JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **165**, 180–191 (1992)

Finite Element Methods for Nonlinear Sobolev Equations with Nonlinear Boundary Conditions

YANPING LIN*

Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

AND

TIE ZHANG

Department of Mathematics, Northeast University of Technology, Shenyang, People's Republic of China

Submitted by Avner Friedman

Received February 6, 1990

We derive optimal L^2 error estimates for semi-discrete finite element methods for nonlinear Sobolev equations with nonlinear boundary conditions. A new projection is introduced and used in the error analysis. © 1992 Academic Press, Inc.

1. INTRODUCTION

Let $\Omega \subset R^d$ $(d \geqslant 1)$ be an open bounded domain with smooth boundary $\partial\Omega$. We consider finite element approximation to the solution of the following Sobolev equations:

$$\begin{aligned} c(u)\, u_t - \nabla \cdot \{a(u)\, \nabla u_t + b(u)\, \nabla u\} &= f(u), \qquad \text{in} \quad \Omega \times J, \\ a(u) \frac{\partial u_t}{\partial \mu} + b(u) \frac{\partial u}{\partial \mu} &= g(u), \qquad \text{on} \quad \partial\Omega \times J, \qquad (1.1) \\ u(\cdot, 0) &= v, \qquad \text{on} \quad \Omega \times \{0\}, \end{aligned}$$

where $J = (0, T)$, $T > 0$, a, b, c, f, g, and v are known functions; $\mu = (\mu_1, ..., \mu_d)$ denotes the outer-normal direction on $\partial\Omega$. We also assume that the functions a, b, c, f, and g are smooth with bounded derivatives and there exists $C_0 > 0$ such that

$$0 < C_0 \leqslant a(u), \qquad c(u), \quad u \in R.$$

* Present address: Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada T6G 2G1.

0022-247X/92 $3.00

The problem (1.1) can arise from many physical processes. For the questions of existence, uniqueness, and continuous dependence of the solutions and their applications, we refer to [2, 7] for extensive literature. For the numerical solution to Dirichlet boundary conditions, both finite difference and finite element methods have been studied by Ewing [5, 6], Ford [8], Ford and Ting [9, 10], and Wahlbin [15]. For problem (1.1), when $g = g(x, t)$ and $d \leqslant 3$, finite element methods with time-stepping have been investigated by Ewing [7]. In these papers the authors obtained quasi-optimal L^2 error estimates except that of [15]. While, Arnold, Douglas, and Thomee [1] and Nakao [12] considered finite element methods to a problem similar to (1.1) for $d = 1$ with periodic boundary conditions, they demonstrated some optimal error estimates and superconvergences. Some further investigations of finite element methods for Sobolev and other related type equations with homogeneous boundary data have been carried out by Lin, Thomee, and Wahlbin [11].

In this paper we study finite element approximation to the solution of (1.1) and show that the finite element solution u_h indeed possesses an optimal order convergence rate to the real solution u as expected. In addition, our results are valid for any space dimension $d \geqslant 1$ provided that u, $u_t \in H^s(\Omega)$ with $s > d/2 + 1$. It is well known that the restriction on $d \leqslant 3$ for parabolic equations [4] and Sobolev equations [7] cannot be removed because of the nonlinearity of the boundary data and the method used there. Hence, our method, at this point, has some advantages and provides, at least, some theoretical significances into the literature.

Let $H^s(\Omega)$ and $H^s(\partial\Omega)$ be Sobolev spaces of order s with norms $\|\cdot\|_s$ and $|\cdot|_s$, and $(\cdot, \cdot)$ and $\langle \cdot, \cdot \rangle$ denote the inner products in $L^2(\Omega)$ and $L^2(\partial\Omega)$, respectively.

Let S_h be a family of finite dimensional subspaces of $H^1(\Omega)$ such that for some $r \geqslant 2$,

$$\inf_{\chi \in S_h} \{\|u - \chi\| + h\|u - \chi\|_1\} \leqslant Ch^s \|u\|_s, \qquad 1 \leqslant s \leqslant r, \quad u \in H^s(\Omega).$$

We also assume that S_h is imposed on quasi-uniform triangulation of Ω such that the usual inverse inequalities hold [3, 7].

The semi-discrete finite element solution $u_h \colon \bar{J} \to S_h$ is now defined by

$$\begin{gathered}(c(u_h)\, u_{h,t}, \chi) + (a(u_h) \nabla u_{h,t} + b(u_h) \nabla u_h, \nabla \chi) \\ = \langle g(u_h), \chi \rangle + (f(u_h), \chi), \qquad \chi \in S_h \\ u_h(0) = v_h\end{gathered} \tag{1.2}$$

where v_h is an appropriate approximation of v into S_h.

