JOURNAL OF INTEGRAL EQUATIONS
AND APPLICATIONS
Volume 10, Number 1, Spring 1998

SEMI-DISCRETE FINITE ELEMENT APPROXIMATIONS FOR LINEAR PARABOLIC INTEGRO-DIFFERENTIAL EQUATIONS WITH INTEGRABLE KERNELS

YANPING LIN

ABSTRACT. In this paper we consider finite element methods for general parabolic integro-differential equations with integrable kernels. A new approach is taken, which allows us to derive optimal L^p, $2 \leq p \leq \infty$, error estimates and superconvergence. The main advantage of our method is that the semi-discrete finite element approximations for linear equations, with both smooth and integrable kernels, can be treated in the same way without the introduction of the Ritz-Volterra projection; therefore, one can make full use of the results of finite element approximations for elliptic problems.

1. Introduction. In this paper we study numerical solutions by finite element methods for the following parabolic integro-differential equation:

$$\begin{cases} u_t + Au = \int_0^t a(t-s)Bu(s)\,ds + f(t) & \text{in } \Omega \times J, \\ u = 0 & \text{on } \partial\Omega \times J, \\ u(\cdot, 0) = v & \text{on } \Omega, \end{cases} \tag{1.1}$$

where $\Omega \subset R^d$, $d \geq 1$, is a bounded domain with smooth boundary $\partial\Omega$, $J = (0, T_0]$, $T_0 > 0$, $a(t) \in L^1(J)$ an integrable kernel, f and v are known smooth functions. A is a positive definite second order elliptic operator,

$$A(t) = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial}{\partial x_j}\right) + a(x)I, \quad a(x) \geq 0,$$

$$a_{ij}(x) = a_{ji}(x), \quad i, j = 1, \dots, d,$$

$$\sum_{i,j=1}^{d} a_{ij}\xi_i\xi_j \geq C_0 \sum_{i=1}^{d} \xi_i^2, \quad C_0 > 0,$$

Received by the editors on June 15, 1995, and in revised form on June 1, 1997.
AMS (MOS) *Mathematics Subject Classification.* 65N30, 45K05.
Key words and phrases. Integrable kernel, finite element, error estimates, maximum norm, superconvergence, parabolic, integro-differential.
This work is supported in part by NSERC (Canada).

and B is any second order operator,

$$B = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(b_{ij}(x)\frac{\partial}{\partial x_j}\right) + \sum_{i=1}^{d} b_i(x)\frac{\partial}{\partial x_i} + b(x)I,$$

with smooth coefficients in x.

In a mathematical model describing the heat flow through a body, very often one has to take some memory effect into consideration. The common feature of such a model consists of introducing some relaxation function into the constitutive relations in order to represent the memory effect. For example, a quite general constitutive assumptions for a homogeneous and isotropic body $\Omega \subset R^n$, $n = 1, 2, 3$ in the applications, is the following [**28**]:

$$e(x,t) = \beta(u(x,t)) + \int_0^\infty h(s)\gamma(u(x,t-s))\,ds, \quad x \in \Omega, \quad t \geq 0,$$

$$q(x,t) = -\rho(\nabla u(x,t)) - \int_0^\infty k(s)\mu(\nabla u(x,t-s))\,ds, \quad x \in \Omega, \quad t \geq 0,$$

where u denotes the body temperature, e and q denote the internal energy and the heat flux, respectively, β, γ, ρ and μ are given functions satisfying certain assumptions and h and k are the internal energy and the heat flux relaxation functions, respectively, representing for the memory effects.

The balance law of the heat energy implies

$$\frac{\partial e}{\partial t}(x,t) + \operatorname{div} q(x,t) = f(x,t), \quad x \in \Omega, t \geq 0,$$

where div is the divergence operator in R^n and f denotes the source. Upon using the constitutive relations, it follows that u satisfies the following partial integro-differential equation:

$$\frac{\partial}{\partial t}\left\{\beta(u(x,t)) + \int_0^\infty h(s)\gamma(u(x,t-s))\,ds\right\} = \operatorname{div}\left\{\rho(\nabla u(x,t)) - \int_0^\infty k(s)\mu(\nabla u(x,t-s))\,ds\right\} + f(x,t).$$

In application it is assumed that the thermal history $u(x,t)$ is known up to $t=0$, then the above equation can be written into the following Volterra parabolic integro-differential equation:

$$\frac{\partial}{\partial t}\left\{\beta(u(x,t))+\int_0^t h(s)\gamma(u(x,t-s))\,ds\right\}$$
$$=\operatorname{div}\left\{\rho(\nabla u(x,t))-\int_0^t k(s)\mu(\nabla u(x,t-s))\,ds\right\}+F(x,t),$$

where F is defined by

$$F=f(x,t)-\frac{\partial}{\partial t}\int_t^\infty h(s)\gamma(u(x,t-s))\,ds$$
$$+\operatorname{div}\int_t^\infty k(s)\mu(\nabla u(x,t-s))\,ds.$$

The initial and boundary conditions are in general as follows.

$$u(x,0)=u_0(x),\quad x\in\Omega,$$

$$\rho(\nabla u(x,t))-\int_0^t k(s)\mu(\nabla u(x,t-s))\,ds=g(x,t)$$

or

$$u(x,t)=g(x,t),\quad (x,t)\in\partial\Omega\times(0,\infty),$$

where u_0 and g are known functions. Therefore, problem (1.1) is just a special case of the above mentioned model. We refer to [**28**] and the references therein for the details of the mathematical modeling in viscoelasticity and thermoelasticity.

Let $\{S_h\}$ be a family of finite dimensional subspaces of $H_0^1(\Omega)$ with the following properties. For some integer $l\geq 2$,

$$\inf_{\chi\in S_h}(\|\chi-u\|+h\|\chi-u\|_1)\leq Ch^r\|u\|_r,\tag{1.2}$$
$$1\leq r\leq l,\quad u\in H^r(\Omega)\cap H_0^1(\Omega),$$

where C is a constant independent of u and h. $H^r(\Omega)$ is a Hilbert space of order r with nroms $\|\cdot\|_r$ and $H_0^1(\Omega)$ is the completion of $C_0^\infty(\Omega)$ under the $\|\cdot\|_1$ norm.

The semi-discrete finite element approximation to the solution u of (1.1) is now defined by $u_h(t) : \bar{J} \to S_h$,

$$(1.3) \qquad (u_{h,t}, \chi) + A(u_h, \chi) = \int_0^t a(t-s)B(u_h(s), \chi)\, ds + (f, \chi), \\ \chi \in S_h, \quad u_h(0) = v_h,$$

where $v_h \in S_h$ is an appropriate approximation of v into S_h, $A(\cdot,\cdot)$ and $B(\cdot,\cdot)$ are the bilinear forms on $H_0^1(\Omega) \times H_0^1(\Omega)$, which are associated with the operators A and B, respectively.

Numerical approximations to the solution of the problem (1.1) have received considerable attention recently. For example, finite difference, collocation methods and the methods of lines are studied in [**4, 11, 12, 14, 22, 23, 29, 37**]. Finite element methods for both smooth and nonsmooth data, with smooth kernels, are studied in [**5, 6, 7, 9, 12, 16, 18, 20, 21, 24, 34**] as well as its nonlinear counterparts [**5, 12, 18, 21**]. Also see [**7, 10, 19**] for the results dealing with weakly singular kernels. Basically speaking, there exist two different approaches in the energy method: the Ritz projection method and the Ritz-Volterra projection method. We shall describe briefly these two methods.

In [**12, 16, 34**] the authors employed the Ritz projection $R_h u : \bar{J} \to S_h$ in the analysis:

$$(1.4) \qquad A(u - R_h u, \chi) = 0, \quad \chi \in S_h.$$

If we write, as is usual for parabolic equations, the error $e(t) = (u - R_h u) + (R_h u - u_h) = \rho + \theta$, we see that it is sufficient to estimate θ only since $R_h u$ approximates u well [**8, 32, 35**]. Thus, we obtain from (1.1) and (1.3) that

$$(1.5) \quad (\theta_t, \chi) + A(\theta, \chi) = \int_0^t a(t-s)B(e(s), \chi)\, ds - (\rho_t, \chi), \quad \chi \in S_h.$$

As shown in [**34**] that the integral term of the righthand side of (1.5) will generate some additional difficulties into the analysis. Therefore, an appropriate splitting $\theta = \theta_1 + \theta_2$ (see [**34**] for detail) is necessary in order to obtain the optimal L^2 error estimates. However, it seems that there is no analogous easy splitting for nonlinear problems, and it seems also difficult to derive maximum norm error estimates and superconvergence by the method used in [**34**].

