

## ON FULLY DISCRETE GALERKIN APPROXIMATIONS FOR PARTIAL INTEGRO-DIFFERENTIAL EQUATIONS OF PARABOLIC TYPE

NAI-YING ZHANG

**ABSTRACT.** The subject of this work is the application of fully discrete Galerkin finite element methods to initial-boundary value problems for linear partial integro-differential equations of parabolic type. We investigate numerical schemes based on the Padé discretization with respect to time and associated with certain quadrature formulas to approximate the integral term. A preliminary error estimate is established, which contains a term related to the quadrature rule to be specified. In particular, we consider quadrature rules with sparse quadrature points so as to limit the storage requirements, without sacrificing the order of overall convergence. For the backward Euler scheme, the Crank-Nicolson scheme, and a third-order (1, 2) Padé-type scheme, the specific quadrature rules analyzed are based on the rectangular, the trapezoidal, and Simpson's rule. For all the schemes studied, optimal-order error estimates are obtained in the case that the solution of the problem is smooth enough. Since this is important for our error analysis, we also discuss the regularity of the exact solutions of our equations. High-order regularity results with respect to both space and time are given for the solution of problems with smooth enough data.

### 1. INTRODUCTION

The main purpose of this work is to formulate and study fully discrete Galerkin finite element approximations of solutions of initial-boundary value problems for linear partial integro-differential equations of parabolic type. The emphasis will be on discretization with respect to time.

Let  $\Omega$  be a bounded domain in  $\mathbf{R}^d$  with sufficiently smooth boundary  $\partial\Omega$ , and let  $0 < t^0 < \infty$ . We shall consider equations of the form  $(u_t = \partial u / \partial t, J = (0, t^0])$

---

Received by the editor November 14, 1990 and, in revised form, May 24, 1991.

1991 *Mathematics Subject Classification.* Primary 65M15; Secondary 65M60, 35K05.

*Key words and phrases.* Initial-boundary value problem, parabolic, partial, integro-differential equation, regularity, finite element method, fully discrete, Galerkin approximation, stability, error estimate, quadrature.

This work was partially supported by the Center for Research on Parallel Computation through NSF Cooperative agreement No. CCR-8809615 and by the State of Texas Governor's Energy Office through contract #1059 for Geophysical Parallel Computation Project.

©1993 American Mathematical Society  
0025-5718/93 \$1.00 + \$.25 per page

$$\begin{aligned}
 (1.1) \quad u_t(x, t) + Au(x, t) &= \int_0^t B(t, s)u(x, s) ds + f(x, t) \\
 &\equiv \tilde{B}u(x, t) + f(x, t), \quad (x, t) \in \Omega \times J, \\
 u(x, t) &= 0, \quad (x, t) \in \partial\Omega \times \bar{J}, \\
 u(x, 0) &= u_0(x), \quad x \in \Omega.
 \end{aligned}$$

Here,  $A$  is an elliptic operator of the form

$$A = - \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial}{\partial x_j} \right) + a_0(x)I,$$

where the matrix  $(a_{ij}(x))_{i,j=1}^d$  is symmetric and uniformly positive definite, and  $a_0(x)$  is nonnegative on  $\bar{\Omega}$ . Further,

$$B(t, s) = \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left( b_{ij}(x; t, s) \frac{\partial}{\partial x_j} \right) + \sum_{j=1}^d b_j(x; t, s) \frac{\partial}{\partial x_j} + b_0(x; t, s)I$$

is a partial differential operator of at most second order. We shall assume that the coefficients  $a_{ij}(x)$ ,  $a_0(x)$ ,  $b_{ij}(x; t, s)$ ,  $b_j(x; t, s)$ ,  $b_0(x; t, s)$ , and  $f = f(x, t)$  are real-valued functions, sufficiently smooth for our purposes.

Such problems and variants of them arise in various applications, for instance, in models for heat conduction in materials with memory, the compression of poro-viscoelastic media, reactor dynamics, the compartment model of a double-porosity system, and epidemic phenomena in biology. We refer to [12, 13, 16] for detailed lists of references.

Denote by  $(\cdot, \cdot)$  the standard inner product in  $L_2 = L_2(\Omega)$  and by  $A(\cdot, \cdot)$  and  $B(t, s; \cdot, \cdot)$  the bilinear forms on  $H_0^1 \times H_0^1 = H_0^1(\Omega) \times H_0^1(\Omega)$  corresponding to  $A$  and  $B(t, s)$ , respectively. We write problem (1.1) in variational form as

$$\begin{aligned}
 (u_t, v) + A(u, v) &= \int_0^t B(t, s; u(s), v) ds + (f, v) \\
 &\equiv \tilde{B}(u(t), v) + (f, v), \quad v \in H_0^1, \quad t \in J, \\
 u(0) &= u_0.
 \end{aligned}$$

We shall now turn to Galerkin finite element approximations of problem (1.1). Let  $\{S_h\}$  be a family of finite-dimensional subspaces of  $H_0^1$  parametrized by a small positive parameter  $h$ . We first pose the analogue of the problem above on the subspace  $S_h$  to get a spatially discrete problem

$$\begin{aligned}
 (1.2) \quad (u_{h,t}, \chi) + A(u_h, \chi) &= \int_0^t B(t, s; u_h(s), \chi) ds + (f, \chi), \\
 &\quad \chi \in S_h, \quad t \in \bar{J}, \\
 u_h(0) &= u_{0h} \in S_h.
 \end{aligned}$$

We assume that  $\{S_h\}$  possesses the standard approximation property such that, for some fixed integer  $r \geq 2$ , we have

$$(1.3) \quad \inf_{\chi \in S_h} \{\|u - \chi\| + h\|u - \chi\|_1\} \leq Ch^s \|u\|_s, \quad u \in H_0^1 \cap H^s, \quad 1 \leq s \leq r,$$

where  $\|\cdot\|$  and  $\|\cdot\|_1$  are the usual norms in  $L_2$  and  $H_0^1$ , respectively, and  $H^s = H^s(\Omega)$  is the standard  $L_2$  Sobolev space of order  $s$  with norm  $\|\cdot\|_s$ .

It was shown in Thomée and Zhang [14] that, for the semidiscrete problem (1.2) with properly chosen approximate initial data  $u_{0h}$ , we have the error estimate

$$(1.4) \quad \|u_h(t) - u(t)\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^t \|u_t\|_r ds \right\}, \quad t \in J,$$

which is formally the same as that for the corresponding purely parabolic problem ( $B \equiv 0$ ).

The main purpose of this paper is to further discretize equation (1.2) with respect to time. A natural way to do this is to replace  $u_{h,t}$  by a backward difference quotient and apply a quadrature rule to the integral term. We consider such a scheme first.

Let  $k > 0$  be the stepsize in time and  $t_n = nk$ . Further, let  $\bar{\partial}_t U_h^n = (U_h^n - U_h^{n-1})/k$  and let  $\{\omega_{nj} \mid 0 \leq j < n, t_n \in J\}$  be a family of quadrature weights such that, for  $\varphi^j = \varphi(t_j)$ , we have

$$\sigma^n(\varphi) \equiv \sum_{j=0}^{n-1} \omega_{nj} \varphi^j \approx \int_0^{t_n} \varphi(s) ds, \quad t_n \in J.$$

We then obtain what we shall refer to as a *backward Euler type scheme*,

$$(1.5) \quad (\bar{\partial}_t U_h^n, \chi) + A(U_h^n, \chi) = \sum_{j=0}^{n-1} \omega_{nj} B(t_n, t_j; U_h^j, \chi) + (f(t_n), \chi),$$

$$\chi \in S_h, \quad t_n \in J,$$

$$U_h^0 = u_{0h} \in S_h.$$

A natural candidate for the quadrature formula is the rectangular rule, whose quadrature weights are  $\omega_{nj} = k$ . However, to then calculate  $U_h^n$ , we must use, and thus store, all the previous values of the solution,  $U_h^0, \dots, U_h^{n-1}$ ; hence, a vast amount of memory will be needed. More precisely, to compute  $U_h^n$ ,  $t_n \in J$ , the solution needs to be stored at  $\lfloor t^0/k \rfloor$  time levels. This becomes a major obstacle in practical calculations. Another disadvantage of the rectangular rule is that it requires a large amount of computation. Thus the number of time levels used in the quadrature will be one of our key criteria in choosing quadrature rules in this work. One way to reduce the storage requirement significantly is to employ quadrature formulas with high-order truncation error, so that a larger stepsize, or fewer quadrature points, may be used, without losing the order of accuracy of the scheme. We will propose quadrature rules based on the trapezoidal rule and on Simpson's rule. We shall therefore focus our attention on a class of quadrature rules whose quadrature weights  $\{\omega_{nj}\}$  are dominated by some weights  $\{\omega_j\}$ , i.e.,  $|\omega_{nj}| \leq \omega_j$ ,  $0 \leq j < n$ ,  $t_n \in J$ , with  $\sum_{j=0}^{n-1} \omega_j \leq C$ ,  $t_n \in J$ . This class contains not only the rectangular rule, but also other rules with some special features.

A second way to approximate the solution of problem (1.1) is to apply higher-order discretization in time, so that fewer time steps are taken in the calculation for the same accuracy. As a first example of this, we consider a *Crank-Nicolson*

type scheme,

$$(1.6) \quad (\bar{\partial}_t U_h^n, \chi) + A(\bar{U}_h^n, \chi) = \sum_{j=0}^{n-1} \omega_{nj} B(t_{n-1/2}, t_j; U_h^j, \chi) + (f(t_{n-1/2}), \chi),$$

$$\chi \in S_h, \quad t_n \in J,$$

$$U_h^0 = u_{0h} \in S_h,$$

where  $\bar{U}_h^n = (U_h^n + U_h^{n-1})/2$  and  $t_{n-1/2} = (t_n + t_{n-1})/2$ .

Keeping these two schemes in mind, we move the discussion to a more general setting, in which we use time-discrete schemes based on Padé approximation of order  $p$ . We will establish a preliminary error estimate

$$\|U_h^n - u(t_n)\| \leq C(u)\{h^r + k^p + \|\text{global quadrature error}\|\}, \quad t_n \in J,$$

where the so-called global quadrature error is a term whose order of convergence is determined by that of the basic quadrature error  $q^i(\varphi) = \sigma^i(\varphi) - \int_0^{t_i} \varphi(s) ds$ ,  $i = 1, \dots, n$ . For schemes based on the backward Euler, Crank-Nicolson, and the third-order (1, 2) Padé approximation, we choose appropriate quadrature formulas so that the overall error bound reads

$$\|U_h^n - u(t_n)\| \leq C(u)\{h^r + k^p\}, \quad t_n \in J.$$

The error estimates we obtain, however, will demand high regularity of the solution of (1.1), particularly when using rules with high-order truncation error. For instance, Simpson's rule requires that  $\int_0^t \|u_t\|_r ds$  and  $\int_0^t \|D_t^4 u\| ds$  be finite. Since the regularity of the solution is of such importance for our numerical methods, and since some of the desired high regularity results with respect to both space and time are not available in the literature, we devote some effort to showing such regularity under appropriate conditions on the prescribed data.

The first contribution to the numerical solution of integro-differential equations of parabolic type known to the author was made by Douglas and Jones [6] in the 1960's, using the finite difference method. The analysis of finite element methods for partial integro-differential equations of parabolic type has become an active research area only recently. Yanik and Fairweather [16] studied fully discrete Galerkin finite element approximations to the solutions of a nonlinear partial integro-differential equation whose integral term contains at most first-order derivatives in space.

Sloan and Thomée [10] considered the discretization in time of a general integro-differential equation in an abstract Hilbert space setting, where  $A$  is a selfadjoint positive definite operator and  $B(t, s) = \kappa(t, s)B$ . Here,  $B$  is an operator satisfying  $\|A^{-1}B\varphi\| \leq C\|\varphi\|$ ,  $\varphi \in D(B)$ , independently of time, and  $\kappa(t, s)$  is a scalar function. In order to reduce the memory and computational requirements of these methods, they first proposed the application of quadrature rules with relatively higher-order truncation error. The backward Euler type scheme with a quadrature formula based on the trapezoidal rule, and the Crank-Nicolson type scheme based on Simpson's rule were analyzed in detail.

