\varepsilon tt}) - t\tilde{b}(u_\varepsilon, u_{\varepsilon t}, u_{\varepsilon tt}) \\
&\le \frac{\nu}{2}\|A_\varepsilon^{\frac{1}{2}} u_{\varepsilon t}\|_0^2 + Ct\|T_{\varepsilon t}\|_0^2 + \frac{t}{2}\|u_{\varepsilon tt}\|_0^2 + Ct\|u_\varepsilon\|_2^2\|A_\varepsilon^{\frac{1}{2}} u_{\varepsilon t}\|_0^2.
\end{aligned}$$

Taking $\psi = tT_{\varepsilon t}$ in Problem (IV), using (14) and Schwarz inequality, we obtain

$$\begin{aligned}
t\|T_{\varepsilon t}\|_0^2 + \frac{\lambda^{-1}}{2}\frac{d}{dt} t\|\nabla T_\varepsilon\|_0^2 &= \frac{\lambda^{-1}}{2}\|\nabla T_\varepsilon\|_0^2 - t\hat{b}(u_\varepsilon, T_\varepsilon, T_{\varepsilon t}) \\
&\le \frac{\lambda^{-1}}{2}\|\nabla T_\varepsilon\|_0^2 + \frac{t}{2}\|T_{\varepsilon t}\|_0^2 + Ct\|u_\varepsilon\|_2^2\|\nabla T_\varepsilon\|_0^2.
\end{aligned}$$

Integrating over $[0, t]$, using Lemma 2.4 and (50) to above two inequalities, we derive

$$t\|A_\varepsilon^{\frac{1}{2}} u_{\varepsilon t}\|_0^2 + \int_0^{t_M} s\|u_{\varepsilon tt}\|_0^2 dt \le C.$$

Taking $\psi = tT_{\varepsilon tt}$ in Problem (49), using Lemma 2.1, (14) and Schwarz inequality, we get

$$\begin{aligned}
& t\|T_{\varepsilon tt}\|_0^2 + \frac{\lambda^{-1}}{2}\frac{d}{dt} t\|A^{\frac{1}{2}} T_{\varepsilon t}\|_0^2 \\
&= \frac{\lambda^{-1}}{2}\|A^{\frac{1}{2}} T_{\varepsilon t}\|_0^2 - t\hat{b}(u_{\varepsilon t}, T_\varepsilon, T_{\varepsilon tt}) - t\hat{b}(u_\varepsilon, T_{\varepsilon t}, T_{\varepsilon tt}) \\
&\le \frac{\lambda^{-1}}{2}\|A^{\frac{1}{2}} T_{\varepsilon t}\|_0^2 + \frac{t}{2}\|T_{\varepsilon tt}\|_0^2 + Ct\|T_\varepsilon\|_2^2\|A_\varepsilon^{\frac{1}{2}} u_{\varepsilon t}\|_0^2 + Ct\|u_\varepsilon\|_2^2\|A^{\frac{1}{2}} T_{\varepsilon t}\|_0^2.
\end{aligned}$$

Integrating over $[0, t]$ and using Lemma 2.4, we derive

$$\int_0^{t_M} s\|T_{\varepsilon tt}\|_0^2 dt \le C.$$

The backward Euler time discretization scheme of the penalized system Problem (IV) is as follow

$$(52) \begin{cases} (\frac{1}{k}(u^{n+1} - u^n, v) + a_\varepsilon(u^{n+1}, v) + \tilde{b}(u^{n+1}, u^{n+1}, v) = \lambda(jT^{n+1}, v), \\ (\frac{1}{k}(T^{n+1} - T^n, \psi) + \bar{a}(T^{n+1}, \psi) + \hat{b}(u^{n+1}, T^{n+1}, \psi) = 0, \end{cases}$$

with $u^0 = 0$, $T^0 = \varphi(x)$.