In [**5, 6**] the authors invented the so-called Ritz-Volterra projection $V_h : C(J, H_0^1) :\to C(J, S_h)$

$$(1.6) \quad A(u - V_h u, \chi) = \int_0^t a(t-s) B(u(s) - V_h u(s), \chi)\, ds, \quad \chi \in S_h.$$

By using this new projection we see easily that if we let the error $e(t) = (u - V_h u) + (V_h u - u_h) = \rho + \theta$, then we find θ satisfies

$$(1.7) \quad (\theta_t, \chi) + A(\theta, \chi) = \int_0^t a(t-s) B(\theta(s), \chi)\, ds - (\rho_t, \chi), \quad \chi \in S_h.$$

Thus, as demonstrated in [**5, 6, 18, 21**] all estimates for various norms of θ can be derived easily regardless of whether the equations being considered are linear or nonlinear. But, it does require some extra efforts to prove the optimality of the Ritz-Volterra projection $V_h u$ to u. It is clear that this extra work is well justified since this approach not only works for the finite element method for parabolic integro-differential equations but it also unifies the analysis in finite element methods for time-dependent problems [**21**].

By looking at the weak form (1.3), we find that the Ritz projection R_h is not consistent with this formulation since we have two elliptic operators in (1.3) while the Ritz projection is just defined for one positive operator. This may be the basic reason that difficulties are encountered if only the Ritz projection is used in the analysis for (1.3). On the other hand, we see that the Ritz-Volterra projection V_h is indeed consistent with our weak form since its definition incorporates the two operators A and B. This is the main reason that the authors of [**1, 5, 6, 18, 20, 21**] have used this projection successfully, not only for parabolic integro-differential equations, but also for hyperbolic integro-differential equations, Sobolev equations and the equations of visco-elasticity. We recall that all of these equations have two elliptic operators of the same order.

Unfortunately, some unexpected difficulties arose when the author of [**7, 19**] investigated semi-discrete finite element approximations for the problem (1.1) with only a weakly singular kernel $a(t) = t^{-\alpha}$, $0 < \alpha < 1$.

That is, it can be seen easily from (1.6) that, in general, the following asymptotic behavior is expected:

$$\left\| \frac{d}{dt}(u - V_h u) \right\| = O(t^{-\alpha}) \quad \text{as } t \to 0. \tag{1.8}$$

As shown in [**7, 19**], optimal L^2 error estimates can be obtained in the same way as in [**6, 21**], but (1.8) makes it difficult to derive maximum norm estimates due to the lack of regularity of $u - V_h u$ in time. This shows that the Ritz-Volterra projection may present some disadvantages when it is used for integrable kernels. However, when the kernel is smooth the maximum norm estimates for the Ritz-Volterra projection via the generalized Green function was obtained in [**20**] with applications to finite element approximations for integro-differential equations of parabolic type, Sobolev equations and parabolic equations with integral boundary conditions.

The above analysis indicates that we need to seek other possible ways in dealing with these problems. The purpose of this paper is to find a way to meet these needs. We shall show that the Ritz projection R_h can be used provided that some changes are made accordingly in the weak form (1.3) since, as stated before, it is not consistent with R_h as given.

Let $A_h : S_h \to S_h$ be defined by

$$(A_h \phi, \psi) = A(\phi, \psi), \quad \phi, \psi \in S_h, \tag{1.9}$$

and $B_h : S_h \to S_h$ by

$$(B_h \phi, \psi) = B(\phi, \psi), \quad \phi, \psi \in S_h. \tag{1.10}$$

Also let $P_h : L^2(\Omega) \to S_h$ be the L^2 projection defined by

$$(P_h \phi, \chi) = (\phi, \chi), \quad \phi \in L^2(\Omega),\ \chi \in S_h. \tag{1.11}$$

Now, using (1.9)–(1.11), we see that (1.3) can be written as

$$u_{h,t} + A_h u_h = \int_0^t a(t-s) B_h u_h(s)\, ds + P_h f, \tag{1.12}$$

and it follows by letting $T_h = A_h^{-1}$ that

$$u_{h,t} + A_h u_h = \int_0^t a(t-s) B_h T_h A_h u_h(s)\, ds + P_h f. \tag{1.13}$$

Thus, we obtain, by solving (1.13) for $A_h u_h$ as an unknown, that

$$\begin{aligned} u_{h,t} + \int_0^t K_h(t-s) u_{h,t}(s)\, ds + A_h u_h \\ = f + \int_0^t K_h(t-s) P_h f(s)\, ds, \end{aligned} \tag{1.14}$$

where $K_h(t)$ is the resolvent of $a(t)B_h T_h$ and is given by

$$K_h(t) = a(t) B_h T_h + \int_0^t a(t-s) B_h T_h K_h(s)\, ds. \tag{1.15}$$

Since we see, from [**8, 32**] that $T_h P_h = T_h$, it follows easily from (1.15) that $K_h(t) P_h = K_h(t)$ for all $t \in \bar{J}$. Hence, (1.14) is equivalent (therefore (1.3) is also equivalent) to the following form:

$$\begin{aligned} \left(u_{h,t} + \int_0^t K_h(t-s) u_{h,t}(s)\, ds, \chi \right) + A(u_h, \chi) \\ = \left(f + \int_0^t K_h(t-s) f(s)\, ds, \chi \right), \quad \chi \in S_h. \end{aligned} \tag{1.16}$$

We see now clearly that the Ritz projection may be used successfully since only one bilinear form appears in (1.16) while $K_h(t)$ is bounded in L^2 (see Section 2 for details). For the same reason the weak form of (1.1) can be written as

$$\begin{aligned} \left(u_t + \int_0^t K(t-s) u_t(s)\, ds, \phi \right) + A(u, \phi) \\ = \left(f + \int_0^t K(t-s) f(s)\, ds, \phi \right), \quad \phi \in H_0^1(\Omega), \end{aligned} \tag{1.17}$$

where $K(t)$ is the resolvent of $a(t)BT$ and is given by

$$K(t) = a(t)BT + \int_0^t a(t-s) BTK(s)\, ds, \tag{1.18}$$

where $T = A^{-1}$ is the solution operator for the elliptic problem

$$Aw = g \quad \text{in } \Omega, \tag{1.19}$$

$$w = 0 \quad \text{on } \partial\Omega. \tag{1.20}$$

We now have our new weak formulation for (1.1) and shall begin our analysis in Section 2. This paper is organized as follows. In Section 2 some necessary lemmas will be proved which are essential in the analysis. In Section 3 optimal L^2 error estimates will be presented, while maximum norm error estimates and superconvergence of the gradients will be demonstrated in Section 4.

We shall throughout this paper assume the inverse assumptions:

$$\|\chi\|_{1,p} \le Ch^{-1}\|\chi\|_{0,p}, \quad 1 \le p \le \infty,\ \chi \in S_h, \tag{1.21}$$

where $W_p^r(\Omega)$, $2 \le p \le \infty$, is the usual Sobolev space with norms $\|\cdot\|_{r,p}$, $\|\cdot\|_r = \|\cdot\|_{r,2}$, $\|\cdot\| = \|\cdot\|_{0,2}$, and $\overset{o}{W}{}_p^1(\Omega)$ is the completion of $C_0^\infty(\Omega)$ under $\|\cdot\|_{1,p}$.

Remark. (i) We assume $a(t) \in L^1(0,T)$, and it certainly covers the following cases:

$$\begin{aligned} a(t) = \sum_{i=1}^{M} c_i t^{-\nu_i} \exp(-\mu_i t), \quad c_i, \mu_i \in R,\\ 0 < \nu_i < 1, \quad i = 1, \dots, M, \end{aligned} \tag{1.22}$$

since each term in the summation is integrable.

(ii) Recently Hornung and Showalter derived, in the study of diffusion models for fractured media [**13**], the following model

$$u_t + \int_0^t b(t-s)u_t(s)\,ds + Au = f \tag{1.23}$$

with

$$b(t) = 6\alpha \sum_{k=1}^{\infty} \exp(-k^2\pi^2\alpha t), \quad t > 0, \quad \alpha > 0. \tag{1.24}$$

Our results in this paper are also valid for this problem since $b(t) = O(t^{-1/2})$ as $t \to 0$ and is also integrable. To see this, we need to observe that the resolvent $K(t)$ is integrable in time, see Section 2.

(iii) In the recent paper by M. Peszynska [**26**], the author dealt with equation (1.23) with a smooth kernel $b(t)$ by finite element methods, which is in fact a special case of (1.1). Only L^2 error estimates are derived for semi-discrete and backward Euler approximations.