As we have mentioned before, time-continuous spatially semidiscrete Galerkin approximations to problem (1.1) have been examined by Thomée and Zhang [14]; optimal-order error estimates (1.4) were given. (An alternative proof of this result by means of a nonconventional projection can also be found in Cannon and Lin [4] and in Lin, Thomée, and Wahlbin [8].)

Comprehensive surveys of the development of this subject are given by Thomée [12, 13].

The present work is based in an essential way on the ideas of Sloan and Thomée [10], and may be considered as an attempt to further develop the results obtained there, and to carry over the results obtained in [14] for the semidiscrete problem to completely discrete schemes. The rest of this paper is organized as follows.

Section 2 is devoted to the existence, uniqueness, and regularity of solutions of integro-differential equations of parabolic type, with emphasis on results needed in our analysis of numerical schemes. We show that the solution of the initial-boundary value problem (1.1) has any desired degree of regularity in both space and time, if the prescribed data satisfy the appropriate regularity and compatibility conditions.

In §3, as a preparation, we first introduce a concept called  $E_k$ -stability and present two sufficient conditions for this. We then give a preliminary error estimate for the fully discrete Galerkin approximation. Finally, we present a bound for the global quadrature error which appears in this estimate and will be recalled frequently afterwards.

Based on these results, we study in the last three sections some concrete quadrature formulas. In §4 we analyze backward Euler type schemes. We concentrate on quadrature rules with dominated weights. Several quadrature formulas are presented and analyzed, with emphasis on how to reduce the storage requirement. Section 5 contains our discussion of Crank-Nicolson type schemes. This time a class of quadrature rules using so-called persistent dominated weights is considered. Two quadrature rules are given as examples. Section 6 discusses the third-order subdiagonal Padé discretization. An overall error estimate with a third-order convergence rate in time is obtained for a scheme that employs a modified Simpson's rule to approximate the integral term and uses a starting procedure to calculate the first two time steps of the solution.

## 2. EXISTENCE, UNIQUENESS, AND REGULARITY

The purpose of this section is to show existence, uniqueness, and regularity of the solutions of integro-differential equations of parabolic type, primarily as groundwork for our analysis of numerical methods. A review of the references considering problem (1.1) can be found in [1].

**2.1. Existence and uniqueness.** Let us first define some notation and recall some results for the purely parabolic case ( $B \equiv 0$ ) of (1.1) (cf., e.g., Pazy [9]).

Let  $X$  be a Banach space. We introduce the Banach space  $C(X) = \{u: \bar{J} \rightarrow X \mid u \text{ is continuous}\}$  with norm  $\|u\|_{C(X)} = \sup_{t \in \bar{J}} \|u(t)\|_X$ . For  $\delta \in (0, 1)$ , we let  $C^\delta(X) = \{u: \bar{J} \rightarrow X \mid u \text{ is Hölder continuous with exponent } \delta\}$  with norm

$$\|u\|_{C^\delta(X)} = \|u\|_{C(X)} + \sup_{t \neq s, s, t \in \bar{J}} \frac{\|u(t) - u(s)\|_X}{|t - s|^\delta}.$$

We also let  $C^1(X) = \{u: \bar{J} \rightarrow X \mid u \text{ is differentiable and } u_t \in C(X)\}$  with norm  $\|u\|_{C^1(X)} = \|u\|_{C(X)} + \|u_t\|_{C(X)}$ .

In addition to  $H^s$ , we shall use the space  $\dot{H}^s = \dot{H}^s(\Omega) = \{v \in H^s \mid A^j v = 0 \text{ on } \partial\Omega \text{ for } j < s/2\}$  with norm  $|v|_s = \|A^{s/2} v\|$ . We recall the fact that  $|\cdot|_s$  and  $\|\cdot\|_s$  are equivalent on  $\dot{H}^s$ .

We know that the homogeneous ( $f \equiv 0$ ) purely parabolic ( $B \equiv 0$ ) case of (1.1) has a unique solution  $u(t) = E(t)u_0$  for  $u_0 \in L_2$ . Furthermore, given  $s \geq 0$ , we know that, for all  $u_0 \in \dot{H}^s$ , we have  $E(t)u_0 \in C(\dot{H}^s)$  and

$$(2.1) \quad |E(t)u_0|_s \leq C|u_0|_s, \quad t \geq 0.$$

The following lemma gives a standard existence and regularity result for the purely parabolic case of (1.1).

**Lemma 2.1.** *Let  $0 < \delta < 1$ . If  $f \in C^\delta(L_2)$  and  $u_0 \in \dot{H}^2$ , then the initial-boundary value problem (1.1) with  $B \equiv 0$  has a unique solution*

$$u(t) = E(t)u_0 + \int_0^t E(t-s)f(s) ds \equiv E(t)u_0 + \tilde{E}f(t)$$

such that  $u \in C^1(L_2) \cap C(\dot{H}^2)$  and

$$(2.2) \quad \|u_t(t)\| + |u(t)|_2 \leq C(|u_0|_2 + \|f\|_{C^\delta(L_2)}), \quad t \in \bar{J}.$$

The proof follows §4.3 of Pazy [9].

We now carry the above result over to the integro-differential equation (1.1). For problem (1.1), by the well-known regularity result for elliptic problems that

$$(2.3) \quad \|\varphi\|_{H^2} \leq C\|A\varphi\| \quad \forall \varphi \in \dot{H}^2,$$

we have

$$\|B(t, s)\varphi\| + \|B^*(t, s)\varphi\| \leq C\|A\varphi\| \quad \forall \varphi \in \dot{H}^2, \quad 0 \leq s \leq t \leq t^0,$$

where  $B^*$  is the adjoint of  $B$  with respect to  $L_2$ . A direct consequence is  $\|B(t, s)A^{-1}\| \leq C$  and  $\|A^{-1}B(t, s)\varphi\| \leq C\|\varphi\| \quad \forall \varphi \in \dot{H}^2, \quad 0 \leq s \leq t \leq t^0$ . The above is also true if we replace  $B(t, s)$  by its time derivatives.

**Theorem 2.2.** *If  $u_0 \in \dot{H}^2$  and  $f \in C^\delta(L_2)$  for some  $\delta \in (0, 1)$ , then the problem (1.1) has a unique solution  $u \in C^1(L_2) \cap C(\dot{H}^2)$ . Furthermore,*

$$(2.4) \quad \|u_t(t)\| + |u(t)|_2 \leq C(|u_0|_2 + \|f\|_{C^\delta(L_2)}), \quad t \in \bar{J}.$$

*Proof.* By Duhamel's principle, we may write (1.1) formally as

$$(2.5) \quad u(t) = \int_0^t E(t-s)\tilde{B}u(s) ds + \left( E(t)u_0 + \int_0^t E(t-s)f(s) ds \right) \\ \equiv Ku(t) + F(t).$$

If we can prove that (2.5) has a solution  $u \in C(\dot{H}^2)$ , then  $\tilde{B}u + f \in C^\delta(L_2)$ , and hence, by Lemma 2.1,  $u$  is the unique solution of a purely parabolic equation that has  $\tilde{B}u + f$  as the right-hand side and  $u \in C^1(L_2) \cap C(\dot{H}^2)$ . Hence,  $u$  is also the unique solution of (1.1). Thus, we shall prove that (2.5) has a unique solution  $u \in C(\dot{H}^2)$  and that (2.4) holds. This will be verified by showing that (2.5) is a well-posed Volterra-type equation in the Banach space  $C(\dot{H}^2)$ .

First we notice that, by Lemma 2.1, we have  $F \in C^1(L_2) \cap C(\dot{H}^2)$  and

$$(2.6) \quad |F(t)|_2 \leq C(|u_0|_2 + \|f\|_{C^\delta(L_2)}), \quad t \in \bar{J}.$$

Next we quote from Thomée and Zhang [14] that the operator  $K$  is bounded in  $C(\dot{H}^2)$  and

$$|Ku(t)|_2 \leq C \int_0^t |u(s)|_2 ds, \quad t \in \bar{J}.$$

Therefore, by the standard argument for the existence of a unique solution of a Volterra integral equation, we conclude that (2.5) has a unique solution  $u \in C(\dot{H}^2)$  and, in view of (2.6), that

$$\|u(t)\|_2 \leq C\|F\|_{C(\dot{H}^2)} \leq C(|u_0|_2 + \|f\|_{C^\delta(L_2)}), \quad t \in \bar{J}.$$

Using the integro-differential equation (1.1), we obtain also that, for  $t \in \bar{J}$ ,

$$\begin{aligned} \|u_t(t)\| &\leq \|Au(t)\| + \|\tilde{B}u(t)\| + \|f(t)\| \leq C\|u\|_{C(\dot{H}^2)} + \|f\|_{C(L_2)} \\ &\leq C(|u_0|_2 + \|f\|_{C^\delta(L_2)}). \quad \square \end{aligned}$$

**2.2. Higher-order regularity.** Later in the numerical analysis, we will need higher-order regularity results for the solution. If we assume that, for some  $m > 2$ ,  $B(t, s): \dot{H}^{i+2} \rightarrow \dot{H}^i$ ,  $0 \leq i \leq m - 2$ , is bounded together with a certain number of its derivatives, then by modifying the technique used in the proof of Theorem 2.2, we can conclude that  $u \in C(\dot{H}^{i+2})$  provided that the data  $u_0$  and  $f$  possess certain regularity properties. However, this condition is unnatural, since these spaces involve boundary conditions associated with  $A$ . In general, we can only expect  $B(t, s): H^{i+2} \rightarrow H^i$ . Therefore, we shall derive a higher-order regularity result, which basically only requires the boundedness of  $B(t, s)$  in Sobolev spaces without boundary conditions associated with  $A$ .

**Theorem 2.3.** *Let  $u$  be the solution of the initial-boundary value problem of (1.1) and let  $n \geq 1$  and  $0 < \delta \leq 1$ . Assume that*

$$(2.7a) \quad D_t^j u(0) \in \dot{H}^2 \cap H^{2(n-j)}, \quad 0 \leq j \leq n - 1,$$

and

$$(2.7b) \quad D_t^j f \in C^\delta(L_2) \cap C(H^{2(n-j-1)}), \quad 0 \leq j \leq n - 1.$$

Then

$$D_t^j u \in C(H^{2(n-j)}), \quad 0 \leq j \leq n,$$

and

$$D_t^j u \in C(\dot{H}^2), \quad 0 \leq j \leq n - 1.$$

Furthermore,

$$(2.8) \quad \begin{aligned} \sum_{j=0}^n \|D_t^j u\|_{C(H^{2(n-j)})} &\leq C \sum_{j=0}^{n-1} \|D_t^j u(0)\|_{H^{2(n-j)}} \\ &\quad + C \sum_{j=0}^{n-1} \|D_t^j f\|_{C^\delta(L_2)} + C \sum_{j=0}^{n-1} \|D_t^j f\|_{C(H^{2(n-j-1)})}. \end{aligned}$$

*Proof.* We shall prove this theorem by induction on  $n$ . By Theorem 2.2, we know that the theorem holds for  $n = 1$ . We now assume that it holds for  $n = m$ ,  $m \geq 1$ . We shall prove that the theorem is true for  $n = m + 1$ . Thus, we assume that

$$(2.9a) \quad D_t^j u(0) \in \dot{H}^2 \cap H^{2(m+1-j)}, \quad 0 \leq j \leq m,$$

and

$$(2.9b) \quad D_t^j f \in C^\delta(L_2) \cap C(H^{2(m-j)}), \quad 0 \leq j \leq m.$$

Differentiating equation (1.1) formally, we obtain

$$(2.10) \quad (u_t)_t + Au_t = B(t, t)u + \widetilde{B}_t u + f_t.$$

We shall write this in the same form as equation (1.1), so that we may use the induction hypothesis.