Now let us state our main theorem in this paper.

THEOREM. *Let u and u_h be the solutions of* (1.1) and (1.2), respectively, *u, ∇u, u_t, and ∇u_t are bounded, and u, $u_t \in H^s(\Omega)$ with $s > d/2 + 1$. Then there exists $C = C(u) > 0$, dependent upon the norms of the solution u mentioned above, such that*

$$\|u(t) - u_h(t)\| + \|u_t(t) - u_{h,t}(t)\| \leqslant C(u)\, h^s, \tag{1.3}$$

provided that $v \in H^s(\Omega)$ and v_h is given by

$$(\nabla(v - v_h), \nabla\chi) + (v - v_h, \chi) = 0, \qquad \chi \in S_h. \tag{1.4}$$

In next sections we shall use the following inequalities:

$$AB \leqslant \varepsilon A^2 + B^2/4\varepsilon, \qquad \varepsilon > 0, \tag{1.5}$$

$$|u|^2_{L^2(\partial\Omega)} \leqslant \varepsilon \|\nabla u\|^2 + C(\varepsilon)\|u\|^2, \qquad \varepsilon > 0, \tag{1.6}$$

$$|u|_r \leqslant C_{T,r} \|u\|_{r+1/2}, \qquad 0 < r \leqslant \tfrac{3}{2}, r \neq 1. \tag{1.7}$$

In Section 2 we introduce a new projection W and study its approximation properties. The proof of the theorem will be given in Section 3.

2. A New Projection $W(t)$

In order to derive optimal error estimates for finite element approximations for diffusion type problems, it is well known [14, 16] that it is necessary to introduce Ritz projection into the analysis. The authors of [1, 7, 8, 12, 15] took this standard approach to treat the problems considered therein.

As a preparation of the proof for the theorem stated in Section 1, instead of using Ritz projection, we shall introduce an auxiliary function, called nonlinear nonclassical elliptic projection of the solution u into finite element spaces S_h, which is now defined in the following way: Let $\lambda \geqslant 0$ and $W(t)$: $[0, T] \to S_h$ be defined by

$$\begin{aligned} &(a(u)\nabla(W_t - u_t) + b(u)\nabla(W - u), \nabla\chi) + \lambda(W_t - u_t, \chi) \\ &\qquad - \langle g(W) - g(u), \chi\rangle = 0, \qquad t > 0, \chi \in S_h, \\ &W(0) = v_h \in S_h. \end{aligned} \tag{2.1}$$

We see, of course, that the study of this new projection will certainly take some extra effort, but optimal error estimates can be demonstrated in a very simple way as usual for parabolic equations [4, 14] if this projection $W(t)$ is used (see Section 3).

We show that $W(t)$ in (2.1) is well defined.

LEMMA 2.1. *For any $\lambda > 0$ there exists a unique $W(t) \in S_h$ satisfying* (2.1).

Proof. First, we shall prove that for any $v \in S_h$ there exists a unique $W_v(t) \in S_h$ such that

$$\begin{aligned}
&(a(u)\nabla(W_{v,t}-u_t)+b(u)\nabla(W_v-u),\nabla\chi)+\lambda(W_{v,t}-u_t,\chi)\\
&\qquad -\langle g(v)-g(u),\chi\rangle=0, \qquad t>0, \chi\in S_h, \qquad (2.2)\\
&W_v(0)=v_h.
\end{aligned}$$

For this purpose we let $S_h = \text{span}\{\psi_k\}_{k=0}^N$, where $\{\psi_k\}_{k=0}^N$ is a linearly independent set, and $W(t) = \sum_{k=1}^N C_k(t)\,\psi_k(x)$. Thus, we can write (2.2) as

$$\begin{aligned}
&(A(t)+\lambda D)\frac{d}{dt}C(t)+B(t)\,C(t)=C(t),\\
&C(0)=\text{determined by } v_h,
\end{aligned} \qquad (2.3)$$

where $A(t)$, $B(t)$, and D are matrices,

$$A(t)=((a(u)\nabla\psi_k,\nabla\psi_l)), \qquad B(t)=((b(u)\nabla\psi_k,\nabla\psi_l)), \qquad D=(\psi_k,\psi_l),$$

and C and F are vectors,

$$\begin{aligned}
&C(t)=(C(t)_1,\ldots,C(t)_N)^{\mathrm{T}}, \qquad F(t)=(F(t)_1,\ldots,F(t)_N)^{\mathrm{T}},\\
&F(t)_l=(a(u)\nabla u_t+b(u)\nabla u,\nabla\psi_l)+\langle g(v)-g(u),\psi_l\rangle, \qquad l=1,\ldots,N.
\end{aligned}$$

It follows from our assumptions that $A(t)+\lambda D$ is positive definite, and we find from general theory of initial value problems of ordinary differential equations that there exists a unique $C(t)$ in (2.3). Consequently, $W_v(t)$ is well defined in (2.2) for any $v \in S_h$.