As the final remark of this introduction section, we notice that the resolvent K_h in (1.15) is well defined since $B_h T_h$ is a matrix or bounded operator on S_h. Similarly, the resolvent $K(t)$ in (1.18) is also well defined since BT is a bounded operator in $L^2(\Omega)$, so that $K(t)$ is a bounded operator in $L^1(J, L^2(\Omega))$ [**25**]. Also the asymptotic constant $C = C(T_0)$ in the error estimates in the next sections will grow with T_0 due to the use of Gronwall's inequality, which limits its validity only to the case T_0 finite. Global error estimates with asymptotic constant C independent of the time, $t \geq 0$, have been recently obtained in [**33**] for a smooth kernel and [**2**] for an integrable kernel.

2. Preliminaries and lemmas.

In this section we shall define some notations and prove a series of lemmas which are needed in the sequel. Without loss of generality it is assumed that the kernel $a(t)$ is nonnegative throughout this paper. We begin by the following result.

Lemma 2.1. *There exists $C > 0$ such that*

$$\|P_h w\|_k \leq C\|w\|_k, \quad w \in H^k(\Omega),\ k = 0, 1; \tag{2.1}$$

$$\|w - P_h w\| \leq Ch^r \|w\|_r, \quad 0 \leq r \leq l; \tag{2.2}$$

$$\|(T - T_h)w\|_k \leq Ch^{r+2-k}\|w\|_r, \quad w \in H^r(\Omega),\ k = 0, 1,\ 0 \leq r \leq l - 2. \tag{2.3}$$

Proof. Equations (2.1) and (2.2) are the stability and optimality of the L^2 projection [**8, 15, 23, 27, 36**], while (2.3) is the well-known error estimate for elliptic problems [**8, 34**] or [**32**, p. 52]. ◻

Definition 2.1. Let $F : H^r(\Omega) \to H^r(\Omega)$ and its operator norm be defined by

$$\|F\|_r = \sup_{0 \neq w \in H^r(\Omega)} \frac{\|Fw\|_r}{\|w\|_r}, \quad r \geq 0. \tag{2.4}$$

Definition 2.2. Let $G : H^r(\Omega) \to L^2(\Omega)$ and its operator norm be defined by

$$\|G\|_r^* = \sup_{0 \neq w \in H^r(\Omega)} \frac{\|Gw\|}{\|w\|_r}, \quad r \geq 0. \tag{2.5}$$

By these definitions, we have

Lemma 2.2. *The operator BT is bounded from $H^r(\Omega) \to H^r(\Omega)$, i.e., there exists $C = C(r) > 0$ such that*

$$\|BT\|_r \leq C(r), \quad \textit{for } r \geq 0. \tag{2.6}$$

Proof. Let $w \in H^r(\Omega)$ and $y = Tw$. Then it follows from elliptic regularity, $\|y\|_{r+2} \leq C(r)\|w\|_r$, and consequently

$$\|BTw\|_r \leq C\|Tw\|_{r+2} = C\|y\|_{r+2} \leq C(r)\|w\|_r.$$

Thus, Lemma 2.2 follows from Definition 2.1. □

Lemma 2.3. *The operator B_hT_h is bounded from $H^k(\Omega) \to H^k(\Omega)$, $k = 0, 1$, i.e., there exists $C > 0$, independent of h, such that*

$$\|B_hT_h\|_k \leq C, \quad k = 0, 1,\ l \geq 3, \quad \textit{and} \quad k = 0,\ l = 2. \tag{2.7}$$

Proof. For $k = 0$, let $w, \phi \in L^2(\Omega)$. It follows from $B_hT_hw \in S_h$ and (1.10) that

$$\begin{aligned}(B_hT_hw, \phi) &= (B_hT_hw, P_h\phi) = B(T_hw, P_h\phi) \\ &= B((T_h - T)w, P_h\phi) + B(Tw, P_h\phi).\end{aligned} \tag{2.8}$$

We find from Lemma 2.1 and the inverse assumption (1.21) that

$$\begin{aligned}B((T_h - T)w, P_h\phi) \le C\|(T_h - T)w\|_1\|P_h\phi\|_1 &\le Ch\|w\|h^{-1}\|P_h\phi\| \\ &\le C\|w\|\,\|\phi\|,\end{aligned}$$

and from $Tw \in H_0^1(\Omega) \cap H^2(\Omega)$ that

$$B(Tw, P_h\phi) = (BTw, P_h\phi) \le C\|w\|\,\|\phi\|. \tag{2.9}$$

Thus, we see that

$$\|B_hT_hw\| \le C\|w\|,$$

which completes the case of $k = 0$.

For $k = 1$, let $w \in H^1(\Omega)$, as we know that

$$\|B_hT_hw\|_1 = \sup_{0\neq\phi\in C^\infty(\Omega)} \frac{(B_hT_hw, \phi)}{\|\phi\|_{-1}}$$

and (2.8) is also valid for $\phi \in C^\infty(\Omega)$. But we have from (2.3) and the inverse assumption (1.21) that

$$\begin{aligned}B((T_h - T)w, P_h\phi) &\le C\|(T - T_h)w\|_1\|P_h\phi\|_1 \\ &\le Ch^2\|w\|_1\|P_h\phi\|_1 \\ &\le Ch\|w\|_1\|P_h\phi\|\end{aligned}$$

and from (2.9) that

$$B(Tw, P_h\phi) \le \|BTw\|_1\|P_h\phi\|_{-1} \le C\|w\|_1\|P_h\phi\|_{-1}.$$

Thus, one finds that

$$(B_hT_hw, \phi) \le C\|w\|_1(h\|P_h\phi\| + \|P_h\phi\|_{-1}). \tag{2.10}$$

But, for any $\xi \in L^2(\Omega)$,

$$\begin{aligned}(P_h\phi, \xi) = (P_h\phi, P_h\xi) &\le \|P_h\phi\|_{-1}\|P_h\xi\|_1 \\ &\le Ch^{-1}\|P_h\phi\|_{-1}\|P_h\xi\| \\ &\le Ch^{-1}\|P_h\phi\|_{-1}\|\xi\|,\end{aligned}$$

so that it follows

$$\|P_h\phi\| \le Ch^{-1}\|P_h\phi\|_{-1}. \tag{2.11}$$

Similarly, for any $\xi \in H^1(\Omega)$, it follows that

$$\begin{aligned}(P_h\phi, \xi) &= (P_h\phi, P_h\xi) = (\phi, P_h\xi)\\ &\le \|\phi\|_{-1}\|P_h\xi\|_1 \le C\|\phi\|_{-1}\|\xi\|_1,\end{aligned}$$

and then we obtain

$$\|P_h\phi\|_{-1} \le C\|\phi\|_{-1}. \tag{2.12}$$

Combining (2.10), (2.11) and (2.12), we obtain that

$$(B_hT_hw, \phi) \le C\|w\|_1\|\phi\|_{-1},$$

which is the case of $k = 1$ in (2.7). Therefore, Lemma 2.3 is complete. □

Lemma 2.4. *There exists $C = C(r) > 0$ such that*

$$\|B_hT_h - BT\|_r^* \le Ch^r, \quad 0 \le r \le l. \tag{2.13}$$

Proof. Since this is trivial for $r = 0$, we consider $1 \le r \le l$. For $w \in H^r(\Omega)$ we have

$$\begin{aligned}(BT - B_hT_h)w = (BT - P_h(BT))w\\ + (P_h(BT) - B_hT_h)(w - P_hw)\\ + (P_h(BT) - B_hT_h)P_hw,\end{aligned}$$

thus it follows that

$$\begin{aligned}\|(BT - B_hT_h)w\| &= Ch^r\|BTw\|_r + C\|(w - P_hw)\|\\ &\quad + C\|(P_h(BT) - B_hT_h)P_hw\|\\ &\le Ch^r\|w\|_r + C\|(P_h(BT) - B_hT_h)P_hw\|.\end{aligned} \tag{2.14}$$

Assume at this moment that

$$\|(P_h(BT) - B_hT_h)^* P_h w\| \leq Ch^r \|w\|_r, \quad w \in H^r(\Omega),$$

where $(P_h(BT) - B_hT_h)^*$ is the adjoint of $P_h(BT) - B_hT_h$ on S_h. Since P_h is a map from $H^r(\Omega)$ onto S_h, by the inequality above there exists $w_0 \in H^r(\Omega)$ such that, for any $w \in H^r(\Omega)$,

$$\begin{aligned} Ch^r\|w_0\|_r &\geq \|(P_h(BT) - B_hT_h)^* P_h w_0\| \\ &= \|(P_h(BT_w) - B_hT_h)^*\| \\ &= \|P_h(BT) - B_hT_h\| \\ &\geq \left\| (P_h(BTw) - B_hT_h) \frac{P_h w}{\|P_h w\|} \right\|, \end{aligned}$$

where the operator norms are taken on S_h, and then

$$\|(P_h(BT) - B_hT_h)P_h w\| \leq Ch^r \|w_0\|_r \|P_h w\| \leq Ch^r \|w\|_r.$$

Therefore, Lemma 2.4 is proved by substituting the above inequality into (2.14).