We easily obtain

$$B(t, t)u(t) = \int_0^t B(t, t)u_t(s) ds + B(t, t)u_0.$$

Similarly, after changing the order of integration, we obtain

$$\begin{aligned} \widetilde{B}_t u(t) &= \int_0^t B_t(t, s) \int_0^s u_t(\tau) d\tau ds + \int_0^t B_t(t, s)u_0 ds \\ &= \int_0^t \int_s^t B_t(t, \tau) d\tau u_t(s) ds + \int_0^t B_t(t, s)u_0 ds. \end{aligned}$$

Using the above facts, we find

$$\begin{aligned} (u_t)_t + Au_t &= \int_0^t \left( B(t, t) + \int_s^t B_t(t, \tau) d\tau \right) u_t(s) ds \\ &\quad + \left( B(t, t) + \int_0^t B_t(t, s) ds \right) u_0 + f_t \\ &\equiv \widetilde{B}_1 u_t + B_2 u_0 + f_t \equiv \widetilde{B}_1 u_t + F. \end{aligned}$$

Let us thus consider the integro-differential equation

$$(2.11) \quad \begin{aligned} v_t + Av &= \widetilde{B}_1 v + F, \quad t \in \overline{J}, \\ v(0) &= u_t(0). \end{aligned}$$

Since the operator  $B_1 = B(t, t) + \int_s^t B_t(t, \tau) d\tau$  is a second-order partial differential operator, and since, by our assumption,  $u_t(0) \in \dot{H}^2$  and  $F \in C^\delta(L_2)$ , we conclude by Theorem 2.2 that (2.11) has a unique solution  $v \in C(\dot{H}^2) \cap C^1(L_2)$ . Let  $U(t) = \int_0^t v(s) ds - u_0$ . We find by integrating (2.11) that  $U$  is the unique solution of (1.1). Thus, we obtain immediately that  $U = u$  and  $v = u_t$ , and hence, by (2.9), we have

$$D_t^j v(0) = D_t^{j+1} u(0) \in \dot{H}^2 \cap H^{2(m-j)}, \quad 0 \leq j \leq m-1.$$

Moreover, by the definition of  $F$  and (2.9), we have

$$\begin{aligned} \|D_t^j F\|_{C^\delta(L_2)} &\leq \|D_t^j B_2 u_0\|_{C^\delta(L_2)} + \|D_t^{j+1} f\|_{C^\delta(L_2)} \\ &\leq C \|u_0\|_{\dot{H}^2} + \|D_t^{j+1} f\|_{C^\delta(L_2)}, \quad 0 \leq j \leq m-1, \end{aligned}$$

and

$$\|D_t^j F\|_{C(H^{2(m-j-1)})} \leq C \|u_0\|_{H^{2(m-j)}} + C \|D_t^{j+1} f\|_{C(H^{2(m-j-1)})}, \quad 0 \leq j \leq m-1.$$

Now by using the induction hypothesis, we obtain

$$D_t^j v \in C(H^{2(m-j)}), \quad 0 \leq j \leq m, \quad D_t^j v \in C(\dot{H}^2), \quad 0 \leq j \leq m-1,$$

and further

$$\begin{aligned}
\sum_{j=0}^m \|D_t^{j+1}u\|_{C(H^{2(m-j)})} &= \sum_{j=0}^m \|D_t^j v\|_{C(H^{2(m-j)})} \\
&\leq C \sum_{j=0}^{m-1} \|D_t^j v(0)\|_{H^{2(m-j)}} + C \sum_{j=0}^{m-1} \|D_t^j F\|_{C^\delta(L_2)} \\
(2.12) \quad &+ C \sum_{j=0}^{m-1} \|D_t^j F\|_{C(H^{2(m-j-1)})} \\
&\leq C \sum_{j=0}^m \|D_t^j u(0)\|_{H^{2(m-j)}} + C \sum_{j=1}^m \|D_t^j f\|_{C^\delta(L_2)} \\
&+ C \sum_{j=1}^m \|D_t^j f\|_{C(H^{2(m-j)})}.
\end{aligned}$$

It remains to show that  $u \in C(H^{2(m+1)})$  and to estimate  $\|u\|_{C(H^{2(m+1)})}$ . We shall accomplish this by showing that  $u$  is the solution of a Volterra equation that is well-posed on  $C(H^{2(m+1)})$ . We write the original equation as

$$(2.13) \quad u = A^{-1} \tilde{B}u + A^{-1}(f - u_t).$$

By the regularity result for elliptic problems, we know that  $A^{-1}: H^{2m} \rightarrow H^{2(m+1)}$  is bounded. Thus  $A^{-1}B(t, s)$  is an operator bounded in  $H^{2(m+1)}$ , uniformly for  $0 \leq s \leq t \leq t^0$ . Hence the operator  $A^{-1}\tilde{B}$  defined by  $A^{-1}\tilde{B}u(t) = \int_0^t A^{-1}B(t, s)u(s) ds$  is a Volterra operator in  $C(H^{2(m+1)})$ .

By (2.12), we have  $u_t \in C(H^{2m})$  and by (2.9),  $f \in C(H^{2m})$ , and hence  $A^{-1}u_t \in C(H^{2(m+1)})$  and  $A^{-1}f \in C(H^{2(m+1)})$ . Therefore, (2.13) is a Volterra equation in  $C(H^{2(m+1)})$  and hence

$$\|u\|_{C(H^{2(m+1)})} \leq C\|A^{-1}(u_t - f)\|_{C(H^{2(m+1)})} \leq C\|u_t\|_{C(H^{2m})} + C\|f\|_{C(H^{2m})}.$$

In view of (2.12), this implies that (2.8) holds for  $n = m + 1$ .  $\square$

From equation (1.1) we obtain  $D_t u(0) = -Au_0 + f(0)$ . Differentiating (1.1), we obtain (2.10), and hence

$$D_t^2 u(0) = -AD_t u(0) + B(0, 0)u_0 + f_t(0) = A^2 u_0 - Af(0) + B(0, 0)u_0 + f_t(0).$$

Repeating this process, we can express  $D_t^j u(0)$  in terms of the prescribed data. In doing so, we see that the conditions required by Theorem 2.3 also implicitly contain certain compatibility conditions for the given data at  $t = 0$ .

### 3. TIME DISCRETIZATION, STABILITY, AND PRELIMINARY ERROR ESTIMATE

This section is devoted to time discretization of integro-differential equations of parabolic type. Since we are primarily interested in the discretization of the time variable, we first discuss an abstract parabolic integro-differential equation in a Hilbert space, and then turn to the concrete situation of a partial integro-differential equation of parabolic type in space and time.

**3.1. Discretization in time.** Let  $A_h$  and  $B_h(t, s): S_h \rightarrow S_h$  be defined by

$$(A_h \mu, \chi) = A(\mu, \chi) \quad \text{and} \quad (B_h(t, s) \mu, \chi) = B(t, s; \mu, \chi), \quad \mu, \chi \in S_h.$$

We can thus rewrite the spatially discrete problem (1.2) as

$$(3.1) \quad \begin{aligned} u_{h,t} + A_h u_h &= \int_0^t B_h(t, s) u_h(s) ds + f_h \equiv \widetilde{B}_h u_h(t) + f_h, \quad t \in \overline{J}, \\ u_h(0) &= u_{0h}, \end{aligned}$$

where  $f_h = P_h f$  with  $P_h: L_2 \rightarrow S_h$  being the  $L_2$ -projector.

In regard to both (1.1) and (3.1), we shall thus consider the time discretization of the following problem on a Hilbert space  $H$ :

$$(3.2) \quad \begin{aligned} u_t + Au &= \int_0^t B(t, s) u(s) ds + f(t) \equiv \widetilde{B} u(t) + f(t), \quad t \in J, \\ u(0) &= u_0, \end{aligned}$$

where  $A$  is a selfadjoint, positive definite linear (unbounded) operator in  $H$  with dense domain  $D(A) \subset H$ . We shall assume that  $A$  has a compact inverse. It follows that  $-A$  generates an analytic semigroup  $E(t) = e^{-At}$ .

For our later discussion, it is convenient to introduce the following concept: we say that *the doubly parametrized operator  $B(t, s)$  is dominated by the operator  $A$*  if  $D(A) \subset D(B(t, s)) = D(B^*(t, s)) \subset H$  for all  $0 \leq s \leq t \leq t^0$ , and if there exists a constant  $C$  such that

$$(3.3) \quad \|B(t, s)\varphi\| + \|B^*(t, s)\varphi\| \leq C\|A\varphi\| \quad \forall \varphi \in D(A), \quad 0 \leq s \leq t \leq t^0,$$

where  $B^*(t, s)$  is the adjoint operator of  $B(t, s)$  with respect to the inner product of  $H$ . If  $B(t, s)$  is dominated by  $A$ , one can easily show that

$$\|B(t, s)A^{-1}\| \leq C, \quad 0 \leq s \leq t \leq t^0,$$

and

$$\|A^{-1}B(t, s)\varphi\| \leq C\|\varphi\| \quad \forall \varphi \in D(B(t, s)), \quad 0 \leq s \leq t \leq t^0.$$

We shall assume that  $B(t, s)$  in (3.2) is dominated by  $A$ , together with some of its derivatives with respect to  $t$  and  $s$ .

For problem (1.1), we have already shown that the partial differential operator  $B(t, s)$  and its derivatives with respect to  $t$  and  $s$  are dominated by  $A$ . For the spatially discrete equation (3.1), we have families of operators  $\{A_h\}$  and  $\{B_h(t, s)\}$ . We thus say that a family of operators  $\{B_h(t, s)\}$  is dominated by  $\{A_h\}$  if there exists a constant  $C$  independent of  $h$  such that

$$\|B_h(t, s)\chi\| + \|B_h^*(t, s)\chi\| \leq C\|A_h\chi\| \quad \forall \chi \in S_h, \quad 0 \leq s \leq t \leq t^0,$$

and similarly for time derivatives of  $B_h$ . This implies

$$\|B_h(t, s)A_h^{-1}\| \leq C \quad \text{and} \quad \|A_h^{-1}B_h(t, s)\| \leq C, \quad 0 \leq s \leq t \leq t^0.$$

When  $B(t, s) = \gamma(t, s)A$ , where  $\gamma(t, s)$  is a bounded scalar function, we have that  $B_h(t, s) = \gamma(t, s)A_h$ ; trivially,  $\{B_h\}$  is dominated by  $\{A_h\}$ . Further, when  $B = B(t, s)$  is a first-order partial differential operator, then since

$$\|\chi\|_1^2 \leq C A(\chi, \chi) = C(A_h \chi, \chi) \leq C\|A_h \chi\| \|\chi\|_1,$$

we have

$$(B_h \chi, \mu) = B(\chi, \mu) \leq C\|\chi\|_1 \|\mu\| \leq C\|A_h \chi\| \|\mu\|,$$

and hence

$$\|B_h \chi\| \leq C \|A_h \chi\|, \quad \chi \in S_h.$$

Similarly,

$$\|B_h^* \chi\| \leq C \|A_h \chi\|, \quad \chi \in S_h.$$

It is now obvious that a family of operators of the form

$$B_h(t, s) = \gamma(t, s)A_h + \text{a discrete first-order operator}$$

is dominated by  $\{A_h\}$ .

When  $B$  is a general second-order partial differential operator, it is more difficult to verify that  $\{B_h\}$  is dominated by  $\{A_h\}$ . However, we claim that this is so as long as the standard inverse inequality

$$(3.4) \quad \|\chi\|_1 \leq Ch^{-1} \|\chi\|, \quad \chi \in S_h,$$

holds for the finite element space  $S_h$ . To show this, we first recall a lemma of Thomée and Zhang [14, Lemma 2.1].