Lemma 5.2. *Under the assumptions of (A1) and (A2), then, for any $0 \le n \le t_M/k$, it is valid that*

$$t_n\|u_\varepsilon(t_n) - u^n\|_0^2 + t_n^2\|\nabla(u_\varepsilon(t_n) - u^n)\|_0^2 \le Ck^2,$$
$$t_n\|T_\varepsilon(t_n) - T^n\|_0^2 + t_n^2\|\nabla(T_\varepsilon(t_n) - T^n)\|_0^2 \le Ck^2.$$

Proof. Letting $e^n = u_\varepsilon(t_n) - u^n$, $\delta^n = T_\varepsilon(t_n) - T^n$ and subtracting (52), from Problem(IV) at $t = t_{n+1}$, we get

$$(53)\quad \begin{cases} (\frac{1}{k}(e^{n+1} - e^n, v) + a_\varepsilon(e^{n+1}, v) + \tilde{b}(u^{n+1}, e^{n+1}, v) + \tilde{b}(e^{n+1}, u_\varepsilon(t_{n+1}), v) \\ = (R^n_{\varepsilon u}, v) + \lambda(j\delta^{n+1}, v), \\ (\frac{1}{k}(\delta^{n+1} - \delta^n, \psi) + \bar{a}(\delta^{n+1}, \psi) + \hat{b}(u^{n+1}, \delta^{n+1}, \psi) + \hat{b}(e^{n+1}, T_\varepsilon(t_{n+1}), \psi) \\ = (R^n_{\varepsilon T}, \psi), \end{cases}$$

where

$$(54)\quad R^n_{\varepsilon u} = u_{\varepsilon t}(t_{n+1}) - \frac{1}{k}(u_\varepsilon(t_{n+1}) - u_\varepsilon(t_n)) = \frac{1}{k}\int_{t_n}^{t_{n+1}} (t - t_n) u_{\varepsilon tt}(t) dt,$$

$$(55)\quad R^n_{\varepsilon T} = T_{\varepsilon t}(t_{n+1}) - \frac{1}{k}(T_\varepsilon(t_{n+1}) - T_\varepsilon(t_n)) = \frac{1}{k}\int_{t_n}^{t_{n+1}} (t - t_n) T_{\varepsilon tt}(t) dt.$$

Taking $(v, \psi) = 2k(e^{n+1}, \delta^{n+1})$ in (53), using (11), (14) and Schwarz inequality, we obtain

$$\begin{aligned} &\|e^{n+1}\|_0^2 - \|e^n\|_0^2 + \|e^{n+1} - e^n\|_0^2 + 2\nu k\|A_\varepsilon^{\frac{1}{2}} e^{n+1}\|_0^2 \\ &= 2k(R^n_{\varepsilon u}, e^{n+1}) - 2k\tilde{b}(e^{n+1}, u_\varepsilon(t_{n+1}), e^{n+1}) + 2\lambda k(j\delta^{n+1}, e^{n+1}) \\ &\le \nu k\|A_\varepsilon^{\frac{1}{2}} e^{n+1}\|_0^2 + Ck^2 \int_{t_n}^{t_{n+1}} \|A_\varepsilon^{-\frac{1}{2}} u_{\varepsilon tt}\|_0^2 dt \\ (56)\quad &\quad + Ck(\|u_\varepsilon(t_{n+1})\|_2^2 + 1)\|e^{n+1}\|_0^2 + k\|\delta^{n+1}\|_0^2. \end{aligned}$$

and

$$\begin{aligned} &\|\delta^{n+1}\|_0^2 - \|\delta^n\|_0^2 + \|\delta^{n+1} - \delta^n\|_0^2 + 2\lambda^{-1} k\|A^{\frac{1}{2}} \delta^{n+1}\|_0^2 \\ &= 2k(R^n_{\varepsilon T}, \delta^{n+1}) - 2k\hat{b}(e^{n+1}, T_\varepsilon(t_{n+1}), \delta^{n+1}) \\ (57)\quad &\le \lambda^{-1} k\|A^{\frac{1}{2}} \delta^{n+1}\|_0^2 + Ck^2 \int_{t_n}^{t_{n+1}} \|A^{-\frac{1}{2}} T_{\varepsilon tt}\|_0^2 dt + Ck\|T_\varepsilon(t_{n+1})\|_2^2 \|e^{n+1}\|_0^2. \end{aligned}$$

Taking the summation of (56), (57) for n from 0 to m, using Lemma 2.5 and Lemma 5.1, for all $0 \le m \le t_M/k - 1$, we have

$$(58)\quad \|e^{m+1}\|_0^2 + \|\delta^{m+1}\|_0^2 + \nu k\sum_{n=0}^{m} \|A_\varepsilon^{\frac{1}{2}} e^{n+1}\|_0^2 + \lambda^{-1} k\sum_{n=0}^{m} \|A^{\frac{1}{2}} \delta^{n+1}\|_0^2 \le Ck^2.$$