It now remains to verify (2.14). For $\psi \in L^2(\Omega)$, we find that

$$\begin{aligned} ((P_h(BT) - B_hT_h)^* P_h w, \psi) &= ((P_h(BTw) - B_hT_h)^* P_h w, P_h \psi) \\ &= (P_h w, (P_h(BT) - B_hT_h) P_h \psi) \\ &= (P_h w, BTP_h \psi) - (P_h w, B_hT_hP_h\psi) \\ &= B^*(P_h w, TP_h\psi) - B^*(P_h w, T_hP_h\psi) \\ &= B^*(P_h w - w, (T - T_h)P_h\psi) \\ &\quad + B^*(w, (T - T_h)P_h\psi) \\ &= B^*(P_h w - w, (T - T_h)P_h\psi) \\ &\quad + (B^* w, (T - T_h)P_h\psi) \\ &\leq Ch^{r-1}\|w\|_r \|(T - T_h)P_h\psi\|_1 \\ &\quad + C\|w\|_r \|(T - T_h)P_h\psi\|_{-r+2} \\ &\leq Ch^r \|w\|_r \|\psi\|, \end{aligned}$$

which implies (2.14). In fact, we have used the negative norm estimates [**32**, p. 77]

$$\|(T - T_h)g\|_{-s} \leq Ch^{s+q}\|g\|_q, \quad 0 \leq s \leq l-2,\ 1 \leq q \leq l. \qquad \square$$

Lemma 2.5. *Assume that* $a(t) \in L^1(0,T)$ *and* $f(t) \in C^1([0,T])$. *Then we have*

$$\frac{d}{dt}\int_0^t a(t-s)f(s)\,ds = f(0)a(t) + \int_0^t a(t-s)f'(s)\,ds. \tag{2.15}$$

Proof. Since

$$\int_0^t a(t-s)f(s)\,ds = \int_0^t a(s)f(t-s)\,ds;$$

thus, (2.15) follows by differentiation. □

Lemma 2.6. *Let* $a(t) \in L^1(0,T)$. *Then we have*

$$\int_0^t \int_0^s a(s-\tau)f(\tau)\,d\tau\,ds = \int_0^t a(t-s)\int_0^s f(\tau)\,d\tau\,ds. \tag{2.16}$$

Proof. It follows by exchanging the order of integration and integration by parts that

$$\begin{aligned}
\int_0^t \int_0^s a(t-\tau)f(\tau)\,d\tau\,ds &= \int_0^t \int_\tau^t a(s-\tau)f(\tau)\,ds\,d\tau \\
&= \int_0^t \int_0^{t-\tau} a(\xi)\,d\xi f(\tau)\,d\tau \\
&= \int_0^{t-\tau} a(\xi)\,d\xi \int_0^\tau f(\xi)\,d\xi\Big|_0^t \\
&\quad + \int_0^t a(t-\tau)\int_0^\tau f(\xi)\,d\xi\,d\tau \\
&= \int_0^t a(t-\tau)\left(\int_0^\tau f(\xi)\,d\xi\right)d\tau. \qquad \square
\end{aligned}$$

In order to estimate the difference between $K(t)$ and $K_h(t)$ defined by (1.18) and (1.15), respectively, we need the following version of Gronwall's inequality of convolution type:

Lemma 2.7. *Let $k(t)$, $g(t) \in L^1(0,T)$ be nonnegative functions and $f(t) \geq 0$ be such that*

$$f(t) \leq Cg(t) + C\int_0^t k(t-s)f(s)\,ds, \tag{2.17}$$

then we have

$$f(t) \leq C\bigg(g(t) + \int_0^t R_k(t-s)g(s)\,ds\bigg), \tag{2.18}$$

where $R_k(t) \in L^1(0,T)$ is the resolvent of $k(t)$ and satisfies

$$R_k(t) = k(t) + \int_0^t k(t-s)R_k(s)\,ds. \tag{2.19}$$

In particular, we have

$$f(t) \leq C\begin{cases} g(t) & \text{if } g \text{ is monotone increasing,} \\ R_k(t) & \text{if } g(t) = k(t). \end{cases} \tag{2.20}$$

Proof. See [**3**, Chapter 1]. □

Lemma 2.8. *Let $F, S : H^r(\Omega) \to H^r(\Omega)$ and $G : H^r(\Omega) \to L^2$. Then it holds*

$$\|FS\|_r \leq \|F\|_r\|S\|_r, \qquad \|F\|_r^* \leq \|F\|_r, \tag{2.21}$$

$$\|GF\|_r^* \leq \|G\|_r^*\|F\|_r, \tag{2.22}$$

$$\|B_hT_hG\|_r^* \leq \|B_hT_h\|\,\|G\|_r^* \leq C\|G\|_r^*. \tag{2.23}$$

Proof. It follows directly from Definition 2.1 and Definition 2.2. □

Lemma 2.9. *There exists $C = C(r) > 0$ such that the resolvent $K(t)$ and $K_h(t)$ in (1.18) and (1.15), respectively, satisfy*

$$\begin{aligned} \|K(t)\|_r &\leq C(r)R_a(t), \quad t \in (0,T],\ 0 \leq r \leq l, \\ \|K_h(t)\| &\leq C(r)R_a(t), \quad t \in (0,T]. \end{aligned} \tag{2.24}$$

Proof. By (1.18) and (2.4), we have

$$\|K(t)\|_r \le a(t)\|BT\|_r + \int_0^t a(t-s)\|BTK(s)\|_r\,ds.$$

It follows from Lemma 2.2 and Lemma 2.8 that

$$\|BTK(s)\|_r \le \|BT\|_r\|K(s)\|_r \le C\|K(s)\|_r$$

so that we obtain

$$\|K(t)\|_r \le Ca(t) + C\int_0^t a(t-s)\|K(s)\|_r\,ds.$$

Hence, Lemma 2.9 follows from Lemma 2.7 with $k(t) = a(t)$. □

Lemma 2.10. *There exists $C = C(r) > 0$ such that*

$$\|K(t) - K_h(t)\|_r^* \le Ch^r m_a(t), \quad t \in (0, T],\ 0 \le r \le l, \tag{2.25}$$

where $m_a(t) \in L^1(0,T)$ and is defined by

$$m_a(t) = R_a(t) + \int_0^t R_a(t-s)R_a(s)\,ds.$$

Proof. Since we have from (1.18) and (1.15) that

$$\begin{aligned} K(t) - K_h(t) = a(t)(BT - B_hT_h) & \\ + \int_0^t a(t-s)(BT - B_hT_h)K(s)\,ds & \\ + \int_0^t a(t-s)B_hT_h(K(s) - K_h(s))\,ds, & \end{aligned}$$

so that it holds

$$\begin{aligned} \|K(t) - K_h(t)\|_r^* \le a(t)\|BT - B_hT_h\|_r^* & \\ + \int_0^t a(t-s)\|(BT - B_hT_h)K(s)\|_r^*\,ds & \\ + \int_0^t a(t-s)\|B_hT_h(K(s) - K_h(s))\|_r^*\,ds. & \end{aligned} \tag{2.26}$$

By using Lemmas 2.8 and 2.9, we have

$$\|(BT - B_hT_h)K(s)\|_r^* \le \|BT - B_hT_h\|_r^* \|K(s)\|_r \le Ch^r R_a(s)$$

and

$$\begin{aligned}\|B_hT_h(K(s) - K_h(s))\|_r^* &\le \|B_hT_h\|\,\|K(s) - K_h(s)\|_r^* \\ &\le C\|K(s) - K_h(s)\|_r^*.\end{aligned}$$

Thus, we obtain that

$$\begin{aligned}\|K(t)-K_h(t)\|_r^* &\le Ch^r a(t) + Ch^r \int_0^t a(t-s)R_a(s)\,ds \\ &\quad + C\int_0^t a(t-s)\|K(s)-K_h(s)\|_r^*\,ds \\ &\le Ch^r R_a(t) + C\int_0^t a(t-s)\|K(s)-K_h(s)\|_r^*\,ds.\end{aligned}$$

Hence, Lemma 2.10 follows from an application of Lemma 2.7. □

Lemma 2.11. *Assume that* $0 \le a(t) \in L^1(0,T)$. *Then* $R_a(t)$ *and* $m_a(t)$ *are nonnegative and are in* $L^1(0,T)$.