**Lemma 3.1.** *Let  $B(t, s; \cdot, \cdot)$  be a bilinear form on  $H_0^1 \times H_0^1$  corresponding to a second-order partial differential operator  $B(t, s)$ . Then*

$$\|B(t, s; g, A_h^{-1} f)\| \leq C \{\|g\| + h\|g\|_1\} \|f\|$$

for  $0 \leq s \leq t \leq \bar{J}$ ,  $\forall f \in L_2$ ,  $g \in H_0^1$ .

With  $\mu, \chi \in S_h$ , the above lemma and (3.4) yield

$$(B_h \chi, \mu) = B(\chi, \mu) \leq C(\|\mu\| + h\|\mu\|_1) \|A_h \chi\| \leq C \|\mu\| \|A_h \chi\|.$$

Since the same argument works for  $B_h^*$ , we conclude that  $\{B_h\}$  is dominated by  $\{A_h\}$ .

Let us recall a time discretization procedure for the corresponding purely parabolic problem of (3.2), i.e., with  $B \equiv 0$ . More details can be found in [2, 3, 11]. Let  $r(z)$  be a rational function approximating the exponential  $e^{-z}$  to order  $p \geq 1$ , i.e., such that

$$(3.5) \quad r(z) = e^{-z} + O(z^{p+1}) \quad \text{for } z \rightarrow 0,$$

and such that

$$(3.6) \quad |r(z)| \leq 1 \quad \text{for } z \geq 0.$$

Let  $\tau_i = \tau_{n,i} \in [0, t_n]$ ,  $i = 1, \dots, m$ , be distinct real numbers, and let  $\{g_i(z)\}_{i=1}^m = \{g_{n,i}(z)\}_{i=1}^m$  be rational functions which are bounded on  $z > 0$ . We consider a scheme of the form

$$(3.7) \quad \begin{aligned} U^n &= r(kA)U^{n-1} + k \sum_{i=1}^m g_i(kA)f(t_n - \tau_i k), \quad t_n \in J, \\ U^0 &= u_0. \end{aligned}$$

By defining  $E_k = r(kA)$  and  $G_k f(t_n) = \sum_{i=1}^m g_i(kA)f(t_n - \tau_i k)$ , we write (3.7) in short form as

$$(3.8) \quad \begin{aligned} U^n &= E_k U^{n-1} + k G_k f(t_n), \quad t_n \in J, \\ U^0 &= u_0. \end{aligned}$$

We shall apply the above time discretization to the integro-differential equation and use various quadrature formulas to approximate the integral term. Let  $t_n^i = t_n - \tau_i k$  and let  $\sigma^{n, \tau_i}$  be a quadrature rule with weights  $\{\omega_{n_j}^{(i)}\}$  such that, for  $\varphi \in C([0, t^0])$  and with  $\varphi^j = \varphi(t_j)$ ,

$$(3.9) \quad \sigma^{n, \tau_i}(\varphi) \equiv \sigma^{n, \tau}(\varphi) = \sum_{j=0}^{n-1} \omega_{n_j}^{(i)} \varphi^j \approx \int_0^{t_n - \tau_i k} \varphi(s) ds.$$

We shall consider the time-discrete scheme

$$(3.10) \quad \begin{aligned} U^n &= E_k U^{n-1} + k G_k \sigma^n(BU) + k G_k f(t_n), & t_n \in J, \\ U^0 &= u_0, \end{aligned}$$

where

$$G_k \sigma^n(BU) \equiv G_k \sigma^n(B(t_n, \cdot)U) = \sum_{i=1}^m g_i(kA) \sum_{j=0}^{n-1} \omega_{n_j}^{(i)} B(t_n^i, t_j) U^j.$$

Note that (3.10) is explicit with respect to the quadrature term.

Now we finally turn to our main consideration, the full discretization of problem (1.1). Applying discretization (3.10) to the semidiscrete equation (3.1), we obtain the following fully discrete Galerkin scheme:

$$(3.11) \quad \begin{aligned} U_h^n &= E_{kh} U_h^{n-1} + k G_{kh} \sigma^n(B_h U) + k G_{kh} P_h f(t_n), & t_n \in J, \\ U_h^0 &= u_{0h}, \end{aligned}$$

where  $E_{kh}$  and  $G_{kh}$  are defined by replacing  $A$  by  $A_h$  in the definitions of  $E_k$  and  $G_k$ , respectively, and  $\sigma^n(B_h U)$  is defined by replacing  $B$  by  $B_h$  in  $\sigma^n(BU)$ .

In this paper, we shall assume that  $g_i(z)$ ,  $i = 1, \dots, m$ , are real fractions. The backward Euler discretization (1.5) is of this form with

$$m = 1, \quad \tau = 0, \quad r(z) = \frac{1}{1+z}, \quad \text{and} \quad g(z) = \frac{1}{1+z},$$

which has order  $p = 1$ . If we choose

$$m = 1, \quad \tau = \frac{1}{2}, \quad r(z) = \frac{1-z/2}{1+z/2}, \quad \text{and} \quad g(z) = \frac{1}{1+z/2},$$

we obtain the Crank-Nicolson discretization (1.6), for which  $p = 2$ .

**3.2. Stability and  $E_k$ -stability.** To study the stability of (3.11), we introduce a concept which we will call  $E_k$ -stability. Let  $\{V^j\}_{j=0}^{n-1}$ ,  $t_n \in J$ , be a sequence in  $D(A)$  and define  $W^n = F_k^n(V)$  iteratively by

$$\begin{aligned} W^n &= E_k W^{n-1} + k G_k \sigma^n(BV), & t_n \in J, \\ W^0 &= 0. \end{aligned}$$

A quadrature formula is called  $E_k$ -stable if there exist nonnegative  $\{\omega_j\}_{j=0}^{n-1}$ ,  $t_n \in J$ , such that  $\sum_{j=0}^{n-1} \omega_j \leq C$  and, for any  $\{V^j\} \subset D(A)$ ,

$$\|F_k^n(V)\| \leq C \sum_{j=0}^{n-1} \omega_j \|V^j\|, \quad t_n \in J.$$

Moreover, we define  $F_{kh}^n$ - and  $E_{kh}$ -stability, respectively, when problem (3.11) is under consideration, with constant  $C$  independent of  $h$ . The following theorem shows the importance of  $E_k$ -stability.

**Theorem 3.2.** *If the quadrature formula defined by (3.9) is  $E_k$ -stable, then scheme (3.10) is stable; that is,*

$$\|U^n\| \leq C\|u_0\| + Ck \sum_{i=1}^m \sum_{j=1}^n \|f(t_j^i)\|, \quad t_n \in J.$$

In the proof we need the known discrete version of Gronwall's Lemma.

**Lemma 3.3.** *Let  $\{\eta_n\}$  be a sequence of nonnegative real numbers satisfying*

$$(3.12) \quad \eta_n \leq \beta_n + \sum_{j=0}^{n-1} \omega_j \eta_j \quad \text{for } n \geq 0,$$

where  $\omega_j \geq 0$  and  $\{\beta_n\}$  is a nondecreasing sequence of nonnegative numbers. Then

$$\eta_n \leq \beta_n \exp\left(\sum_{j=0}^{n-1} \omega_j\right) \quad \text{for } n \geq 1.$$

We give a proof here for the reader's convenience.

*Proof of Lemma 3.3.* Let  $S_n = \sum_{i=0}^{n-1} \omega_i \eta_i$ . It is sufficient to show that

$$(3.13) \quad S_n \leq \beta_n \left( \exp\left(\sum_{i=0}^{n-1} \omega_i\right) - 1 \right), \quad n \geq 1.$$

We shall use induction to prove this. Since  $\eta_0 \leq \beta_0$ , we have

$$S_1 = \omega_0 \eta_0 \leq \omega_0 \beta_0 \leq \beta_0(e^{\omega_0} - 1) \leq \beta_1(e^{\omega_0} - 1)$$

for  $n = 1$ . Assume now that (3.13) holds for  $S_i$ ,  $1 \leq i \leq n$ . To complete the proof we shall prove that it holds for  $S_{n+1}$ . By definition of  $S_n$  and (3.12), we have

$$S_{n+1} - S_n = \omega_n \eta_n \leq \omega_n(\beta_n + S_n),$$

and hence

$$S_{n+1} \leq \omega_n \beta_n + (1 + \omega_n)S_n.$$

By our induction assumption and the monotonicity of  $\beta_n$  we then obtain

$$\begin{aligned} S_{n+1} &\leq \beta_n \left\{ \omega_n + (1 + \omega_n) \left( \exp\left(\sum_{i=0}^{n-1} \omega_i\right) - 1 \right) \right\} \\ &= \beta_n \left\{ (1 + \omega_n) \exp\left(\sum_{i=0}^{n-1} \omega_i\right) - 1 \right\} \\ &\leq \beta_n \left\{ e^{\omega_n} \exp\left(\sum_{i=0}^{n-1} \omega_i\right) - 1 \right\} \leq \beta_{n+1} \left( \exp\left(\sum_{i=0}^n \omega_i\right) - 1 \right), \end{aligned}$$

and hence the proof is complete.  $\square$

*Proof of Theorem 3.2.* The proof is quite straightforward. Let  $U^n = U_1^n + U_2^n$ , where we define  $U_1^n, U_2^n \in D(A)$  by

$$U_1^n = E_k U_1^{n-1} + k G_k \sigma^n(BU), \quad t_n \in J, \quad \text{subject to } U_1^0 = 0,$$

and

$$U_2^n = E_k U_2^{n-1} + k G_k f(t_n), \quad t_n \in J, \quad \text{subject to } U_2^0 = U^0.$$

By (3.6) and some spectral analysis, we have  $\|E_k\| = \|r(kA)\| = \sup_{\lambda>0} |r(k\lambda)| \leq 1$ , and similarly,  $\|G_k f(t_n)\| \leq C \sum_{i=1}^m \|f(t_n^i)\|$ ; hence,

$$\|U_2^n\| \leq \|U_2^{n-1}\| + Ck \sum_{i=1}^m \|f(t_n^i)\| \leq \cdots \leq \|U_0\| + Ck \sum_{i=1}^m \sum_{j=1}^n \|f(t_j^i)\|, \quad t_n \in J.$$

Therefore, by  $U_1^n = F^n(U)$  and the  $E_k$ -stability of this scheme, we obtain

$$\begin{aligned} \|U^n\| &\leq \|U_1^n\| + \|U_2^n\| \\ &\leq \|U_0\| + Ck \sum_{i=1}^m \sum_{j=1}^n \|f(t_j^i)\| + C \sum_{j=0}^{n-1} \omega_j \|U^j\|, \quad t_n \in J, \end{aligned}$$

which leads to the conclusion by using Lemma 3.3.  $\square$

**3.3. Some sufficient conditions for  $E_k$ -stability.** The  $E_k$ -stability of a scheme is important not only to prove the stability of the time discretization, but also to obtain error estimates. We now give some sufficient conditions for a quadrature formula to be  $E_k$ -stable.

We say that a quadrature rule has *dominated quadrature weights*  $\{\omega_{ji}\}$  if there are weights  $\{\omega_i\}$  such that  $|\omega_{ji}| \leq \omega_i$ ,  $0 \leq t_i \leq t_j \leq t^0$ , and  $\sum_{i=0}^{n-1} \omega_i \leq C$ ,  $t_n \in J$ . If the time-stepping is based on the subdiagonal Padé approximation, i.e.,  $r(z) = p(z)/q(z)$ , where both  $p(z)$  and  $q(z)$  are real polynomials with  $\deg p < \deg q$ , then our first theorem shows that the domination of the quadrature weights is sufficient for  $E_k$ -stability.