Thus

$$(59)\quad \|\nabla u^{m+1}\|_0 \le C, \quad \|\nabla T^{m+1}\|_0 \le C, \quad \forall m \le t_M/k - 1.$$

Taking $v = 2kt_{n+1}e^{n+1}$ in (53), using Lemma 2.1, (11) and Schwarz inequality, we obtain

$$\begin{aligned} &t_{n+1}(\|e^{n+1}\|_0^2 - \|e^n\|_0^2 + \|e^{n+1} - e^n\|_0^2) + 2\nu k t_{n+1}\|A_\varepsilon^{\frac{1}{2}} e^{n+1}\|_0^2 \\ &= 2kt_{n+1}(R^n_{\varepsilon u}, e^{n+1}) + 2kt_{n+1}\lambda(j\delta^{n+1}, e^{n+1}) \\ &\quad - 2kt_{n+1}\tilde{b}(e^{n+1}, u_\varepsilon(t_{n+1}), e^{n+1}) \\ &\le \nu k t_{n+1}\|A_\varepsilon^{\frac{1}{2}} e^{n+1}\|_0^2 + Ck^2 \int_{t_n}^{t_{n+1}} t\|u_{\varepsilon tt}\|_0^2 dt + kt_{n+1}\|\delta^{n+1}\|_0^2 \\ (60)\quad &\quad + Ckt_{n+1}(\|u_\varepsilon(t_{n+1})\|_2^2 + 1)\|e^{n+1}\|_0^2. \end{aligned}$$

Taking $\psi = 2kt_{n+1}\delta^{n+1}$ in (53), using (14) and Schwarz inequality, we get

$$\begin{aligned} & t_{n+1}(\|\delta^{n+1}\|_0^2 - \|\delta^n\|_0^2 + \|\delta^{n+1} - \delta^n\|_0^2) + 2\lambda^{-1}kt_{n+1}\|A^{\frac{1}{2}}\delta^{n+1}\|_0^2 \\ & = 2kt_{n+1}(R^n_{\varepsilon T}, \delta^{n+1}) - 2kt_{n+1}\hat{b}(e^{n+1}, T_\varepsilon(t_{n+1}), \delta^{n+1}) \\ & \le \lambda^{-1}kt_{n+1}\|A^{\frac{1}{2}}\delta^{n+1}\|_0^2 + Ck^2\int_{t_n}^{t_{n+1}} t\|T_{\varepsilon tt}\|_0^2 dt \\ & \quad + Ckt_{n+1}\|T_\varepsilon(t_{n+1})\|_2^2\|e^{n+1}\|_0^2. \end{aligned} \tag{61}$$

We known that

$$t_{n+1}(\|v^{n+1}\|_0^2 - \|v^n\|_0^2) = t_{n+1}\|v^{n+1}\|_0^2 - t_n\|v^n\|_0^2 - k\|v^n\|_0^2. \tag{62}$$

Taking the summation of (60), (61) for n from 0 to m, using (58), (62), Lemma 2.5 and Lemma 5.1, for all $0 \le m \le t_M/k - 1$, we derive

$$t_{m+1}(\|e^{m+1}\|_0^2 + \|\delta^{m+1}\|_0^2) + k\sum_{n=0}^{m} t_{n+1}(\nu\|A_\varepsilon^{\frac{1}{2}}e^{n+1}\|_0^2 + \lambda^{-1}\|A^{\frac{1}{2}}\delta^{n+1}\|_0^2) \le Ck^2.$$