Proof. It follows from the definitions of R_a and m_a. □

3. Optimal L^2 error estimates. In this section the optimal L^2 error estimates will be proved for the semi-discrete finite element approximation. Theorem 3.1 (without f on the righthand side of (3.1)) below with smooth kernels has been proved by using the Ritz projection [**34**] and the Ritz-Volterra projection [**5, 6, 21**]. Since our proof based on our weak formulations is very different from that of [**5, 6, 21, 34**] and can be used in the next sections, we give the proof here for completeness.

Theorem 3.1. *Assume that* u *and* u_h *are the solutions of* (1.1) *and* (1.3), *respectively,* $a(t) \in L^1(0,T)$. *If* $u_t,\ f \in L^1(J; H^r(\Omega))$, $v \in H^r(\Omega) \cap H_0^1(\Omega)$ *and* $\|v - v_h\| + h\|v - v_h\|_1 \le Ch^r\|v\|_r$, *then there exists* $C > 0$, *independent of* h *and* u, *such that*

$$\|u(t) - u_h(t)\| \le Ch^r\left(\|v\|_r + \int_0^t (\|u_t(s)\|_r + \|f(s)\|_r)\,ds\right). \tag{3.1}$$

Proof. By using Lemma 2.5 we see that

$$\frac{d}{dt}\int_0^t K_h(t-s)u_h(s)\,ds = \int_0^t K_h(t-s)u_{h,t}(s)\,ds + K_h(t)u_h(0);$$

thus, (1.16) can be written as

$$\begin{aligned}(3.2)\quad &\left(\frac{d}{dt}\left(u_h + \int_0^t K_h(t-s)u_h(s)\,ds\right),\chi\right) + A(u_h,\chi)\\ &\qquad = \left(f + \int_0^t K_h(t-s)f(s)\,ds,\chi\right) + (K_h(t)v_h,\chi), \quad \chi\in S_h.\end{aligned}$$

Similarly, (1.17) can be written as

$$\begin{aligned}(3.3)\quad &\left(\frac{d}{dt}\left(u + \int_0^t K(t-s)u(s)\,ds\right),\chi\right) + A(u,\chi)\\ &\qquad = \left(f + \int_0^t K(t-s)f(s)\,ds,\chi\right) + (K(t)v,\chi), \quad \chi\in S_h.\end{aligned}$$

Let the error $e(t) = \rho(t) + \theta(t)$ where $\rho(t) = u(t) - R_h u(t)$ and $\theta(t) = R_h u(t) - u_h(t)$. We have from [**8, 15, 29, 35, 37**] that

$$(3.4)\qquad \|\rho(t)\| \le Ch^r\|u(t)\|_r \le Ch^r\left(\|v\|_r + \int_0^t \|u_t(s)\|_r\,ds\right).$$

Thus, it remains to estimate $\theta(t)$ in L^2 only. Let

$$(3.5)\qquad \Theta = \theta + \int_0^t K_h(t-s)\theta(s)\,ds,$$

and it follows from (3.2) and (3.3) that

$$\begin{aligned}(\Theta_t,\chi) + A(\theta,\chi) &= -\left(\rho_t + \int_0^t K_h(t-s)\rho_t(s)\,ds,\chi\right) + (K_h(t)\theta(0),\chi)\\ (3.6)\qquad &\quad + \left(\int_0^t (K(t-s) - K_h(t-s))(f(s)-u_t(s))\,ds,\chi\right)\\ &= \sum_{i=1}^3 (J_i,\chi), \quad \chi\in S_h,\end{aligned}$$

where

$$J_1 = -\rho_t + \int_0^t K_h(t-s)\rho_t(s)\,ds, \tag{3.7}$$

$$J_2 = K_h(t)\theta(0), \tag{3.8}$$

$$J_3 = \int_0^t (K(t-s) - K_h(t-s))(f(s) - u_t(s))\,ds. \tag{3.9}$$

If we let $\chi = \Theta \in S_h$, it follows from Lemma 2.9 that

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\Theta\|^2 + A(\theta,\theta) &= \sum_{i=1}^{3}(J_i,\Theta) - A\left(\theta(t), \int_0^t K_h(t-s)\theta(s)\,ds\right)\\ &\le \sum_{i=1}^{3}(J_i,\Theta) + C\|\theta(t)\|_1\\ &\quad\cdot \int_0^t \|K_h(t-s)\theta(s)\|_1\,ds\\ &\le \sum_{i=1}^{3}(J_i,\Theta) + C\|\theta(t)\|_1\\ &\quad\cdot \int_0^t R_a(t-s)\|\theta(s)\|_1\,ds.\end{aligned}$$

Integrating from 0 to t and using $\Theta(0) = \theta(0)$, we obtain that

$$\begin{aligned}\|\Theta(t)\|^2 + \int_0^t \|\theta(s)\|_1^2\,ds &\le C\|\theta(0)\|^2 + \int_0^t \sum_{i=1}^{3}\|J_i\|\,\|\Theta\|\,ds\\ &\quad + C\int_0^t \|\theta(s)\|_1\left(\int_0^s R_a(s-\tau)\|\theta(\tau)\|_1\,d\tau\right)ds\\ &= C\|\theta(0)\|^2 + Q_1(t) + Q_2(t).\end{aligned} \tag{3.10}$$

It is easy to see, by changing the order of integration and integration by parts, from Lemma 2.11 for Q_2 that

$$(3.11)\quad \begin{aligned} Q_2(t) &\le \varepsilon \int_0^t \int_0^s R_a(s-\tau)\|\theta(s)\|_1^2\, d\tau\, ds \\ &\quad + C(\varepsilon)\int_0^t \int_0^s R_a(s-\tau)\|\theta(\tau)\|_1^2\, d\tau\, ds \\ &\le \varepsilon C \int_0^t \|\theta(s)\|_1^2\, ds \\ &\quad + C(\varepsilon)\int_0^t R_a(t-s)\bigg(\int_0^s \|\theta(\tau)\|_1^2\, d\tau\bigg)\, ds. \end{aligned}$$

Thus, we substitute (3.11) into (3.10) with $\varepsilon > 0$ small and fixed, and use Lemma 2.7 to obtain that

$$\|\Theta(t)\|^2 + \int_0^t \|\theta(s)\|_1^2\, ds \le C(\|\theta(0)\|^2 + Q_1(t)),$$

where the monotonic nondecreasing property of $Q_1(t)$ was used.

But we have for Q_1 that

$$Q_1(t) \le \frac{1}{2}\sup_{s\le t}\|\Theta(s)\|^2 + C\bigg(\int_0^t \sum_{i=1}^3 \|J_i\|\, ds\bigg)^2$$

and then

$$(3.12)\quad \|\Theta(t)\|^2 \le \frac{1}{2}\sup_{s\le t}\|\Theta(s)\|^2 + C\bigg(\|\theta(0)\| + \sum_{i=1}^3 \int_0^t \|J_i\|\, ds\bigg)^2.$$

Since (3.12) holds for all $t \in J$, we conclude that

$$(3.13)\quad \|\Theta(t)\| \le C\bigg(\|\theta(0)\| + \sum_{i=1}^3 \int_0^t \|J_i\|\, ds\bigg).$$

Notice from Lemma 2.9 that

$$(3.14)\quad \begin{aligned} \|\Theta(t)\| &\ge \|\theta(t)\| - \int_0^t \|K_h(t-s)\theta(s)\|\, ds \\ &\ge \|\theta(t)\| - C\int_0^t R_a(t-s)\|\theta(s)\|\, ds, \end{aligned}$$

and hence, substituting (3.14) into (3.13) and applying Lemma 2.7 as before, it follows that

$$\|\theta(t)\| \leq C\bigg(\|\theta(0)\| + \sum_{i=1}^{3} \int_0^t \|J_i\|\, ds\bigg). \tag{3.15}$$

It is easy to see from (3.7) and Lemma 2.9 that

$$\begin{aligned}\|J_1\| &\leq \|\rho_t\| + \int_0^t \|K_h(t-s)\rho_t(s)\|\, ds \\ &\leq Ch^r\|u_t\|_r + Ch^r \int_0^t R_a(t-s)\|u_t(s)\|_r\, ds,\end{aligned}$$

and then

$$\int_0^t \|J_1\|\, ds \leq Ch^r \int_0^t \|u_t(s)\|_r\, ds.$$

Similarly, we have

$$\int_0^t \|J_2\|\, ds \leq Ch^r\|\theta(0)\|$$

and

$$\int_0^t \|J_3\|\, ds \leq Ch^r \int_0^t (\|u_t(s)\|_r + \|f(s)\|_r)\, ds.$$

Combining the above estimates and (3.15), we have

$$\|\theta(t)\| \leq C\bigg(\|\theta(0)\| + h^r\bigg(\|v\|_r + \int_0^t (\|u_t(s)\|_r + \|f(s)\|_r)\, ds\bigg)\bigg). \tag{3.16}$$