Now let $W^0 = v_h$ and $\{W^m\}$ be defined by

$$\begin{aligned}
&(a(u)\nabla(W_t^{m+1}-u_t)+b(u)\nabla(W^{m+1}-u),\nabla\chi)+\lambda(W_t^{m+1}-u_t,\chi)\\
&\qquad -\langle g(W^m)-g(u),\chi\rangle=0, \qquad t>0, \chi\in S_h, \qquad (2.4)\\
&W^{m+1}(0)=v_h, \qquad m=0,1,2,\ldots.
\end{aligned}$$

We know from the above that $\{W^m\}$ is well defined. It we set $Z^m = W^{m+1} - W^m$, we obtain from (2.4) that

$$\begin{aligned}
&(a(u)\nabla Z_t^m+b(u)\nabla Z^m,\nabla\chi)+\lambda(Z_t^m,\chi)-\langle g(W^m)\\
&\qquad -g(W^{m-1}),\chi\rangle=0, \qquad t>0, \chi\in S_h, \qquad (2.5)\\
&Z^m(0)=0, \qquad m=1,2,\ldots.
\end{aligned}$$

By setting $\chi = Z_t^m \in S_h$, it is easy to see from (2.5) and (1.5)–(1.6) that

$$\begin{aligned}
&C_0 \|\nabla Z_t^m\|^2 + \lambda \|Z_t^m\|^2 + \lambda \|Z_t^m\|^2 \\
&\leqslant C \|\nabla Z^m\| \, \|\nabla Z_t^m\| + C \|Z^{m-1}\|_{L^2(\partial\Omega)} \|Z_t^m\|_{L^2(\partial\Omega)} \\
&\leqslant \frac{C_0}{4} \|\nabla Z_t^m\|^2 + C \|\nabla Z^m\|^2 + \frac{C_0}{4} \|\nabla Z_t^m\|^2 + \frac{\lambda}{2} \|Z_t^m\|^2 + C \|Z^{m-1}\|_1^2 \\
&\leqslant \frac{C_0}{2} \|\nabla Z_t^m\|^2 + \frac{\lambda}{2} \|Z_t^m\|^2 + C \int_0^t (\|Z_t^m\|_1^2 + \|Z_t^{m-1}\|_1^2) \, d\tau. \qquad (2.6)
\end{aligned}$$

Here we have used $Z^m(0) = 0$ and

$$\|Z^m\|_1^2 \leqslant C \int_0^t \|Z_t^m(\tau)\|_1^2 \, d\tau.$$

Thus, (2.6) and Gronwall's lemma imply that

$$\|Z_t^m\|_1^2 \leqslant C \int_0^t \|Z_t^{m-1}\|_1^2 \, d\tau \leqslant \cdots \leqslant \frac{(CT)^m}{m!}, \qquad m \geqslant 1.$$

Here we have assumed that $\|Z_t^0(t)\|_1$ is uniformly bounded in $[0, T]$. In fact, it can be seen easily from (2.4) and an argument similar to the above that

$$\|W_t^m\|_1^2 \leqslant C(u) + C \int_0^t \|W_t^{m-1}(\tau)\|_1^2 \, d\tau, \qquad m \geqslant 1.$$

Since $W^0 = v_h$ it follows that $\|W_t^1(t)\|_1$ is uniformly bounded and so is $\|Z_t^0(t)\|_1$ by its definition and the above inequality. Hence, $\{W^m\}$ is a Cauchy sequence in S_h, so there exists a unique $W(t) \in S_h$ such that $W^m \to W$ and $W_t^m \to W_t$ in S_h. Lemma 2.1 is now complete by letting $m \to \infty$ in (2.4). Q.E.D.