Taking $v = 2kt_{n+1}^2 A_\varepsilon e^{n+1}$ in (53), using (10), (11), (58), (59), Young inequality and Lemma 2.1, we obtain

$$\begin{aligned} & t_{n+1}^2\{\|A_\varepsilon^{\frac{1}{2}}e^{n+1}\|_0^2 - \|A_\varepsilon^{\frac{1}{2}}e^n\|_0^2 + \|A_\varepsilon^{\frac{1}{2}}(e^{n+1} - e^n)\|_0^2\} + 2\nu kt_{n+1}^2\|A_\varepsilon e^{n+1}\|_0^2 \\ & = 2kt_{n+1}^2(R^n_{\varepsilon u}, A_\varepsilon e^{n+1}) - 2kt_{n+1}^2\tilde{b}(u^{n+1}, e^{n+1}, A_\varepsilon e^{n+1}) \\ & \quad - 2kt_{n+1}^2\tilde{b}(e^{n+1}, u_\varepsilon(t_{n+1}), A_\varepsilon e^{n+1}) + 2kt_{n+1}^2\lambda(j\delta^{n+1}, A_\varepsilon e^{n+1}) \\ & \le \nu kt_{n+1}^2\|A_\varepsilon e^{n+1}\|_0^2 + Ck^2\int_{t_n}^{t_{n+1}} t\|u_{\varepsilon tt}\|_0^2 dt + Ckt_{n+1}^2\|\delta^{n+1}\|_0^2 \\ & \quad + Ckt_{n+1}^2(\|u_\varepsilon(t_{n+1})\|_2^2 + 1)\|A_\varepsilon^{\frac{1}{2}}e^{n+1}\|_0^2. \end{aligned} \tag{63}$$

Then, taking $\psi = 2kt_{n+1}^2 A\delta^{n+1}$ in (53), using (13), (14), (59), Young inequality and Lemma 2.1, we obtain

$$\begin{aligned} & t_{n+1}^2\{\|A^{\frac{1}{2}}\delta^{n+1}\|_0^2 - \|A^{\frac{1}{2}}\delta^n\|_0^2 + \|A^{\frac{1}{2}}(\delta^{n+1} - \delta^n)\|_0^2\} + 2\lambda^{-1}kt_{n+1}^2\|A\delta^{n+1}\|_0^2 \\ & = 2kt_{n+1}^2\{(R^n_{\varepsilon T}, A\delta^{n+1}) - 2kt_{n+1}^2\hat{b}(u^{n+1}, \delta^{n+1}, A\delta^{n+1}) \\ & \quad - 2kt_{n+1}^2\hat{b}(e^{n+1}, T_\varepsilon(t_{n+1}), A\delta^{n+1})\} \\ & \le \lambda^{-1}kt_{n+1}^2\|A\delta^{n+1}\|_0^2 + Ck^2\int_{t_n}^{t_{n+1}} t\|T_{\varepsilon tt}\|_0^2 dt \\ & \quad + Ckt_{n+1}^2\|T_\varepsilon(t_{n+1})\|_2^2\|A_\varepsilon^{\frac{1}{2}}e^{n+1}\|_0^2 + Ckt_{n+1}^2\|A^{\frac{1}{2}}\delta^{n+1}\|_0^2. \end{aligned} \tag{64}$$

We known that

$$t_{n+1}^2(\|v^{n+1}\|_0^2 - \|v^n\|_0^2) = t_{n+1}^2\|v^{n+1}\|_0^2 - t_n^2\|v^n\|_0^2 - (2kt_n + k^2)\|v^n\|_0^2. \tag{65}$$

Taking the summation of (63), (64) for n from 0 to m, using (65), Lemma 2.5 and Lemma 5.1, for all $0 \le m \le t_M/k - 1$, we derive

$$t_{m+1}^2(\|A_\varepsilon^{\frac{1}{2}}e^{m+1}\|_0^2 + \|A^{\frac{1}{2}}\delta^{m+1}\|_0^2) + k\sum_{n=0}^{m} t_{n+1}^2(\nu\|A_\varepsilon e^{n+1}\|_0^2 + \lambda^{-1}\|A\delta^{n+1}\|_0^2) \le Ck^2.$$