From our assumption on v_h, it follows that $\|\theta(0)\| \leq Ch^r\|v\|_r$. Hence, Theorem 3.1 is completed by (3.4), (3.16) and the triangle inequality. □

4. Maximum norm estimates and superconvergence in R^2. Let Ω be a bounded domain in R^2 with polygonal boundary $\partial\Omega$. For $k \geq 2$, $0 < h \leq 1$, let S_h^k be one parameter family of finite element subspaces of $\overset{o}{W}{}_2^1(\Omega)$ [**8, 15, 34**], consisting of piecewise polynomial functions of degree at most $k-1$, defined on a quasi-uniform partition of Ω [**8**]. It is required that S_h^k possess the following approximation properties. For any $w \in \overset{o}{W}{}_2^1(\Omega) \cap W_p^k(\Omega)$,

$$\inf_{\chi \in S_h^k} (\|w-\chi\|_p + h\|w-\chi\|_{1,p}) \le Ch^r \|w\|_{r,p}, \quad p \ge 2,\ 1 \le r \le k. \tag{4.1}$$

Lemma 4.1. *Let $P_h : L^2(\Omega) \to S_h^k$ be the L^2 projection. Then for $w \in \overset{o}{W}{}_p^1(\Omega) \cap W_p^r(\Omega)$, it holds*

$$\|P_h w\|_{r,p} \le C\|w\|_{r,p}, \quad r = 0, 1,\ 2 \le p \le \infty. \tag{4.2}$$

Proof. See [**8, 15, 36**]. □

Let $z \in \Omega$ and $\delta_h^z \in S_h^k$ be the discrete δ-function at $x = z$ such that

$$(\delta_h^z, \chi) = \chi(z), \quad \chi \in S_h^k. \tag{4.3}$$

Let $G_z \in \overset{o}{W}{}_2^1(\Omega) \cap W_2^2(\Omega)$ be the smooth Green's function at $x = z$ defined by

$$AG^z = \delta_h^z \quad \text{in } \Omega. \tag{4.4}$$

It is obvious from (4.3)–(4.4) that

$$A(G^z, w) = P_h w(z), \quad w \in W_2^1(\Omega). \tag{4.5}$$

Let $G_h^z \in S_h^k$ be the Ritz projection of G^z, i.e.,

$$A(G^z - G_h^z, \chi) = 0, \quad \chi \in S_h^k. \tag{4.6}$$

Lemma 4.2. *For Green's function G^z and its Ritz projection G_h^z, we have*

$$\|G^z - G_h^z\|_{1,1} \le Ch\left(\log\frac{1}{h}\right)^{k^*}, \quad k^* = \begin{cases} 1 & \text{if } k = 2, \\ 0 & \text{if } k \ge 3, \end{cases} \tag{4.7}$$

$$\|G^z\|_{1,1} + \|G_h^z\|_{1,1} + \|G_h^z\| \leq C. \tag{4.8}$$

Proof. See [**17, 27, 36**]. □

Theorem 4.1. *For $k = 2$, we assume that $u(t) \in L^1(J; \overset{o}{W}{}_2^1 \cap W_\infty^2)$, $u_t(t) \in L^2(J; W_2^2)$ and $u_h(0) = R_h(0)v$. Then we have, for $t \in J$,*

$$\begin{aligned} \|u(t) - u_h(t)\|_{0,\infty} \leq Ch^2 \bigg(& \log \frac{1}{h}(\|v\|_{2,\infty} + \|u(t)\|_{2,\infty}) \\ & + \left[\log \frac{1}{h} \int_0^t (\|u_t(s)\|_{2,2}^2 + \|f(s)\|_{2,2}^2)\, ds \right]^{1/2} \bigg). \end{aligned} \tag{4.9}$$

For $k \geq 3$, we assume that $u(t) \in L^1(J; \overset{o}{W}{}_2^1 \cap W_\infty^k)$, $u_h(0) = R_h(0)v$, $u_{tt}(t) \in L^1(J; W_2^k)$ and $u_t(0) \in W_2^k$. Thus we have, for $t \in J$,

$$\begin{aligned} \|u(t) - u_h(t)\|_{0,\infty} \leq Ch^k \bigg(& \|v\|_{k,2} + \|u(t)\|_{k,\infty} \\ & + \|f(0)\|_{k,2} + \|u_t(0)\|_{k,2} \\ & + \int_0^t (\|u_{tt}(s)\|_{k,2} + \|f_t(s)\|_{k,2})\, ds \bigg). \end{aligned} \tag{4.10}$$

Proof. We first show the result of $k = 2$. It is well known under our assumptions on S_h^k that we have

$$\|\rho(t)\|_{0,\infty} \leq Ch^2 \log \frac{1}{h} \|u(t)\|_{2,\infty}, \tag{4.11}$$

which is the standard error estimate for elliptic problems [**8, 27, 36**]. Thus we need to estimate θ only. Since $v_h = R_h v$, then $\theta(0) = 0$. We see now from (3.8) that $J_2 = 0$. We find from Lemma 2.5 that

$$\left(\theta_t + \int_0^t K_h(t-s)\theta_t(s)\, ds, \chi \right) + A(\theta, \chi) = (J_1, \chi) + (J_3, \chi). \tag{4.12}$$

Now let $\chi = \theta_t$. We have

$$\begin{aligned} \|\theta_t\|^2 + \frac{1}{2}\frac{d}{dt}A(\theta,\theta) &= (J_1 + J_3, \theta_t) - \left(\int_0^t K_h(t-s)\theta_t(s)\, ds, \theta_t(t) \right) \\ &= K_1 + K_2. \end{aligned} \tag{4.13}$$

Since

$$K_1 \le \varepsilon\|\theta_t\|^2 + C(\varepsilon)(\|J_1\|^2 + \|J_2\|^2) \tag{4.14}$$

and

$$\begin{aligned} K_2 &\le C\int_0^t R_a(t-s)\|\theta_t(s)\|\,\|\theta_t(t)\|\,ds \\ &\le \varepsilon C\|\theta_t(t)\|^2 + C(\varepsilon)\int_0^t R_a(t-s)\|\theta_t(s)\|^2\,ds; \end{aligned} \tag{4.15}$$

thus, if we select an $\varepsilon > 0$ small and fixed, substitute (4.14) and (4.15) into (4.13) and integrate the resultant inequality, we have

$$\begin{aligned} \int_0^t \|\theta_t\|^2\,ds + \|\theta\|_1^2 \le C\bigg(&\int_0^t (\|J_1\|^2 + \|J_3\|^2)\,ds \\ &+ \int_0^t\int_0^s R_a(s-\tau)\|\theta_t(\tau)\|^2\,d\tau\,ds\bigg). \end{aligned} \tag{4.16}$$

Then we obtain, by applying Lemmas 2.6 and 2.7, that

$$\int_0^t \|\theta_t\|_2\,ds + \|\theta\|_1^2 \le C\bigg(\int_0^t \|J_1\|^2 + \|J_3\|^2\bigg)\,ds. \tag{4.17}$$

It is obvious from the Cauchy inequality and (3.7) that

$$\|J_1\|^2 \le \|\rho_t\|^2 + C\int_0^t R_a(t-s)\|\rho_t(s)\|^2\,ds, \tag{4.18}$$

and then

$$\int_0^t \|J_1\|^2\,ds \le C\int_0^t \|\rho_t(s)\|^2\,ds \le Ch^4\int_0^t \|u_t(s)\|_{2,2}^2\,ds. \tag{4.19}$$

Similarly, we have from (3.9),

$$\int_0^t \|J_3\|^2\,ds \le Ch^4\int_0^t (\|u_t(s)\|_{2,2}^2 + \|f(s)\|_{2,2}^2)\,ds. \tag{4.20}$$

Thus, it follows that

$$\|\theta(t)\|_1^2 \le Ch^4 \int_0^t (\|u_t(s)\|_{2,2}^2 + \|f(s)\|_{2,2}^2)\, ds. \tag{4.21}$$

The inverse assumptions (1.21) imply that

$$\begin{aligned}\|\theta(t)\|_{0,\infty} &\le C\left(\left(\log\frac{1}{h}\right)^{1/2}\|\theta(t)\|_1\right.\\ &\le Ch^2\left(\log\frac{1}{h}\int_0^t (\|u_t(s)\|_{2,2}^2 + \|f(s)\|_{2,2}^2)\, ds\right)^{1/2}.\end{aligned} \tag{4.22}$$

Hence (4.9) is completed by (4.11), (4.22) and the triangle inequality.