LEMMA 2.2. *Assume that* $\|v - v_h\| + h\|v - v_h\|_1 \leqslant Ch^s \|v\|_s$, u, $u_t \in H^s(\Omega)$ *with* $s > d/2 + 1$, *u and* ∇u *bounded. Then there exists a* $C = C(u) > 0$ *such that*

$$\|\nabla(W - u)\| + \|\nabla(W_t - u_t)\| \leqslant Ch^{s-1}(\|u\|_{1,s} + \|v\|_s), \qquad (2.7)$$

where

$$\|u\|_{k,s}^2 = \sum_{j=0}^{k} \left\{ \|D_t^j u(t)\|_s^2 + \int_0^t \|D_t^j u(\tau)\|_s^2 \, d\tau \right\}, \qquad k \geqslant 0, \quad t \in \bar{J}.$$

Proof. Let $\eta = W - u$ and $P_h \colon L^2(\Omega) \to S_h$ be L^2 projection [3, 14]. We know from (2.1), (1.5)–(1.7), and [3, 14] that

$$\begin{aligned}
&(a(u)\nabla\eta_t + b(u)\nabla\eta, \nabla\eta_t) + \lambda(\eta_t, \eta_t) - \langle g(W) - g(u), \eta_t\rangle \\
&\quad \leqslant (a(u)\nabla\eta_t + b(u)\nabla\eta, \nabla(u - P_h u)_t) \\
&\qquad + \lambda(\eta_t, (u - P_h u)_t) - \langle g(W) - g(u), (u - P_h u)_t\rangle \\
&\quad \leqslant C(\|\nabla\eta_t\| + \|\nabla\eta\|)\|\nabla(u - P_h u)_t)\| + C\|\eta_t\|\,\|(u - P_h u)_t\| \\
&\qquad + C\|\eta\|_{L^2(\partial\Omega)}\|(u - P_h u)_t\|_{L^2(\partial\Omega)} \\
&\quad \leqslant \frac{C_0}{4}\|\nabla\eta_t\|^2 + \frac{\lambda}{4}\|\eta_t\|^2 + C\|\eta\|_1^2 + Ch^{2s-2}(\|u_t\|_s^2 + \|u\|_s^2),
\end{aligned}$$

and then,

$$\begin{aligned}
&C_0\|\nabla\eta_t\|^2 + \lambda\|\eta_t\|^2 \\
&\quad \leqslant \frac{C_0}{2}\|\nabla\eta_t\|^2 + \frac{\lambda}{2}\|\eta_t\|^2 + C\|\eta\|_1^2 + Ch^{2s-2}(\|u\|_s^2 + \|u_t\|_s^2) \\
&\quad \leqslant \frac{C_0}{2}\|\nabla\eta_t\|^2 + \frac{\lambda}{2}\|\eta_t\|^2 + C(\|\nabla\eta(0)\|^2 + \|\eta(0)\|^2) \\
&\qquad + Ch^{2s-2}(\|u\|_s^2 + \|u_t\|_s^2) + C\int_0^t \|\eta_t\|_1^2\,d\tau.
\end{aligned}$$

From Gronwall's lemma and our assumptions on v_h we obtain

$$\|\eta_t\|_1^2 \leqslant Ch^{2s-2}(\|u\|_{1,s}^2 + \|v\|_s^2). \tag{2.8}$$

Hence, Lemma 2.2 follows from

$$\|\eta\|_1^2 \leqslant C\|\eta(0)\|_1^2 + C\int_0^t \|\eta_t\|_1^2\,d\tau. \qquad \text{Q.E.D.} \tag{2.9}$$

We now turn to an estimate for $\eta = W - u$ in $L^2(\Omega)$.

LEMMA 2.3. *Under assumptions of Lemma 2.2 and* $|v - v_h|_{-1/2} \leqslant Ch^s\|v\|_s$. *We have*

$$\|\eta\| + \|\eta_t\| \leqslant Ch^s(\|u\|_{1,s} + \|v\|_s). \tag{2.10}$$

Proof. Let $\alpha \in H^1(\Omega)$ such that

$$(a(u)\nabla\alpha + b(u)\nabla\eta, \nabla V) + \lambda(\alpha, V) - \langle G\eta, V\rangle = 0, \qquad V \in H^1(\Omega), \tag{2.11}$$

where

$$G=\int_0^1 \frac{dg}{du}(\theta u+(1-\theta)W)\,d\theta.$$

The existence and uniqueness of such $\alpha\in H^1(\Omega)$ is guaranteed by taking λ large enough. We shall, from now on, assume that λ is big enough (fixed) for all our needs. We see from (2.1) and (2.11) that

$$(a(u)\nabla(\eta_t-\alpha),\nabla\chi)+\lambda(\eta_t-\alpha,\chi)=0,\qquad \chi\in S_h. \tag{2.12}$$