**Theorem 3.4.** *Let the time-stepping be based on the subdiagonal Padé approximation and accurate of order  $p = 1$ . Assume that  $B(t, s)$  is an operator such that the  $D_i^j B$ ,  $i = 0, 1$ , are dominated by  $A$ . If the quadrature rule defined by (3.9) has dominated weights  $\{\omega_{ni}\}$ , then the quadrature rule is  $E_k$ -stable.*

For subdiagonal Padé approximation, we shall assume throughout this paper that  $|r(z)| < 1$  for  $z > 0$ . A fact that we shall use in the proof is that, in this case, the generated time-stepping procedure is known to have the smoothing property [11], i.e.,

$$(3.14) \quad \|AE_k^n\| \leq C t_n^{-1}, \quad t_n \in J.$$

*Proof of Theorem 3.4.* Without loss of generality, we assume  $m = 1$ . Denote  $g(kA)$  by  $\overline{G}_k$  and  $t_j - \tau k$  by  $t'_j$  for short. To estimate  $F_k^n(V)$ , we split it as

$$\begin{aligned} F_k^n &= k \sum_{j=1}^n E_k^{n-j} \overline{G}_k \sum_{i=0}^{j-1} \omega_{j,i} (B(t'_j, t_i) - B(t'_n, t_i)) V^i \\ &\quad + k \sum_{j=1}^n E_k^{n-j} \overline{G}_k A \sum_{i=0}^{j-1} \omega_{j,i} A^{-1} B(t'_n, t_i) V^i \\ &\equiv \text{I} + \text{II}. \end{aligned}$$

Since  $B_i$  is dominated by  $A$ , we have

$$\|A^{-1}(B(t'_j, t_i) - B(t'_n, t_i))\varphi\| \leq C(t_n - t_j)\|\varphi\|, \quad \varphi \in D(A),$$

and hence, by the smoothing property (3.14) of  $E_k$ , we obtain

$$\begin{aligned} \|\mathbb{I}\| &\leq Ck \sum_{j=1}^n (t_n - t_j)^{-1} \sum_{i=0}^{j-1} \omega_i \|A^{-1}(B(t'_j, t_i) - B(t'_n, t_i))V^i\| \\ &\leq Ck \sum_{j=1}^n (t_n - t_j)^{-1} (t_n - t_j) \sum_{i=0}^{j-1} \omega_i \|V^i\| \leq C \sum_{i=0}^{n-1} \omega_i \|V^i\|. \end{aligned}$$

It remains to estimate  $\mathbb{II}$ . Changing the order of summation, we have

$$\begin{aligned} (3.15) \quad \mathbb{II} &= \sum_{i=1}^{n-1} \left( k \sum_{j=i+1}^n \omega_{j,i} \bar{G}_k E_k^{n-j} A \right) A^{-1} B(t'_n, t_i) V^i \\ &\equiv \sum_{i=1}^n G_{ni} A^{-1} B(t'_n, t_i) V^i. \end{aligned}$$

By spectral analysis, we obtain

$$\|G_{ni}\| \leq C\omega_i \sup_{\lambda>0} \sum_{j=0}^n k\lambda |r(k\lambda)|^{n-j} |g(k\lambda)| \leq C\omega_i \sup_{\lambda>0} \frac{\lambda |g(\lambda)|}{1 - |r(\lambda)|}.$$

Set  $s(\lambda) = \lambda |g(\lambda)| / (1 - |r(\lambda)|)$ . Since  $|r(\lambda)| < 1$  for  $\lambda > 0$ , we need only bound  $s(\lambda)$  as  $\lambda \rightarrow 0$  and  $\lambda \rightarrow +\infty$ . By (3.5), it easily follows that  $|s(\lambda)| \leq C$  as  $\lambda \rightarrow +\infty$ . Furthermore, since  $\deg(p) < \deg(q)$  and  $|\lambda g(\lambda)|$  is bounded, we obtain  $|s(\lambda)| \leq C$  as  $\lambda \rightarrow 0$ . Altogether, we obtain  $\|G_{ni}\| \leq C\omega_i$ , and so

$$\|\mathbb{II}\| \leq C \sum_{i=0}^{n-1} \omega_i \|A^{-1} B(t_n, t_i) V^i\| \leq C \sum_{i=0}^{n-1} \omega_i \|V^i\|. \quad \square$$

Besides what has been discussed above, there are other time-stepping procedures that do not have the smoothing property, for instance, the Crank-Nicolson discretization. Let us consider the class of time discretizations that are *strictly accurate of order  $p = 1$* , i.e.,

$$(3.16) \quad r(z) - 1 = -z \sum_{i=1}^m g_i(z).$$

For more discussion on this, we refer to Thomée [11, Chaps. 7 and 8] and the references therein. For simplicity, we shall restrict ourselves to the case  $m = 1$ .

For a quadrature formula with dominated weights  $\{\omega_i\}$ , if the dominated weights satisfy

$$(3.17) \quad \sum_{j=i+1}^{n-1} |\omega_{j+1,i} - \omega_{ji}| \leq C\omega_i, \quad 0 \leq i < n-1, \quad t_n \in J,$$

we say that the quadrature rule has *persistent dominated quadrature weights*.

**Theorem 3.5.** *Let the time discretization be strictly accurate of order  $p = 1$  and  $m = 1$ . Assume that  $B$  and  $B_t$  are dominated by  $A$ . If the quadrature rule defined by (3.9) has persistent dominated weights  $\{\omega_{nj}\}$ , then the quadrature rule is  $E_k$ -stable.*

*Proof.* In this proof, let us denote  $\sigma^{j,\tau}$  by  $\sigma^{j'}$ , and  $t_j - \tau k$  by  $t'_j$  for short. Since the time discretization is strictly of order 1, we have  $\bar{G}_k = g(kA) = -(E_k - I)A^{-1}/k$ , and therefore, by summation by parts, we obtain

$$\begin{aligned} F_k^n(V) &= k \sum_{j=1}^n E_k^{n-j} \bar{G}_k \sigma^{j'}(BV) = - \sum_{j=1}^n (E_k^{n-j+1} - E_k^{n-j}) A^{-1} \sigma^{j'}(BV) \\ &= - \sum_{j=1}^n E_k^{n-j} (\sigma^{j+1'} - \sigma^{j'}) (A^{-1}BV) + (E_k^n \sigma^{n'}(A^{-1}BV) - \sigma^{1'}(A^{-1}BV)) \\ &\equiv \text{I} + \text{II}. \end{aligned}$$

By the stability of  $E_k^n$ , since  $B$  is dominated by  $A$ , and since the quadrature formula has dominated weights, we obtain immediately

$$\|\text{II}\| \leq \sum_{i=0}^{n-1} \omega_i \|A^{-1}B(t'_n, t_i)V^i\| \leq C \sum_{i=0}^{n-1} \omega_i \|V^i\|.$$

It remains to estimate I. We split I into three terms as follows:

$$\begin{aligned} \text{I} &= - \sum_{j=1}^{n-1} E_k^{n-j} A^{-1} \left( \sum_{i=0}^j \omega_{j+1,i} B(t'_{j+1}, t_i)V^i - \sum_{i=0}^{j-1} \omega_{ji} B(t'_j, t_i)V^i \right) \\ &= - \sum_{j=1}^{n-1} E_k^{n-j} A^{-1} \omega_{j+1,j} B(t'_{j+1}, t_j)V^j \\ &\quad - \sum_{j=1}^{n-1} E_k^{n-j} A^{-1} \left( \sum_{i=0}^{j-1} (\omega_{j+1,i} - \omega_{ji}) B(t'_{j+1}, t_i)V^i \right) \\ &\quad - \sum_{j=1}^{n-1} E_k^{n-j} A^{-1} \left( \sum_{i=0}^{j-1} \omega_{ji} (B(t'_{j+1}, t_i) - B(t'_j, t_i))V^i \right) \\ &\equiv \text{I}_1 + \text{I}_2 + \text{I}_3. \end{aligned}$$

In a manner similar to the estimate of II, we obtain

$$\|\text{I}_1\| \leq C \sum_{j=1}^{n-1} \omega_j \|V^j\|.$$

Since  $B_t$  is also dominated by  $A$ , we have

$$\begin{aligned} \|A^{-1}(B(t'_{j+1}, t_i) - B(t'_j, t_i))\varphi\| &= \left\| \int_{t'_j}^{t'_{j+1}} A^{-1}B_t(\tau, t_i) d\tau \varphi \right\| \\ &\leq Ck \|\varphi\|, \quad t'_j, t'_{j+1} \in J, \quad t_i \in J, \end{aligned}$$

for  $\varphi \in D(A)$ , and hence

$$\|\text{I}_3\| \leq Ck \sum_{j=1}^{n-1} \sum_{i=0}^{j-1} \omega_i \|V^i\| \leq C \sum_{i=0}^{n-2} \omega_i \|V^i\|.$$

Changing the order of summation, and using our assumption that the quadrature rule has persistent dominated weights, and that  $B$  is dominated by  $A$ , we have

$$\begin{aligned} \|I_2\| &= \left\| \sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} E_k^{n-j} (\omega_{j+1,i} - \omega_{ji}) A^{-1} B(t'_{j+1}, t_i) V^i \right\| \\ &\leq C \sum_{i=0}^{n-2} \left( \sum_{j=i+1}^{n-1} |\omega_{j+1,i} - \omega_{ji}| \right) \|V^i\| \leq C \sum_{i=0}^{n-2} \omega_i \|V^i\|. \end{aligned}$$

These estimates lead to

$$\|I\| \leq C \sum_{i=0}^{n-1} \omega_i \|V^i\|. \quad \square$$

Clearly, the above results hold also for  $E_{kh}$ -stability.

**3.4. A preliminary error estimate.** This subsection will prepare us for our later discussion of the error estimate for fully discrete Galerkin approximations. Let us introduce an auxiliary approximate solution  $\widehat{U}_h^n \in S_h$  obtained by applying the discretization method (3.8) to a purely parabolic equation with right-hand side  $(\widetilde{B}u + f)(t)$ , i.e.,

$$\begin{aligned} \widehat{U}_h^n &= E_{kh} \widehat{U}_h^{n-1} + k G_{kh} P_h (\widetilde{B}u + f)(t_n), \quad t_n \in J, \\ \widehat{U}_h^0 &= u_{0h}. \end{aligned}$$

We shall denote the *basic time-stepping error* by

$$(3.18) \quad \widehat{e}^n = \widehat{U}_h^n - u(t_n),$$

which has been well studied in the literature (see Thomée [11] and the references therein).

Denote by  $q^{n,i}(\varphi)$  the *basic quadrature error*, i.e.,

$$q^n(\varphi) = q^{n,\tau_i}(\varphi) = q^{n,\tau}(\varphi) = \sigma^{n,\tau_i}(\varphi) - \int_0^{t_n^i} \varphi(s) ds.$$

We define the local quadrature error

$$q^n(B_h \varphi) = q^{n,\tau_i}(B_h \varphi) = \sigma^{n,\tau_i}(B_h \varphi) - \int_0^{t_n^i} B_h(t_n^i, s) \varphi(s) ds$$

and the global quadrature error

$$\begin{aligned} Q_{kh}^n(\varphi) &= k \sum_{j=1}^n E_{kh}^{n-j} G_{kh} q^j(B_h \varphi) \\ (3.19) \quad &= k \sum_{j=1}^n E_{kh}^{n-j} \sum_{i=1}^m g_i(kA) q^{j,\tau_i}(B_h \varphi). \end{aligned}$$

We shall frequently make use of the elliptic projector  $R_h: H_0^1 \rightarrow S_h$  defined by

$$A(R_h \varphi, \chi) = A(\varphi, \chi), \quad \varphi \in H_0^1, \chi \in S_h,$$

which has the well-known approximation property

$$(3.20) \quad \|(\mathcal{R}_h - I)u\| + h\|(\mathcal{R}_h - I)u\|_1 \leq Ch^s \|u\|_s, \quad 1 \leq s \leq r.$$

For  $\{\varphi^j\}_{j=0}^n$ , we write  $\|\varphi^n\| = \max_{0 \leq j \leq n} \|\varphi^j\|$ .