Now we consider the case of $k \ge 3$. By writing (4.12) with $e = u(t) - u_h(t)$ as

$$\begin{aligned}A(\theta,\chi) &= \left(\int_0^t (K(t-s) - K_h(t-s))(f-u_t)\, ds, \chi\right)\\ &\quad - \left(e_t + \int_0^t K_h(t-s)e_t(s)\, ds, \chi\right)\end{aligned} \tag{4.23}$$

and letting $\chi = G_h^z$ in (4.23), it follows from (4.5) and Lemma 4.2 that

$$\begin{aligned}|\theta(z,t)| &= \left(\left\|\int_0^t (K(t-s) - K_h(t-s))(f-u_t)\, ds\right\|\right.\\ &\quad + \left.\left\|e_t + \int_0^t K_h(t-s)e_t(s)\, ds\right\|\right)\|G_h^z\|\\ &\le Ch^k\int_0^t (\|f\|_{k,2} + \|u_t\|_{k,2})\, ds\\ &\quad + C\left(\|e_t\| + \int_0^t R_a(t-s)\|e_t(s)\|\, ds\right).\end{aligned} \tag{4.24}$$

We now assume that

$$\begin{aligned}\|e_t(t)\| \le Ch^k\Big(&\|u_t(0)\|_{k,2} + \|f(0)\|_{k,2}\\ &+ \int_0^t (\|u_{tt}\|_{k,2} + \|f\|_{k,2})\, ds\Big).\end{aligned} \tag{4.25}$$

Then it follows from the arbitrariness of $z \in \Omega$ that

$$\|\theta(t)\|_{0,\infty} \leq Ch^k\bigg(\|u_t(0)\|_{k,2} + \|f(0)\|_{k,2} + \int_0^t (\|u_{tt}\|_{k,2} + \|f\|_{k,2})\,ds\bigg). \tag{4.26}$$

Hence (4.10) follows. □

It now remains to verify (4.25). But it suffices to show the following result.

Theorem 4.2. *Under assumptions of Theorem* 4.2*, we have for* $k \geq 2$*,*

$$\|\theta_t(t)\| \leq Ch^k\bigg(\|u_t(0)\|_{k,2} + \|f(0)\|_{k,2} + \int_0^t (\|u_{tt}\|_{k,2} + \|f\|_{k,2})\,ds\bigg). \tag{4.27}$$

Proof. Since $\theta(0) = 0$, it follows from differentiating (3.6) that

$$(\Theta_{tt}, \chi) + A(\theta_t, \chi) = (J_{1,t} + J_{3,t}, \chi) \tag{4.28}$$

and then by letting $\chi = \Theta_t$ in (4.28) that

$$\frac{1}{2}\frac{d}{dt}\|\Theta_t\|^2 + A(\theta_t, \theta_t) \leq \|J_{1,t} + J_{3,t}\|\,\|\Theta_t\| - A\bigg(\theta_t, \int_0^t K_h(t-s)\theta_t(s)\,ds\bigg). \tag{4.29}$$

By repeating an argument similar to the proof of Theorem 3.1, we can obtain that

$$\|\Theta_t(t)\| \leq C\bigg(\|\Theta_t(0)\| + \int_0^t \|J_{1,t} + J_{3,t}\|\,ds\bigg). \tag{4.30}$$

It is easy to see by letting $t = 0$ in (3.6) that

$$(\Theta_t(0), \chi) = -(\rho_t(0), \chi);$$

thus, it follows that

$$\|\Theta_t(0)\| \le \|\rho_t(0)\| \le Ch^k \|u_t(0)\|_k. \tag{4.31}$$

Also we see from Lemma 2.5 that

$$J_{1,t} = \rho_{tt} + \int_0^t K_h(t-s)\rho_{tt}(s)\, ds + K_h(t)\rho_t(0), \tag{4.32}$$

and then

$$\int_0^t \|J_{1,t}\|\, ds \le Ch^k \biggl(\|u_t(0)\|_k + \int_0^t \|u_{tt}(s)\|_{k,2}\, ds \biggr). \tag{4.33}$$

For the same reason, we have

$$\begin{aligned} J_{3,t} = &\int_0^t (K(t-s) - K_h(t-s))(f_t - u_{tt})\, ds \\ &+ (K(t) - K_h(t))(f(0) + u_t(0)) \end{aligned}$$

and then

$$\begin{aligned} \int_0^t \|J_{3,t}\|\, ds \le Ch^k \biggl(&\|u_t(0)\|_{k,2} + \|f(0)\|_{k,2} \\ &+ \int_0^t (\|u_{tt}\|_{k,2} + \|f\|_{k,2})\, ds \biggr). \end{aligned} \tag{4.35}$$

Finally we notice that

$$\|\Theta_t(t)\| \ge \|\theta_t(t)\| - C \int_0^t R_a(t-s)\|\theta_t(s)\|\, ds, \tag{4.36}$$

and, hence, Theorem 4.2 follows from substituting (4.31), (4.33) and (4.35) into (4.30) and using (4.36) and Lemma 2.7. □

Remark. The case of $k = 2$ has been proved in [**20**] via the generalized Green function and weighted norm estimates technique, which is very different from that given above.

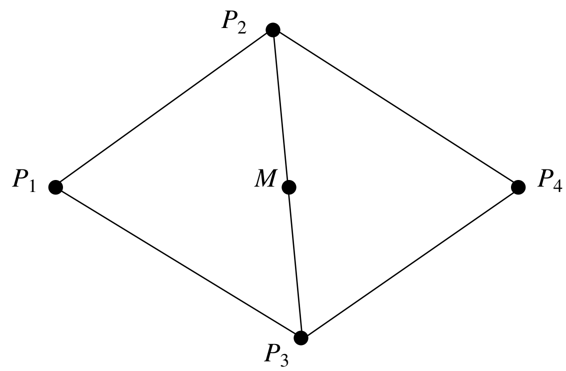

Figure 1.

In the remainder of this section we shall show a stronger maximum norm (without logarithm factor) and superconvergence estimates for piecewise linear element approximations. For this purpose, we require more restrictions on S_h^2. That is, in addition to the quasi-uniform triangulation of Ω, any two adjacent elements form an approximate parallelogram [**17, 36**]. There exists $C > 0$, independent of h, such that (see Figure 1).

$$|\overline{P_1P_2} - \overline{P_3P_4}| \leq Ch^2. \tag{4.37}$$

Theorem 4.3. *Assume that the linear finite element spaces* S_h^2 *satisfies* (4.37) *and that the assumptions of Theorem* 4.1 *for* $k = 2$. *If* $u(t) \in \overset{o}{W}{}_2^1 \cap W_\infty^3(\Omega)$*, then we have*

$$\|u(t) - u_h(t)\|_{0,\infty} = O(h^2) \tag{4.38}$$

and

$$\max_{M \in Q} |\nabla u(M,t) - \overline{\nabla} u_h(M,t))| = O\left(h^2 \log \frac{1}{h}\right), \tag{4.39}$$

where Q *is the set of optimal points of stress, i.e., all middle points of sides of the triangles, and* $\overline{\nabla}$ *is the averaging gradient of two elements at* $x = M$ [**15, 36**].

Proof. First let us show (4.38). Since we know from [**17, 36**] that, under our assumptions on S_h^2, we have

$$\|\rho(t)\|_{0,\infty} = \|u(t) - R_h u(t)\|_{0,\infty} = O(h^2). \tag{4.40}$$

Estimate (4.38) is proved by using (4.40) and Theorem 4.2 for $k = 2$.