Noticing that $\eta_t-\alpha=W_t-u_t-\alpha=W_t-(u_t+\alpha)$, we see that W_t is actually a standard Ritz projection of $u_t+\alpha$ into S_h, so that we have from Lemma 2.2 and [3, 14, 16],

$$\begin{aligned}\|W_t-u_t-\alpha\| &\leqslant Ch\,\|W_t-u_t-\alpha\|_1\leqslant Ch(\|\eta_t\|_1+\|\alpha\|_1)\\ &\leqslant Ch^s(\|u\|_{1,s}+\|v\|_s)+Ch\,\|\alpha\|_1.\end{aligned} \tag{2.13}$$

Also, we see that the elliptic regularity and (2.11) imply

$$\|\alpha\|_1\leqslant C\,\|\nabla\eta\|+C\,\|\eta\|_{L^2(\partial\Omega)}\leqslant C\,\|\eta\|_1\leqslant Ch^{s-1}(\|u\|_{1,s}+\|v\|_s),$$

and hence,

$$\|\eta_t-\alpha\|\leqslant Ch^s(\|u\|_{1,s}+\|v\|_s). \tag{2.14}$$

It remains now to estimate $\|\alpha\|$. Let $\beta\in H^1(\Omega)$ be defined by

$$(a(u)\nabla\beta+b(u)\nabla\eta,\nabla V)+\lambda(\beta,V)-\langle G\eta,V\rangle=(\alpha,V),\qquad V\in H^1(\Omega). \tag{2.15}$$

If we set $V=\alpha$ in (2.15) and $V=\beta$ in (2.11), we obtain from integration by parts, (1.5)–(1.7), and our assumptions that

$$\begin{aligned}\|\alpha\|^2 &= (a(u)\nabla\beta+b(u)\nabla\eta,\nabla\alpha)+\lambda(\beta,\alpha)-\langle G\eta,\alpha\rangle\\ &=(b(u)\nabla\eta,\nabla(\alpha-\beta))-\langle G\eta,\alpha-\beta\rangle\\ &=\left\langle \eta,\frac{\partial}{\partial\mu}b(\alpha-b)\right\rangle-(\eta,\nabla\cdot b\nabla(\alpha-\beta))-\langle G\eta,\alpha-\beta\rangle\\ &\leqslant C\,|\eta|^2_{-1/2}+\varepsilon\,|\alpha-\beta|^2_{3/2}+C\,\|\eta\|^2\\ &\quad+\varepsilon\,\|\alpha-\beta\|^2_2+C\,|\eta|^2_{-1/2}+\varepsilon\,|\alpha-\beta|^2_{1/2}\\ &\leqslant C\,|\eta|^2_{-1/2}+C\,\|\eta\|^2+C\varepsilon\,\|\alpha-\beta\|^2_2\\ &\leqslant \varepsilon C\,\|\alpha-\beta\|^2_2+C\,\|\eta(0)\|^2+C\,|\eta(0)|^2_{-1/2}+C\int_0^t(\|\eta_t\|^2+|\eta_t|^2_{-1/2})\,d\tau\\ &\leqslant C\varepsilon\,\|\alpha-\beta\|^2_2+Ch^{2s}\,\|v\|^2_s+C\int_0^t(\|\eta_t\|^2+|\eta_t|^2_{-1/2})\,d\tau.\end{aligned} \tag{2.16}$$

We obtain by subtracting (2.15) from (2.11) that

$$(a(u)\nabla(\alpha-\beta), \nabla V)+\lambda(\alpha-\beta, V)=(\alpha, V), \qquad V\in H^1(\Omega),$$

and, thus, $\|\alpha-\beta\|_2\leqslant C\|\alpha\|$. It is now easy to see by taking ε small (fixed) in (2.16) that

$$\|\alpha\|^2\leqslant Ch^{2s}(\|u\|_{1,s}^2+\|v\|_s^2)+C\int_0^t(\|\eta_t\|^2+|\eta_t|_{-1/2}^2)\,d\tau. \tag{2.17}$$