We shall now give a preliminary error estimate, in which only the basic time-stepping error and the global quadrature error remain to be specified.

**Theorem 3.6.** *Let  $u$  be the solution of the initial-boundary value problem in (1.1), and let  $U_h^n$  be the approximate solution given by the scheme (3.11) using a time-stepping strictly accurate of order  $p = 1$  and a quadrature formula defined by (3.9). Let  $\hat{e}^n$  be the basic time-stepping error and  $Q_{kh}^n$  the global quadrature error defined by (3.18) and (3.19), respectively. Assume  $B_h$  and  $B_{h,t}$  are dominated by  $A_h$ . If the quadrature rule is  $E_{kh}$ -stable, and if*

$$\|u_{0h} - u_0\| \leq Ch^r \|u_0\|_r,$$

then, for  $t_n \in J$ ,

$$\|U_h^n - u(t_n)\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t(s)\|_r ds \right\} + C\|\hat{e}^n\| + C\|Q_{kh}^n(\mathcal{R}_h u)\|.$$

*Proof.* We write

$$(3.21) \quad e^n = (U_h^n - \widehat{U}_h^n) + (\widehat{U}_h^n - u(t_n)) \equiv Z^n + \hat{e}^n,$$

where  $Z^n \in S_h$  is the only term that needs to be estimated. Following Wheeler [15], let  $\theta^n = U_h^n - \mathcal{R}_h u(t_n)$ . Then, by definition, we have

$$\begin{aligned} Z^n &= E_{kh} Z^{n-1} + kG_{kh} \sigma^n(B_h U) - kG_{kh} P_h \widetilde{B} u(t_n) \\ &= E_{kh} Z^{n-1} + kG_{kh} \sigma^n(B_h \theta) + kG_{kh} (\sigma^n(B_h \mathcal{R}_h u) - \widetilde{B}_h \mathcal{R}_h u(t_n)) \\ &\quad + kG_{kh} (\widetilde{B}_h \mathcal{R}_h u(t_n) - P_h \widetilde{B} u(t_n)) \\ &\equiv E_{kh} Z^{n-1} + \sum_{j=1}^3 I_j, \quad t_n \in J, \end{aligned}$$

$$Z^0 = 0.$$

We now split  $Z^n$  further into  $Z^n = Z_1^n + Z_2^n + Z_3^n$ , where  $Z_i^n = E_{kh} Z_i^{n-1} + I_i$ ,  $t_n \in J$ , and  $Z_i^0 = 0$ . By the  $E_{kh}$ -stability of the quadrature rule, we have

$$(3.22) \quad \|Z_1^n\| = \|F_{kh}^n(\theta)\| \leq C \sum_{i=0}^{n-1} \omega_i \|\theta^i\|, \quad t_n \in J.$$

By the definition of the global quadrature error, we have

$$Z_2^n = Q_{kh}^n(\mathcal{R}_h u), \quad t_n \in J.$$

Assuming for a moment that

$$(3.23) \quad \|Z_3^n\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\}, \quad t_n \in J,$$

we obtain

$$\|Z^n\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} + \|Q_{kh}^n(\mathcal{R}_h u)\| + C \sum_{j=0}^{n-1} \omega_j \|\theta^j\|,$$

and hence, for  $t_n \in J$ ,

$$\begin{aligned} \|\theta^n\| &\leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t(s)\|_r \right\} \\ &\quad + \|\hat{\varepsilon}^n\| + \|\|Q_{kh}^n(R_h u)\| + C \sum_{j=0}^{n-1} \omega_j \|\theta^j\|. \end{aligned}$$

The proof will be completed by applying the discrete Gronwall Lemma 3.3.

It remains to prove (3.23). Denote  $\sum_{i=1}^m g_i(kA_h)$  by  $\bar{G}_{kh}$ . We further split  $Z_3^n$  into  $Z_3^n = Z_{31}^n + Z_{32}^n$  such that

$$\begin{aligned} Z_{31}^n &= E_{kh} Z_{31}^{n-1} + k \bar{G}_{kh} (\widetilde{B}_h R_h u(t_n) - P_h \widetilde{B} u(t_n)), \\ Z_{32}^n &= E_{kh} Z_{32}^{n-1} + k (G_{kh} - \bar{G}_{kh}) (\widetilde{B}_h R_h u(t_n) - P_h \widetilde{B} u(t_n)), \\ Z_{31}^0 &= Z_{32}^0 = 0. \end{aligned}$$

Let us first estimate  $Z_{31}^n$ . By iteration we have

$$Z_{31}^n = k \sum_{j=1}^n E_{kh}^{n-j} \bar{G}_{kh} (\widetilde{B}_h R_h u(t_j) - P_h \widetilde{B} u(t_j)).$$

Let  $\widetilde{B}(\varphi(t), \psi) = \int_0^t B(t, s; \varphi(s), \psi) ds$ . For  $\chi \in S_h$ , we then have

$$\begin{aligned} (Z_{31}^n, \chi) &= k \sum_{j=1}^n (\widetilde{B}_h R_h u(t_j) - \widetilde{B} u(t_j), E_{kh}^{n-j} \bar{G}_{kh} \chi) \\ &= k \sum_{j=1}^n \widetilde{B}(\rho(t_j), E_{kh}^{n-j} \bar{G}_{kh} \chi). \end{aligned}$$

Since the discretization is strictly accurate of order  $p = 1$ , we have  $\bar{G}_{kh} = k^{-1} A_h^{-1} (E_{kh} - I)$ ; hence, summation by parts yields

$$\begin{aligned} (Z_{31}^n, \chi) &= \sum_{j=1}^n \widetilde{B}(\rho(t_j), (E_{kh}^{n-j+1} - E_{kh}^{n-j}) A_h^{-1} \chi) \\ &= -\widetilde{B}(\rho(t_n), A_h^{-1} \chi) + \widetilde{B}(\rho(t_1), A_h^{-1} E_{kh}^n \chi) \\ &\quad + \sum_{j=2}^{n-1} \widetilde{B}(\rho(t_{j+1}) - \rho(t_j), A_h^{-1} E_{kh}^{n-j} \chi). \end{aligned}$$

By Lemma 3.1 and (3.20), we obtain

$$\begin{aligned} &|\widetilde{B}(\rho(t_n), A_h^{-1} \chi)| + |\widetilde{B}(\rho(t_1), A_h^{-1} E_{kh}^n \chi)| \\ &\leq C \int_0^{t_n} (\|\rho\| + h \|\rho\|_1) ds \|\chi\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} \|\chi\|. \end{aligned}$$

Similarly, we obtain

$$\begin{aligned}
& \left| \sum_{j=2}^{n-1} \widetilde{B}(\rho(t_{j+1}) - \rho(t_j), A_h^{-1} E_{kh}^{n-j} \chi) \right| \\
& \leq \sum_{j=2}^{n-1} \left| \int_{t_j}^{t_{j+1}} \frac{d}{ds} \widetilde{B}(\rho(s), A_h^{-1} E_{kh}^{n-j} \chi) ds \right| \\
& \leq \sum_{j=1}^{n-1} \left\{ \left| \int_{t_j}^{t_{j+1}} B(s, s; \rho(s), A_h^{-1} E_{kh}^{n-j} \chi) ds \right| \right. \\
& \quad \left. + \left| \int_{t_j}^{t_{j+1}} \widetilde{B}(\rho(s), A_h^{-1} E_{kh}^{n-j} \chi) ds \right| \right\} \\
& \leq C \sum_{j=1}^{n-1} \left\{ \int_{t_j}^{t_{j+1}} (\|\rho(\tau)\| + h\|\rho(\tau)\|_1) d\tau \|E_{kh}^{n-j} \chi\| \right. \\
& \quad \left. + \int_{t_j}^{t_{j+1}} \int_0^s (\|\rho(\tau)\| + h\|\rho(\tau)\|_1) d\tau ds \|E_{kh}^{n-j} \chi\| \right\} \\
& \leq Ch^r \int_0^{t_n} \|u\|_r d\tau \|\chi\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} \|\chi\|.
\end{aligned}$$

Therefore, we obtain

$$(3.24) \quad |(Z_{31}^n, \chi)| \leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} \|\chi\|,$$

which leads to the required bound for  $Z_{31}^n$ .

We now bound  $Z_{32}^n$ . By our definitions, we have

$$(G_{kh} - \overline{G}_{kh})\varphi(t_j) = - \sum_{i=1}^m g_i(kA_h) \int_{t'_j}^{t_j} D_s \varphi(s) ds,$$

and hence, since  $k\|A_h g_i(kA_h)\| \leq C$  and denoting  $T_h = A_h^{-1}$ ,

$$\begin{aligned}
\|(G_{kh} - \overline{G}_{kh})\varphi(t_j)\| & \leq Ck^{-1} \sum_{i=1}^m \int_{t'_j}^{t_j} \|T_h D_s \varphi(s)\| ds \\
& \leq C \max_{0 \leq s \leq t_j} \|T_h D_s \varphi(s)\|.
\end{aligned}$$

By the stability of  $E_{kh}$  and using the above inequality with  $\varphi = \widetilde{B}_h R_h u - P_h \widetilde{B} u$ , we obtain

$$(3.25) \quad \|Z_{32}^n\| \leq C \max_{0 \leq s \leq t_n} \|T_h D_s (\widetilde{B}_h R_h u(s) - P_h \widetilde{B} u(s))\|.$$

For any  $\chi \in S_h$ , we have

$$\begin{aligned}
& (A_h^{-1} D_s (\widetilde{B}_h R_h u(s) - P_h \widetilde{B} u(s)), \chi) \\
& = B(s, s; \rho(s), A_h^{-1} \chi) + \widetilde{B}_t(\rho(s), A_h^{-1} \chi),
\end{aligned}$$

which yields, by Lemma 3.1 and (3.20), that

$$\begin{aligned} & |(T_h D_s(\widetilde{B}_h R_h u(s) - P_h \widetilde{B} u(s)), \chi)| \\ & \leq C \left\{ \|\rho(s)\| + h \|\rho(s)\|_1 + \int_0^s (\|\rho\| + h \|\rho\|_1) d\tau \right\} \|\chi\| \\ & \leq Ch^r \left\{ \|u_0\|_r + \int_0^s \|u_t\|_r d\tau \right\} \|\chi\|. \end{aligned}$$

This and (3.25) lead to the desired bound for  $Z_{32}^n$ .  $\square$

**3.5. A bound for the global quadrature error.** To estimate the global quadrature error, we could use the fact that, by the stability of  $E_{kh}$ ,

$$\|Q_{kh}^n\| \leq \sum_{i=1}^m \|q^{n, \tau_i}(B_h R_h u)\|,$$

and then estimate the local quadrature error instead. For instance, for the rectangular rule,  $\|q^n(\varphi)\| \leq Ck \int_0^{t_n} \|D_s \varphi(s)\| ds$ . This implies, since  $R_h = A_h^{-1} A$  and  $B_h$  is dominated by  $A_h$ , that

$$\|Q_{kh}^n\| \leq Ck \int_0^{t_n} (\|Au\| + \|Au_t\|) ds.$$

However, the regularity of the solution with respect to space required by the above error bound is unnecessarily high. In the following lemma we present a bound for the global quadrature error which leads to an error estimate demanding less regularity of the solution.

Let  $\Phi_h(t, s) = A_h^{-1} B_h P_h$ . If  $B_h$  is dominated by  $A_h$ , then  $\|\Phi_h\| \leq C$ . We shall frequently use the boundedness of this operator and its derivatives.

**Lemma 3.7.** *Let the time-stepping be strictly accurate of order  $p = 1$  and  $m = 1$ , and let  $Q_{kh}^n(\varphi)$  be the global quadrature error defined by (3.19). Assume  $B_h$  and  $B_{h,t}$  are dominated by  $A_h$ . If the quadrature formula is  $E_{kh}$ -stable, then, for  $t_n \in J$ ,*

$$\|Q_{kh}^n(R_h u)\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} + C \sum_{j=0}^{n-1} \|(q^{j+1, \tau} - q^{j, \tau})(\Phi_h u)\|,$$

where  $q^j$  is the basic quadrature error and  $q^{0, \tau} = 0$ .