Lemma 5.3. *Under the assumptions of (A1) and (A2), then, for any $0 \le m \le t_M/k$, it is valid that*

$$k\sum_{n=1}^{m} t_n^2\|p_\varepsilon(t_n) - p^n\|_0^2 \le Ck^2.$$

Proof. Taking $v = t_{n+1}^2(e^{n+1} - e^n)$ in (53), using Lemma 2.1, (11) and Schwarz inequality, we obtain

$$
\begin{aligned}
&kt_{n+1}^2\|\frac{e^{n+1}-e^n}{k}\|_0^2 + \nu t_{n+1}^2\{\|A_\varepsilon^{\frac{1}{2}}e^{n+1}\|_0^2 - \|A_\varepsilon^{\frac{1}{2}}e^n\|_0^2 + \|A_\varepsilon^{\frac{1}{2}}(e^{n+1}-e^n)\|_0^2\}\\
&= t_{n+1}^2(R_{\varepsilon u}^n, e^{n+1}-e^n) - t_{n+1}^2\tilde{b}(u^{n+1}, e^{n+1}, e^{n+1}-e^n)\\
&\quad - t_{n+1}^2\tilde{b}(e^{n+1}, u_\varepsilon(t_{n+1}), e^{n+1}-e^n) + \lambda t_{n+1}^2(j\delta^{n+1}, e^{n+1}-e^n)\\
&\le \frac{kt_{n+1}^2}{2}\|\frac{e^{n+1}-e^n}{k}\|_0^2 + Ck^2\int_{t_n}^{t_{n+1}} t\|u_{\varepsilon tt}\|_0^2 dt + Ckt_{n+1}^2\|A^{\frac{1}{2}}\delta^{n+1}\|_0^2\\
&\quad + Ckt_{n+1}^2\|\nabla u^{n+1}\|_0^2\|A_\varepsilon e^{n+1}\|_0^2 + Ckt_{n+1}^2\|u_\varepsilon(t_{n+1})\|_2^2\|A_\varepsilon^{\frac{1}{2}}e^{n+1}\|_0^2.
\end{aligned}
\tag{66}
$$

Taking the summation of (66) for n from 0 to m, using Lemma 2.5, (59) and Lemma 5.2, we derive

$$
k\sum_{n=1}^{m} t_{n+1}^2\|\frac{e^{n+1}-e^n}{k}\|_0^2 \le Ck^2, \quad \forall m \le t_M/k - 1. \tag{67}
$$

Then using the equation (53) and the available estimates for e^n and δ^n, we can prove

$$
k\sum_{n=1}^{m} t_n^2\|p_\varepsilon(t_n) - p^n\|_0^2 \le Ck^2, \quad \forall m \le t_M/k.
$$

Finally, by combining Theorem 4.3 with Lemma 5.2, Lemma 5.3, we obtain the following theorem.

Theorem 5.4. *Under the assumptions of (A1) and (A2), ε sufficiently small, then, for any $0 \le m \le t_M/k$, it is valid that*

$$
\begin{aligned}
t_m(\|u(t_m) - u^m\|_0^2 + \|T(t_m) - T^m\|_0^2) + t_m^2(\|\nabla(u(t_m) - u^m)\|_0^2 + \|\nabla(T(t_m) - T^m)\|_0^2)&\\
+k\sum_{n=1}^{m} t_n^2\|p(t_n) - p^n\|_0^2 \le C(k^2 + \varepsilon^2).&
\end{aligned}
$$

6. Conclusions

In this paper, we studied the penalty method for the two-dimensional unsteady conduction-convection problem under some realistically assumptions. By using the penalty method, we overcome the coupled problem and can efficiently sperate the computation of the velocity from that of the pressure. Optimal error estimate of the numerical velocity, pressure and temperature for the penalized system and the backward Euler scheme are derived, we will extend the present analysis to a fully discrete scheme by combining it with the finite element approximation results in our future work.

References

[1] R.A. Adams, Sobolev Spaces, Academic Press, New York, 1975.

[2] G.F. Carey and R. Krishnan, Penalty finite element method for the Navier-Stokes equations, Comput. Methods Appl. Mech. Engrg. 42(2)(1984) 183-224.

[3] G.F. Carey and J.T. Oden, Finite Elements: Fluid Mechanics, Vol. VI, Prentice-Hall, 1986.

[4] R.Courant, Variational methods for the solution of problems of equilibrium and vibrations, Bull. Amer. Math. Soc. 49(1943) 1-23.

[5] C. Cuvelier, A. Segal and A.A. Steenhoven, Finite Element Methods and Navier-Stokes Equations, D. Reidei Publishing Company, 1986.

[6] E. Dibenedetto and A. Friedman, Conduction-convection problems with change of phase, J. Differ. Equations, 62(2) (1986) 129-185.