By (2.14) and (2.17) we see that we need to estimate $|\eta_t|_{-1/2}$. For this purpose we let $\gamma\in H^1(\Omega)$ be such that

$$(a(u)\nabla\gamma+b(u)\nabla\eta, \nabla V)+\lambda(\gamma, V)-\langle G\eta, V\rangle=\langle\delta, V\rangle, \qquad V\in H^1(\Omega), \tag{2.18}$$

where $\delta\in H^{1/2}(\partial\Omega)$ and $|\delta|_{1/2}=|\eta_t|_{-1/2}$, $\langle\delta, \eta_t\rangle=|\eta|_{-1/2}^2$. Setting $\Phi=\gamma-\alpha$, it follows from (2.11) and (2.18) that

$$(a(u)\nabla\Phi, \nabla V)+\lambda(\Phi, V)=\langle\delta, V\rangle, \qquad V\in H^1(\Omega) \tag{2.19}$$

and, then, the elliptic regularity implies

$$\|\Phi\|_2\leqslant C|\delta|_{1/2}=C|\eta_t|_{-1/2}. \tag{2.20}$$

Now let $V=\eta_t$ in (2.19) together with integration by parts and (1.5)–(1.7). We obtain

$$\begin{aligned}
|\eta|_{-1/2}^2&=(a(u)\nabla\Phi, \nabla\eta_t)+\lambda(\Phi, \eta_t)=(a(u)\nabla\eta_t, \nabla\Phi)+\lambda(\eta_t, \Phi)\\
&=(a(u)\nabla\eta_t+b(u)\nabla\eta, \nabla\Phi)+\lambda(\eta_t, \Phi)\\
&\quad-\langle G\eta, \Phi\rangle+\langle G\eta, \Phi\rangle-(b(u)\nabla\eta, \nabla\Phi)\\
&=(a(u)\nabla\eta_t+b(u)\nabla\eta, \nabla(\Phi-I_h\Phi))\\
&\quad+\lambda(\eta_t, \Phi-I_h\Phi)-\langle G\eta, \Phi-I_h\Phi\rangle\\
&\quad+\langle G\eta, \Phi\rangle-\left\langle\eta, \frac{\partial}{\partial\mu}b(u)\Phi\right\rangle+(\eta, \nabla\cdot b(u)\nabla\Phi)\\
&\leqslant\frac{C}{\varepsilon}h(\|\eta_t\|_1^2+\|\eta\|_1^2)+\varepsilon\|\Phi\|_2^2+\frac{C}{\varepsilon}|\eta|_{-1/2}^2\\
&\quad+\varepsilon(|\Phi-I_h\Phi|_{3/2}^2+|\Phi|_{1/2})+C\|\eta\|\,\|\Phi\|_2\\
&\leqslant\frac{C}{\varepsilon}h^s(\|u\|_{1,s}^2+\|v\|_s^2)+C\varepsilon\|\Phi\|_2^2+\frac{C}{\varepsilon}\int_0^t|\eta_t|_{-1/2}^2\,d\tau.
\end{aligned} \tag{2.21}$$

Taking ε small (fixed), we find from (2.20) and (2.21) that

$$|\eta_t|_{-1/2}^2 \leqslant Ch^{2s}(\|u\|_{1,s}^2 + \|v\|_s^2) + C\int_0^t |\eta_t|_{-1/2}^2 \, d\tau.$$

Applying Gronwall's lemma to above inequality, we see

$$|\eta_t|_{-1/2}^2 \leqslant Ch^{2s}(\|u\|_{1,2}^2 + \|v\|_s^2). \tag{2.22}$$

Using (2.22), (2.17), and (2.14), we see again from Gronwall's lemma that

$$\|\eta_t\|^2 \leqslant Ch^{2s}(\|u\|_{1,s}^2 + \|v\|_s^2) \tag{2.23}$$

and, then,

$$\|\eta\|^2 \leqslant C\|\eta(0)\|^2 + C\int_0^t \|\eta_t\|^2 \, d\tau \leqslant Ch^{2s}(\|u\|_{1,s}^2 + \|v\|_s^2). \quad \text{Q.E.D.} \tag{2.24}$$

In Lemma 2.3 we assumed that $|v - v_h|_{-1/2} \leqslant Ch^s \|v\|_s$, this is not valid in general. Thus we are required to approximate our initial data v in a suitable way so that this assumption holds.