*Proof.* By the definition of  $Q_{kh}^n(\varphi)$ , after changing the order of summation, we obtain

$$\begin{aligned} (3.26) \quad Q_{kh}^n(\varphi) &= k \sum_{i=1}^n E_{kh}^{n-i} G_{kh} q^i(B_h \varphi) \\ &= k \sum_{i=1}^n E_{kh}^{n-i} \overline{G}_{kh} \sum_{j=0}^{i-1} (q^{j+1, \tau} - q^{j, \tau})(B_h \varphi) \\ &= \sum_{j=0}^{n-1} \left( k \sum_{i=j+1}^n E_{kh}^{n-i} \overline{G}_{kh} A_h \right) (q^{j+1, \tau} - q^{j, \tau})(\Phi_h \varphi), \end{aligned}$$

where, for convenience, we have defined  $q^{0,\tau} = 0$ . Since the discretization is strictly accurate of order 1, we have

$$\left\| k \sum_{i=j+1}^n E_{kh}^{n-i} \bar{G}_{kh} A_h \right\| = \left\| \sum_{i=j+1}^n (E_{kh}^{n-i} - E_{kh}^{n-i+1}) \right\| = \|I - E_{kh}^{n-j}\| \leq C,$$

and hence

$$(3.27) \quad \|Q_{kh}^n(\varphi)\| \leq C \sum_{j=0}^{n-1} \|(q^{j+1,\tau} - q^{j,\tau})(\Phi_h \varphi)\|.$$

It therefore suffices to show

$$(3.28) \quad \|Q_{kh}^n((R_h - P_h)u)\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\}, \quad t_n \in J.$$

Denote  $(R_h - P_h)u(t)$  by  $\varepsilon(t)$ ; we have

$$\|\varepsilon(t)\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^t \|u_t(s)\|_r ds \right\}.$$

Note that by our definitions, we have

$$(3.29) \quad Q_{kh}^n(\varepsilon) \equiv F_{kh}^n(\varepsilon) - k \sum_{j=1}^n E_{kh}^{n-j} G_{kh} \widetilde{B}_h \varepsilon(t_j).$$

Since the quadrature rule is  $E_{kh}$ -stable, we obtain

$$(3.30) \quad \|F_{kh}^n(\varepsilon)\| \leq C \sum_{i=0}^{n-1} \omega_i \|\varepsilon(t_i)\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t(s)\|_r ds \right\}.$$

Similarly to the proof of (3.27), denoting  $t_j - \tau k$  by  $t'_j$ , we have

$$(3.31) \quad \left\| k \sum_{j=1}^n E_k^{n-j} G_{kh} \widetilde{B}_h \varepsilon(t_j) \right\| \leq C \sum_{j=0}^{n-1} \|\widetilde{\Phi}_h \varepsilon(t'_{j+1}) - \widetilde{\Phi}_h \varepsilon(t'_j)\|.$$

Since

$$(3.32) \quad \begin{aligned} \widetilde{\Phi}_h \varepsilon(t'_{j+1}) - \widetilde{\Phi}_h \varepsilon(t'_j) &= \int_{t'_j}^{t'_{j+1}} \Phi_h(t'_{j+1}, s) \varepsilon(s) ds \\ &+ \int_0^{t'_j} (\Phi_h(t'_{j+1}, s) - \Phi_h(t'_j, s)) \varepsilon(s) ds, \end{aligned}$$

we obtain immediately

$$\|\widetilde{\Phi}_h \varepsilon(t'_{j+1}) - \widetilde{\Phi}_h \varepsilon(t'_j)\| \leq C \int_{t'_j}^{t'_{j+1}} \|\varepsilon(s)\| ds + Ck \int_0^{t'_j} \|\varepsilon(s)\| ds,$$

and hence, by (3.31) and (3.32),

$$(3.33) \quad \begin{aligned} \left\| k \sum_{j=1}^n E_k^{n-j} G_{kh} \widetilde{B}_h \varepsilon(t_j) \right\| &\leq C \int_0^{t_n} \|\varepsilon(s)\| ds \\ &\leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\}. \end{aligned}$$

Taking (3.29), (3.30), and (3.33) together completes the proof.  $\square$

4. BACKWARD EULER TYPE SCHEMES

The purpose of this section is to analyze the backward Euler type scheme (1.5) in detail. Various quadrature formulas are presented, with the emphasis on reducing the memory storage requirement.

Let  $S(n)$  be the number of nonzero quadrature weights  $\{\omega_{ni}\}_{i=0}^{n-1}$  used in the quadrature rule  $\sigma^n(\varphi)$ . For each rule discussed, we shall give an upper bound of  $S(n)$  for  $0 \leq t_n \leq t^0$ , denoted by  $S_{\max}$ . This is also the upper bound for the number of the levels at which the solution needs to be stored in calculating the numerical solution  $U_h^n$  on the whole interval  $[0, t^0]$ .

The backward Euler time discretization is strictly accurate of order  $p = 1$ , and the basic time-stepping error is bounded by (see Thomée [11])

$$(4.1) \quad \|\hat{e}^n\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t(s)\|_r ds \right\} + Ck \int_0^{t_n} \|u_{tt}(s)\| ds, \quad t_n \in J.$$

In this section, we shall refer to BE-stability when we mean  $E_k$ -stability for the backward Euler type scheme. Thus, we have:

**Theorem 4.1.** *Let  $u$  be the solution of (1.1) and  $U_h^n$  the solution of the backward Euler type scheme (1.5). Assume that the quadrature rule is BE-stable. If  $\|u_{0h} - u_0\| \leq Ch^r \|u_0\|_r$ , then, for  $t_n \in J$ ,*

$$\begin{aligned} \|U_h^n - u(t_n)\| &\leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} \\ &\quad + Ck \int_0^{t_n} \|u_{tt}\| ds + C \|Q_{kh}^n(R_h u)\|. \end{aligned}$$

Since the smoothing property holds, by Theorem 3.4, we shall give quadrature formulas with dominated weights in order to keep the BE-stability.

**4.1. The rectangular rule.** The simplest quadrature rule that we shall discuss is the *rectangular rule*, i.e.,

$$(4.2) \quad \sigma^n(\varphi) = k \sum_{j=0}^{n-1} \varphi^j.$$

Obviously, this rule has dominated weights, and hence is BE-stable.

**Theorem 4.2.** *Let  $u$  be the solution of (1.1) and  $U_h^n$  the solution of the backward Euler type scheme (1.5) using the rectangular rule (4.2). Assume  $B_h, B_{h,t}^{(1)}, B_{h,s}^{(1)}$ , and  $B_{h,ts}^{(2)}$  are dominated by  $A_h$ . If  $\|u_{0h} - u_0\| \leq Ch^r \|u_0\|_r$ , then, for  $t_n \in J$ ,*

$$\begin{aligned} \|U_h^n - u(t_n)\| &\leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} \\ &\quad + Ck \left\{ \|u_0\| + \int_0^{t_n} (\|u_t\| + \|u_{tt}\|) ds \right\}. \end{aligned}$$

*Proof.* By Theorem 4.1 and Lemma 3.7, we need only estimate the quadrature-related term  $\sum_{j=0}^{n-1} \|(q^{j+1} - q^j)(\Phi_h u)\|$ . We have

$$(4.3) \quad \begin{aligned} (q^{j+1} - q^j)(\Phi_h u) &= (q^{j+1} - q^j)(\Phi_h(t_{j+1}, \cdot)u(\cdot)) \\ &\quad + q^j((\Phi_h(t_{j+1}, \cdot) - \Phi_h(t_j, \cdot))u(\cdot)) \\ &\equiv I_1 + I_2. \end{aligned}$$

Since

$$(q^{j+1} - q^j)(\varphi) = \int_{t_j}^{t_{j+1}} (t_{j+1} - s) D_s \varphi(s) ds$$

and  $B_h$  and  $B_{h,s}^{(1)}$  are dominated by  $A_h$ , we obtain

$$\|I_1\| \leq Ck \int_{t_j}^{t_{j+1}} (\|u(s)\| + \|u_t(s)\|) ds \leq Ck^2 \|u_0\| + Ck \int_{t_j}^{t_{j+1}} \|u_t(s)\| ds.$$

Similarly, since  $B_{h,t}^{(1)}$  and  $B_{h,ts}^{(2)}$  are dominated by  $A_h$ , we obtain

$$\|I_2\| \leq Ck^2 \int_0^{t_j} (\|u(s)\| + \|u_t(s)\|) ds \leq Ck^2 \|u_0\| + Ck^2 \int_0^{t_j} \|u_t(s)\| ds.$$

Therefore, we get

$$\sum_{j=0}^{n-1} \|(q^{j+1} - q^j)(\Phi_h u)\| \leq Ck \left\{ \|u_0\| + \int_0^{t_n} \|u_t(s)\| ds \right\}. \quad \square$$

**4.2. Modified trapezoidal rule.** As we have mentioned before, we may reduce the memory requirement by using a trapezoidal rule based on longer subintervals. We shall discuss a modified trapezoidal rule which is similar to a quadrature formula introduced by Sloan and Thomée [10].

Let  $m_1 = \lfloor k^{-1/2} \rfloor$ , where  $\lfloor x \rfloor$  denotes the largest integer less than or equal to  $x$ , and set  $k_1 = m_1 k$  and  $\bar{t}_j = j k_1$ . We define  $j_n$  to be the largest integer such that  $\bar{t}_{j_n} < t_n$ . We apply the trapezoidal rule with stepsize  $k_1$  on  $[0, \bar{t}_{j_n}]$  and then the rectangular rule with stepsize  $k$  on the remaining part  $[\bar{t}_{j_n}, t_n]$ . More precisely, we introduce the following *modified trapezoidal rule*:

$$(4.4) \quad \sigma^n(\varphi) = \frac{k_1}{2} \sum_{j=1}^{j_n} (\varphi(\bar{t}_j) + \varphi(\bar{t}_{j+1})) + k \sum_{j=j_n m_1}^{n-1} \varphi(t_j) \equiv \sigma_2^n(\varphi) + \sigma_1^n(\varphi).$$

An upper bound of the storage for this rule is given by  $S_{\max} = t^0/m_1 k + m_1$ . Since  $m_1 = O(k^{-1/2})$ , we have  $S_{\max} = O(k^{-1/2})$ .

Let  $\omega_j^1 = k$  and

$$\omega_j^2 = \begin{cases} k_1, & j \equiv 0 \pmod{m_1}, \\ 0, & \text{otherwise.} \end{cases}$$

We define  $\omega_j = \omega_j^2 + \omega_j^1$ , and find easily that  $\sum_{j=0}^{n-1} \omega_j^1 \leq t^0$  and

$$\sum_{j=0}^{n-1} \omega_j^2 \leq \sum_{j=1}^{j_n} k_1 \leq 2\bar{t}_{j_n+1} \leq C t^0.$$

Therefore, the  $\omega_j$  are dominating weights, and hence this rule is BE-stable.

We now give an error estimate for the backward Euler type scheme using the trapezoidal rule. The regularity requirement of this scheme is the same as that for the purely parabolic problem.