[7] Y.N. He, Optimal error estimate of the penalty finite element method for the time-dipendent Navier-Stokes equations, Math. Comp. 74(251) (2005) 1201-1216.
[8] Y.N. He and W.W. Sun, Stability and convergence of the Crank-Nicolson/Adams-Bashforth scheme for the time-dependent Navier-Stokes equations, SIAM J. Numer. Anal. 45(2) (2007)837-869.
[9] J.G. Heywood and R. Rannacher, Finite element approximations of the nonstationary Navier-Stokes problem. Part I: Regularity of solutions and second-order spatial discretization, SIAM J. Numer. Anal. 19(2) (1982) 275-311.
[10] V. Girault and P.A. Raviart, Finite Element Method for Navier-Stokes equations, Springer-Verlag, Berlin, 1986.
[11] D.C. Kim and Y.D. Choi, Analysis of conduction-natural convection conjugate heat transfer in the gap between concentric cylinders under solar irradiation, Int. J. Therm. Sci. 48(6) (2009) 1247-1258.
[12] Z.D. Luo, The bases and applications of mixed finite element methods, Beijing Science Press, 2006. (in Chinese)
[13] Z.D. Luo, On the note of mixed finite element analysis for the conduction-convection problems, Acta Math. Appl. Sinica, 2(22) (1999) 261-270. (in Chinese)
[14] Z.D. Luo, J. Chen, I.M. Navon and J. Zhu, An optimizing reduced PLSMFE formulation for non-stationary conduction-convection problems, Int. J. Numer. Meth. Fluids, 60(4) (2009) 409-436.
[15] Z.D. Luo and X.M. Lu, A least squares Galerkin/Petrov mixed finite element method for the stationary conduction-convection problems, Math. Numer. Sinica, 25 (2) (2003) 231-244. (in Chinese)
[16] J. Shen, On error estimates of the penalty method for unsteady Navier-Stokes equations, SIAM J. Numer. Anal. 32(2) (1995) 386-403.
[17] J. Shen, On a new pseudocompressibility method for the incompressible Navier-Stokes equations, Appl. Numer. Math. 21(1) (1996) 71-90.
[18] D.Y. Shi and J.C. Ren, Nonconforming mixed finite element method for the stationary conduction-convection problem, Int. J. Numer. Anal. Mod. 6(2) (2009) 293-310.
[19] S.M. Shen, The finite element analysis for the conduction-convection problems, Math. Numer. Sinica, 2 (1994) 170-182. (in Chinese)
[20] H.Y. Sun, Y.N. He and X.L. Feng, On error estimates of the pressure-correction projection methods for the time-dependent Navier-Stokes equations, Int. J. Numer. Anal. Mod. 8(1) (2011) 70-85.
[21] R. Temam, Navier-Stokes equations: theory and numerical analysis(Third edition), North-Holland, Amsterdam, 1984.
[22] R. Temam, Une méthode d'approximation des solution des équations de Navier-Stokes, Bull. Soc. Math. France, 98(1968) 115-152.
[23] K. Wang, Y.N. He and X.L. Feng, On error estimates of the penalty method for the viscoelastic flow problem I: time discretization, Appl. Math. Model. 34(12) (2010) 4089-4105.
[24] K. Wang, Y.N. He and X.L. Feng, On error estimates of the fully discrete penalty method for the viscoelastic flow problem, Int. J. Comput. Math. 88(10) (2011) 2199-2220.
[25] T. Zhang and Y.N. He, Penalty finite element for the stationary conduction-convection problems, Submitted.
[26] T. Zhang, Penalty mixed finite element method for non-stationary conduction-convection problem, Submitted.
[27] Y. Zhang, M.F. Feng and Y.N. He, Subgrid model for the stationary incompressible Navier-Stokes equations based on the high order polynomial interpolation, Int. J. Numer. Anal. Mod. 7(4) (2010) 734-748.

Faculty of Science, Xi'an Jiaotong University, Xi'an 710049, P.R. China.
E-mail: `yanyan513@gmail.com`

Faculty of Science, Xi'an Jiaotong University, Xi'an 710049, P.R. China.
E-mail: `heyn@mail.xjtu.edu.cn`

College of Mathematics and Systems Science, Xinjiang University, Urumqi 830046, P.R. China.
E-mail: `fxlmath@gmail.com`