Now let us consider (4.39). From [**17, 36**], we see that

$$\max_{M \in Q} |\nabla u(M,t) - \overline{\nabla} R_h u(M,t))| = O\left(h^2 \log \frac{1}{h}\right). \tag{4.41}$$

Thus we need to estimate $\nabla\theta(t)$. Following [**36**], we define

$$\partial_z G^z = \lim_{\Delta z \to 0, \Delta z \| L} \frac{G^{z+\Delta z} - G^z}{|\Delta z|}, \tag{4.42}$$

where G^z is defined in (4.4) and L is any fixed direction in R^2. Also, we have from [**15, 17, 36**]

$$A(\partial_z G^z, \phi) = \partial_z \phi, \quad \phi \in \overset{o}{W}{}_2^1, \tag{4.43}$$

$$A(\partial_z G^z - \partial_z G_h^z, \chi) = 0, \quad \chi \in S_h^2, \tag{4.44}$$

$$\|\partial_z G_h^z\| \le C\left(\log \frac{1}{h}\right)^{1/2}. \tag{4.45}$$

We see now that if we let $\chi = \partial_z G_h^z \in S_h^2$ in (4.23) and use (4.25) and (4.45), we find that

$$\begin{aligned} |\partial_z \theta(z,t)| &= \left(\left\| \int_0^t (K(t-s) - K_h(t-s))(f-u_t)\, ds \right\| \right. \\ &\quad \left. + \left\| e_t + \int_0^t K_h(t-s) e_t(s)\, ds \right\| \right) \|\partial_z G_h^z\| \\ &\le Ch^2 \left(\log \frac{1}{h}\right)^{1/2}. \end{aligned} \tag{4.46}$$

Thus, we have

$$\|\theta(t)\|_{1,\infty} = O\left(h^2 \log \frac{1}{h}\right), \tag{4.47}$$

and, hence, (4.39) follows. □

Remark. We have proved optimal error estimates in L^2 and maximum norm estimates and superconvergence of gradients in two-dimensional spaces. In fact, it can be proved that all results in this paper are valid for the following general equations:

$$u_t + A(t)u = \int_0^t a(t-s)B(t,s)u(s)\,ds + f, \quad \text{in } \Omega \times J, \tag{4.48}$$

with homogeneous boundary conditions and initial data $u(x,0) = v$, where $A(t)$ is a positive definite second order elliptic operator,

$$A(t) = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(a_{ij}(x,t)\frac{\partial}{\partial x_j}\right) + a(x,t)I, \quad a(x,t) \geq 0,$$
$$a_{ij}(x,t) = a_{ji}(x,t), \quad i,j = 1,\dots,d,$$
$$\sum_{i,j=1}^{d} a_{ij}\xi_i\xi_j \geq C_0 \sum_{i=1}^{d} \xi_i^2$$

and B is any second order operator,

$$B(t,s) = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(b_{ij}(x,t,s)\frac{\partial}{\partial x_j}\right) + \sum_{i=1}^{d} b_i(x,t,s)\frac{\partial}{\partial x_i} + b(x,t,s)I,$$

with smooth coefficients in x, t and s. Since the proofs of these results are similar to those given in Sections 2–4, we omit them.

Acknowledgment. The authors would like to thank the referees for their suggestions and comments which led to the improvement of the manuscripts. He would also like to express his appreciation to Professors W. Allegretto and A. Zhou for some fruitful discussions.

REFERENCES

1. W. Allegretto and Y. Lin, *Numerical solutions for a class of differential equations in linear viscoelasticity*, Calcolo **30** (1993), 69–88.

2. W. Allegretto, A. Zhou and Y. Lin, *Long-time stability and error estimates for parabolic integro-differential equations with integrable kernels*, Department of Mathematics, University of Alberta, 1995.

3. H. Brunner and P.J. van der Houwen, *The numerical solution of Volterra equations*, North-Holland, Amsterdam, 1986.

4. H. Brunner, J. Kauthen and A. Ostermann, *Runge-Kutta time discretizations of parabolic Volterra integro-differential equations*, J. Integral Equations Appl., to appear.

5. J.R. Cannon and Y. Lin, *A priori L^2 error estimates for finite element methods for nonlinear diffusion equations with memory*, SIAM J. Numer. Anal. **27** 1990), 595–607.

6. ———, *Non-classical H^1 projection and Galerkin methods for nonlinear parabolic integro-differential equation*, Calcolo **25** (1988), 187–201.

7. C. Chen, V. Thomee and L. Wahlbin, *Finite element approximation of a parabolic integro-differential equations with a weakly singular kernel*, Math. Comp. **58** (1992), 587–602.

8. P.G. Ciarlet, *The finite element method for elliptic problems*, North Holland, Amsterdam, 1978.

9. U. Jin Choi and R.C. MacCamy, *Fractional order Volterra equations*, in *Volterra integrodifferential equations in Banach spaces and applications*, Pitman Res. Notes Math. Ser. **190** (1989), 231–249.

10. Xu Da, *On the discretization in time for a parabolic integrodifferential equation with a weakly singular kernel* I: *Smooth initial data*, Appl. Math. Comp. **58** (1993), 1–27.

11. G. Fairweather, *Galerkin and collocation methods for partial integro-differential equations*, in *Integral equations and inverse problems*, Pitman Res. Notes Math. Ser. **235** (1991), 76–85.

12. C.E. Greenwell-Yanik and G. Fairweather, *Finite element methods for parabolic and hyperbolic partial integro-differential equations*, Nonlinear Anal. **12** (1988), 785–809.

13. U. Hornung and R. Showalter, *Diffusion models for fractures media*, J. Math. Anal. Appl. **147** (1990), 69–80.

14. J.P. Kauthen, *Theoretical and computational aspects of continuous times collocation methods for Volterra-type integral and partial integro-differential equations*, Ph.D. thesis, Universite de Fribourg, 1989.

15. M. Krizek and P. Neittaanmarki, *On superconvergence techniques*, Acta Appl. Math. **9** (1987), 175–198.

16. M.N. LeRoux and V. Thomee, *Numerical solution of semilinear integro-differential equations of parabolic type with nonsmooth data*, SIAM J. Numer. Anal. **26** (1989), 1291–1309.

17. Q. Lin, T. Lu and S. Shen, *Maximum norm estimates, extrapolation and optimal point of stress for the finite element methods on the strongly regular triangulation*, J. Comp. Math. **1** (1983), 376–383.

18. Y. Lin, *Galerkin methods for nonlinear parabolic integro-differential equations with nonlinear boundary conditions*, SIAM J. Numer. Anal. **27** (1990), 608–621.

19. ———, *Numerical solutions for integro-differential equations of parabolic type with weakly singular kernels*, in *Comparison methods and stability theory*, Lecture Notes in Pure Appl. Math. **162** (1994), 261–268.

20. ———, *On maximum norm estimates for Ritz-Volterra projections and applications to some time-dependent problems*, J. Comp. Math. **15** (1997), 159–178.

21. Y. Lin, V. Thomee and L. Wahlbin, *Ritz-Volterra projection onto finite element spaces and applications to integro-differential and related equations*, SIAM J. Numer. Anal. **28** (1991), 1047–1070.

22. J.C. Lopez-Marcos, *A difference scheme for a nonlinear partial integrodifferential equation*, SIAM J. Numer. Anal. **27** (1990), 20–31.

23. J.A. Nitshe, *L_∞-convergence of finite element Galerkin approximations for parabolic problems*, RAIRO **13** (1979), 31–51.

24. A. Pani and T. Peterson, *The effect of numerical quadrature on semidiscrete finite element methods for parabolic integro-differential equations*, SIAM J. Numer. Anal., to appear.

25. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer Verlag, New York, 1983.

26. M. Peszynska, *Finite element approximation of differential equations with convolution term*, Math. Comp. **65** (1996), 1019–1037.

27. R. Rannacher and R. Scott, *Some optimal error estimates for piecewise linear finite element approximations*, Math. Comp. **38** (1982), 1–22.

28. M. Renardy, W.J. Hrusa and J.A. Nohel, *Mathematical problems in viscoelasticity*, Longman Scientific & Technical, England, 1987.

29. J.M. Sanz-Serna, *A numerical method for a partial integro-differential equation*, SIAM J. Numer. Math. **25** (1988), 319–327.

30. A.H. Schatz, V. Thomee and L. Wahlbin, *Maximum norm stability and error estimates in parabolic finite element equations*, Comm. Pure Appl. Math. **33** (1980), 265–304.

31. I.H. Sloan and V. Thomee, *Time discretization of an integro-differential equation of parabolic type*, SIAM J. Numer. Anal. **23** (1986), 1052–1061.

32. V. Thomee, *Galerkin finite element methods for parabolic problems*, Lecture Notes in Math. **1054** (1984).

33. V. Thomee and L. Wahlbin, *Long-time numerical solution of a parabolic equation with memory*, Math. Comp. **62** (1994), 477–496.

34. V. Thomee and N.Y. Zhang, *Error estimates for semi-discrete finite element methods for parabolic integro-differential equations*, Math. Comp. **53** (1989), 121–139.

35. M.F. Wheeler, *A priori L_2 error estimates for Galerkin approximation to parabolic partial differential equations*, SIAM J. Numer. Anal. **19** (1973), 723–759.

36. Qiding Zu and Qun Lin, *Superconvergence theory for finite element methods*, Hunan Scientific Press, Changsha, 1989.

37. Yi Yan and G. Fairweather, *Orthogonal spline collocation methods for some partial integrodifferential equations*, SIAM J. Numer. Anal. **29** (1992), 755–768.

Department of Mathematics, University of Alberta, Edmonton, Alberta, T6G 2G1, Canada
E-mail address: ylin@hilbert.math.ualberta.ca
http://vega.math.ualberta.ca/~ylin/ylin.html/