LEMMA 2.4. *If v_h is defined by* (1.4), *then we have for some $C > 0$,*

$$|v - v_h|_{-1/2} < Ch^s \|v\|_s. \tag{2.25}$$

Proof. Since

$$|v - v_h|_{-1/2} = \sup\{\langle v - v_h, \phi\rangle \mid |\phi|_{1/2} = 1, \phi \in H^{1/2}(\partial\Omega)\}.$$

If we let $\psi \in H^2(\Omega)$ such that

$$-\nabla^2\psi + \psi = 0, \text{ in } \Omega, \qquad \frac{\partial\psi}{\partial\mu} = \phi, \text{ on } \partial\Omega,$$

then we see that

$$\begin{aligned}\langle v - v_h, \phi\rangle &= (\nabla(v - v_h), \nabla\psi) + (v - v_h, \psi)\\ &= (\nabla(v - v_h), \nabla(\phi - I_h\psi)) + (v - v_h, \psi - I_h\psi)\\ &\leqslant Ch^{s-1}\|v\|_s h\|\psi\|_2 \leqslant Ch^s\|v\|_s |\phi|_{1/2}.\end{aligned}$$

Thus, we complete Lemma 2.4. Q.E.D.

LEMMA 2.5. *If $u, u_t \in H^s(\Omega)$ with $s > d/2 + 1$ and ∇u, ∇u_t are bounded, then we have*

$$\|W\|_\infty + \|\nabla W\|_\infty + \|W_t\|_\infty + \|\nabla W_t\|_\infty \leqslant C(u). \tag{2.26}$$

Proof. Let $R_h u$ be the Ritz projection of u into S_h; we then see from Lemma 2.2 and [3, 13, 14] that

$$\begin{aligned}\|\nabla W\|_\infty &\leqslant \|\nabla R_h u\|_\infty + \|\nabla(R_h u - W)\|_\infty \\ &\leqslant C(u) + Ch^{-d/2}\|\nabla(R_h u - W)\| \\ &\leqslant C(u) + Ch^{-d/2}(\|\nabla(R_h u - u)\| + \|\nabla(u - W)\|) \\ &\leqslant C(u) + Ch^{-d/2}h^{s-1}\|u\|_{1,s} \leqslant C(u).\end{aligned}$$

The remainder of the proofs is similar to the above, so we omit it. Q.E.D.

3. Proof of Main Theorem

In this section we shall prove our main theorem stated in Section 1 by using nonclassical projection W introduced and studied in Section 2. By definition of W in (2.1) we see that

$$\begin{aligned}&(c(u)\,W_t, \chi) + (a(u)\,\nabla W_t + b(u)\,\nabla W, \nabla\chi) - \langle g(W), \chi\rangle \\ &\quad = -((\lambda + c(u))\,\eta_t, \chi) + (f(u), \chi), \qquad \chi \in S_h, \qquad (3.1)\end{aligned}$$

If we write the error $u - u_h = (u - W) + (W - u_h) = \eta + \theta$, we know from Lemma 2.2 and Lemma 2.3 that

$$\|\eta\| + \|\eta_t\| \leqslant Ch^s(\|u\|_{1,s} + \|v\|_s). \qquad (3.2)$$

Thus, it remains to estimate θ only. We see from (1.2) and (3.1) that θ satisfies

$$\begin{aligned}&(c(u_h)\,\theta_t, \chi) + (a(u_h)\,\nabla\theta_t + b(u_h)\,\nabla\theta, \nabla\chi) - \langle G^*\theta, \chi\rangle \\ &\quad = -((a(u) - a(u_h))\,\nabla W_t + (b(u) - b(u_h))\,\nabla W, \nabla\chi) \\ &\qquad - ((c(u) - c(u_h))\,W_t + (\lambda + c(u))\,\eta_t, \chi) + (f(u) - f(u_h), \chi), \quad \chi \in S_h, \qquad (3.3)\end{aligned}$$

where

$$G^* = \int_0^1 \frac{dg}{du}(\xi W + (1 - \xi)\,u_h)\,d\xi,$$

which is bounded from our assumption on g. Letting $\chi = \theta_t$ and using Lemma 2.5, we obtain

$$\begin{aligned}&C_0\|\theta_t\|_1^2 + (b(u_h)\,\nabla\theta, \nabla\theta_t) - \langle G^*\theta, \theta_t\rangle \\ &\quad \leqslant C\|\theta + \eta\|(\|\nabla\theta_t\| + \|\theta_t\|) + C\|\eta_t\|\,\|\theta\| \\ &\quad \leqslant \frac{C_0}{4}\|\theta_t\|_1^2 + C(\|\theta\|^2 + \|\eta_t\|^2 + \|\eta\|^2). \qquad (3.4)\end{aligned}$$

Also, by (1.5)–(1.7),

$$|(b(u_h \nabla\theta, \nabla\theta_t) - \langle G^*\theta, \theta_t\rangle| \leqslant \frac{C_0}{4} \|\theta_t\|_1^2 + C\|\theta\|_1^2. \tag{3.5}$$

Combining (3.2)–(3.5) we find

$$\|\theta_t\|_1^2 \leqslant C\|\theta\|_1^2 + C\|\eta\|^2 + \|\eta_t\|^2 \leqslant Ch^{2s}(\|u\|_{1,s}^2 + \|v\|_s^2) + C\int_0^t \|\theta_t\|_1^2\, d\tau. \tag{3.6}$$

Here we have used $\theta(0) = W(0) - v_h = 0$. Hence Gronwall's lemma yields

$$\|\theta_t\|_1^2 \leqslant Ch^{2s}(\|u\|_{1,s}^2 + \|v\|_s^2), \tag{3.7}$$

so that

$$\|\theta\|_1^2 \leqslant C\int_0^t \|\theta_t\|_1^2\, d\tau \leqslant Ch^{2s}(\|u\|_{1,s}^2 + \|v\|_s^2). \tag{3.8}$$

Hence, we have from (3.7)–(3.8), (3.2), and triangle inequality that

$$\|u - u_h\| + \|u_t - u_{h,t}\| \leqslant Ch^s(\|u\|_{1,s} + \|v\|_s). \quad \text{Q.E.D.} \tag{3.9}$$

Remark. As mentioned in the Section 1 we assumed that $s > 1 + d/2$ which is needed in Lemma 2.5 in order to bound W etc. [16]. In the case of linear equations, i.e., $a = a(x, t)$, $b = b(x, t)$, and $c = c(x, t)$, the restriction on $s > 1 + d/2$ can be removed, since Lemma 2.5 will be no longer necessary in our error estimates.

Acknowledgments

The authors thank Professors Alfred H. Schatz and Lars B. Wahlbin for many comments and criticism during the preparation of this work. The authors also thank the referee for comments and suggestions which led to the improvement of the manuscript.

References

1. D. N. Arnold, J. Douglas, Jr., and V. Thomee, Superconvergence of a finite element approximation to the solution of a Sobolev equation in a single space variable, *Math. Comp.* **36** (1981), 53–63.
2. R. W. Carroll and R. E. Showalter, "Singular and Degenerate Cauchy Problems," Mathematics in Sciences and Engineering, Vol. 127, Academic Press, New York, 1976.
3. P. G. Ciarlet, "The Finite Element Methods for Elliptic Problems," North Holland, Amsterdam, 1978.

4. J. DOUGLAS, JR. AND T. DUPONT, Galerkin methods for parabolic equations with nonlinear boundary conditions, *Numer. Math.* **20** (1973), 213–237.
5. R. E. EWING, The approximation of certain parabolic equations backward in time by Sobolev equations, *SIAM J. Math. Anal.* **6** (1975), 283–294.
6. R. E. EWING, Numerical solution of Sobolev partial differential equations, *SIAM J. Numer. Anal.* **12** (1975), 345–363.
7. R. E. EWING, Time-stepping Galerkin methods for nonlinear Sobolev partial differential equations, *SIAM J. Numer. Anal.* **15** (1978), 1125–1150.
8. W. H. FORD, Galerkin approximation to nonlinear pseudoparabolic partial differential equation, *Aequationes Math.* **14** (1976), 271–291.
9. W. H. FORD AND T. W. TING, Stability and convergence of difference approximations to pseudoparabolic partial equations, *Math. Comp.* **27** (1973), 737–743.
10. W. H. FORD AND T. W. TING, Uniform error estimates for difference approximations to nonlinear pseudoparabolic partial differential equations, *SIAM J. Numer.* **11** (1974), 155–169.
11. Y. LIN, V. THOMEE, AND L. WAHLBIN, Rizt–Volterra projection on finite element spaces and applications to integro-differential and related equations, in preparation.
12. M. T. NAKAO, Error estimates of a Galerkin method for some nonlinear Sobolev equations in one space dimension, *Numer. Math.* **47** (1985), 139–157.
13. R. RANNACHER AND R. SCOTT, Some optimal error estimates for piecewise linear finite element approximations, *Math. Comp.* **38** (1982), 1–22.
14. V. THOMEE, "Galerkin Finite Element Methods for Parabolic Equations," Lecture Notes in Mathematics, Vol. 1054, Springer-Verlag, New York/Berlin, 1984.
15. L. WAHLBIN, Error estimates for a Galerkin method for a class of model equations for long waves, *Numer. Math.* **23** (1975), 289–303.
16. M. F. WHEELER, A priori L_2 error estimates for Galerkin approximation to parabolic partial differential equations, *SIAM J. Numer. Anal.* **10** (1973), 723–759.