**Theorem 4.3.** *Let  $u$  be the solution of (1.1) and  $U_h^n$  the solution of the backward Euler type scheme (1.5) using the trapezoidal rule defined by (4.4). Assume  $B_{h,t}$  and  $B_{h,s}^{(i)}$ ,  $0 \leq i \leq 2$ , are dominated by  $A_h$ . If  $\|u_{0h} - u_0\| \leq Ch^r \|u_0\|_r$ , then, for  $t_n \in J$ ,*

$$\begin{aligned} \|U_h^n - u(t_n)\| &\leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} \\ &\quad + Ck \left\{ \|u_0\| + \int_0^{t_n} (\|u_t\| + \|u_{tt}\|) ds \right\}. \end{aligned}$$

*Proof.* We know that rule (4.4) is BE-stable. By the preliminary error estimate in Theorem 3.6 and Lemma 3.7, we need only prove

$$(4.5) \quad \sum_{j=0}^{n-1} \|(q^{j+1} - q^j)(\Phi_h u)\| \leq Ck \left\{ \|u_0\| + \int_0^{t_n} (\|u_t\| + \|u_{tt}\|) ds \right\}.$$

We consider (4.3) again. Since the step length of the trapezoidal part is bounded by  $Ck_1$ , we easily obtain

$$\|q^n(\varphi)\| \leq Ck \int_0^{t_n} (\|D_s^2 \varphi(s)\| + \|D_s \varphi(s)\|) ds, \quad t_n \in J.$$

Thus,  $I_2$  of (4.3) is bounded by

$$\|I_2\| \leq Ck^2 \int_0^{t_j} (\|u(s)\| + \|u_t(s)\| + \|u_{tt}(s)\|) ds.$$

We further define

$$\begin{aligned} q^n(\varphi) &= \left( \sigma_2^n(\varphi) - \int_0^{\bar{t}_{j_n}} \varphi(s) ds \right) + \left( \sigma_1^n(\varphi) - \int_{\bar{t}_{j_n}}^{t_n} \varphi(s) ds \right) \\ &\equiv q_2^n(\varphi) + q_1^n(\varphi). \end{aligned}$$

Let

$$\psi_2(s) = \begin{cases} (s - \bar{t}_{j-1})(s - \bar{t}_{j-1/2}), & s \in [\bar{t}_{j-1}, \bar{t}_{j-1/2}], \quad j \geq 1, \\ (s - \bar{t}_j)(s - \bar{t}_{j-1/2}), & s \in [\bar{t}_{j-1/2}, \bar{t}_j], \quad j \geq 1, \end{cases}$$

and

$$\psi_1(s) = -(s - t_{j+1}), \quad s \in (t_j, t_{j+1}], \quad j \geq 0;$$

we have

$$q_2^n(\varphi) = \int_0^{\bar{t}_{j_n}} \psi_2(s) D_s^2 \varphi(s) ds \quad \text{and} \quad q_1^n(\varphi) = \int_{\bar{t}_{j_n}}^{t_n} \psi_1(s) D_s \varphi(s) ds.$$

Since

$$(q_2^{n+1} - q_2^n)(\varphi) = \int_{\bar{t}_{j_n}}^{\bar{t}_{j_{n+1}}} \psi_2(s) D_s^2 \varphi(s) ds$$

and

$$(q_1^{n+1} - q_1^n)(\varphi) = \int_{t_n}^{t_{n+1}} \psi_1(s) D_s \varphi(s) ds - \int_{\bar{t}_{j_n}}^{\bar{t}_{j_{n+1}}} \psi_1(s) D_s \varphi(s) ds,$$

we obtain

$$\begin{aligned} \|I_1\| &\leq Ck \int_{t_j}^{t_{j+1}} (\|u(s)\| + \|D_s u(s)\|) ds \\ &\quad + Ck \int_{\bar{t}_{j_j}}^{\bar{t}_{j_{j+1}}} (\|u(s)\| + \|D_s u(s)\| + \|D_s^2 u(s)\|) ds. \end{aligned}$$

Now (4.5) is a direct consequence of the bounds for  $I_1$  and  $I_2$ .  $\square$

**4.3. A modified Simpson's rule.** In the previous subsection, in order to reduce the storage requirement without the loss of overall accuracy, we used the trapezoidal rule with a larger meshsize to approximate the main part of the integral term. Pursuing this idea one step further, we now propose a quadrature formula based on Simpson's rule in order to reduce the number of quadrature nodes even further.

This will be done by first using Simpson's rule on subintervals of length  $O(k^{1/4})$ . The number of such subintervals of  $[0, t_n]$  is  $O(k^{-1/4})$ . The length of the remaining subinterval is at most  $O(k^{1/4})$ . On this remainder, we apply the trapezoidal rules with meshsizes first  $O(k^{1/2})$  and then  $O(k^{3/4})$ . The remaining subinterval is now of length  $O(k^{3/4})$ , and here we use the rectangular rule with meshsize  $k$ . The quadrature error of this combined rule is then  $O(k)$  and the storage requirement is  $O(k^{-1/4})$ .

We now make the above precise. Let  $m_0 = \lfloor k^{-1/4} \rfloor$  and define  $k_i = m_0^{i-1} k$ ,  $1 \leq i \leq 4$ . We shall now describe the choice of the quadrature points  $\{\bar{t}_{nj}\}$  in  $[0, t_n]$ . We shall often denote these by  $\{\bar{t}_j\}$  for short, since the dependence on  $n$  will be clear. First define  $\bar{t}_j = jk_4$ ,  $0 \leq j \leq p_{4n}$ , where  $p_{4n}$  is the largest even integer such that  $\bar{t}_{p_{4n}} < t_n$ . Next, on the remaining subinterval  $[\bar{t}_{p_{4n}}, t_n]$ , whose length is at most  $2k_4$ , we use quadrature points with meshsize  $k_3$ , and thus define  $\bar{t}_j = \bar{t}_{p_{4n}} + (j - p_{4n})k_3$ ,  $p_{4n} < j \leq p_{3n}$ , where  $p_{3n}$  is the largest integer such that  $\bar{t}_{p_{3n}} < t_n$ . We then define the remaining quadrature points in  $[0, t_n]$  by using meshsizes  $k_2$  and  $k_1$  in turn. In this way, we can write the quadrature points by

$$\bar{t}_j = \begin{cases} jk_4, & 0 \leq j \leq p_{4n}, \\ \bar{t}_{p_{4n}} + (j - p_{4n})k_3, & p_{4n} < j \leq p_{3n}, \\ \bar{t}_{p_{3n}} + (j - p_{3n})k_2, & p_{3n} < j \leq p_{2n}, \\ \bar{t}_{p_{2n}} + (j - p_{2n})k_1, & p_{2n} < j \leq p_{1n}, \end{cases}$$

where  $p_{2n}$  and  $p_{1n}$  are the largest integers such that  $\bar{t}_{p_{2n}} < t_n$  and  $\bar{t}_{p_{1n}} \leq t_n$ . Thus, we divide  $[0, t_n]$  as

$$\begin{aligned} [0, t_n] &= \left( \bigcup_{j=1}^{p_{4n}} [\bar{t}_{j-1}, \bar{t}_j] \right) \cup \left( \bigcup_{j=p_{4n}+1}^{p_{3n}} [\bar{t}_{j-1}, \bar{t}_j] \right) \\ &\quad \cup \left( \bigcup_{j=p_{3n}+1}^{p_{2n}} [\bar{t}_{j-1}, \bar{t}_j] \right) \cup \left( \bigcup_{j=p_{2n}+1}^{p_{1n}} [\bar{t}_{j-1}, \bar{t}_j] \right) \\ &\equiv I_4 \cup I_3 \cup I_2 \cup I_1. \end{aligned}$$

We shall use Simpson's rule with stepsize  $k_4$  on  $I_4$ , and the trapezoidal rule with stepsize  $k_3$  on  $I_3$  and with stepsize  $k_2$  on  $I_2$ . On  $I_1$ , we shall use the

rectangular rule with stepsize  $k_1 = k$ . Thus, the *modified Simpson's rule* is defined by

$$\begin{aligned}
 \sigma^n(\varphi) &= \frac{k_4}{3} \sum_{\substack{j=1 \\ j \text{ is odd}}}^{p_{4n}-1} \{\varphi(\bar{t}_{j-1}) + 4\varphi(\bar{t}_j) + \varphi(\bar{t}_{j+1})\} \\
 (4.6) \quad &+ \frac{k_3}{2} \sum_{j=p_{4n}+1}^{p_{3n}} \{\varphi(\bar{t}_{j-1}) + \varphi(\bar{t}_j)\} \\
 &+ \frac{k_2}{2} \sum_{j=p_{3n}+1}^{p_{2n}} \{\varphi(\bar{t}_{j-1}) + \varphi(\bar{t}_j)\} + k \sum_{j=p_{2n}}^{p_{1n}-1} \varphi(\bar{t}_j) \\
 &\equiv \sigma_4^n(\varphi) + \sigma_3^n(\varphi) + \sigma_2^n(\varphi) + \sigma_1^n(\varphi).
 \end{aligned}$$

To give an upper bound of the number of levels that need to be stored, we first notice that the number of quadrature points in  $I_4$  is bounded by  $t^0/k_4$ . Since the length of  $I_3$  is less than  $k_4$ , the number of quadrature points in  $I_3$  is bounded by  $k_4/k_3$ , etc. Thus, we have

$$S_{\max} = t^0/k_4 + k_4/k_3 + k_3/k_2 + k_2/k_1 = t^0/k_4 + 3m_0 = O(k^{-1/4}).$$

The analysis of Simpson's rule is similar to that of the rectangular and trapezoidal rules. We first note that this rule is BE-stable by showing that it has dominated quadrature weights. For  $1 \leq i \leq 4$ , we define

$$\omega_j^i = \begin{cases} k_i, & j \equiv 0 \pmod{m_0^{i-1}}, \\ 0 & \text{otherwise.} \end{cases}$$

Since we have

$$\sum_{j=0}^{n-1} \omega_j^i \leq k_i \frac{n}{m_0^{i-1}} \leq k_i \frac{t^0}{k_i} = t^0$$

and

$$|\omega_{nj}| \leq \sum_{i=1}^4 \omega_j^i \equiv \omega_j,$$

we can choose the  $\omega_j$  thus defined as the dominating weights.

Moreover, since the difference of  $q^{j+1}(\varphi)$  and  $q^j(\varphi)$  occurs only on

$$\bigcup_{i=1}^4 [\bar{t}_{p_{i,j}}, \bar{t}_{p_{i,j+1}}],$$

we obtain

$$\|(q^{j+1} - q^j)(\varphi)\| \leq Ck \sum_{i=1}^4 \int_{\bar{t}_{p_{i,j}}}^{\bar{t}_{p_{i,j+1}}} \|D_s^i \varphi(s)\| ds.$$

On the other hand, we easily obtain

$$\|q^j(\varphi)\| \leq Ck^2 \int_0^{t_j} \sum_{i=1}^4 \|D_s^i u(s)\| ds.$$

Therefore, similarly to the proof of Theorem 4.2, we obtain

$$\begin{aligned} \sum_{j=0}^{n-1} \|(q^{j+1} - q^j)(\Phi_h u)\| &\leq Ck \int_0^{t_n} \sum_{i=0}^4 \|D_s^i u(s)\| ds \\ &\leq Ck \left\{ \|u_0\| + \int_0^{t_n} \sum_{i=1}^4 \|D_s^i u(s)\| ds \right\}. \end{aligned}$$

Hence, by Theorem 3.6 and Lemma 3.7, we obtain:

**Theorem 4.4.** *Let  $u$  be the solution of (1.1) and  $U_h^n$  the solution of the backward Euler type scheme (1.5) using the modified Simpson's rule defined by (4.6). Assume  $B_{h,t}$  and  $B_{h,s}^{(i)}$ ,  $0 \leq i \leq 4$ , are dominated by  $A_h$ . If  $\|u_{0h} - u_0\| \leq Ch^r \|u_0\|_r$ , then, for  $t_n \in J$ ,*

$$\begin{aligned} \|U_h^n - u(t_n)\| &\leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} \\ &\quad + Ck \left\{ \|u_0\| + \int_0^{t_n} \sum_{i=1}^4 \|D_s^i u(s)\| ds \right\}. \end{aligned}$$

Note that the application of Simpson's rule requires higher regularity with respect to time than the last two quadrature formulas.

## 5. CRANK-NICOLSON TYPE SCHEMES

In this section we discuss the fully discrete Crank-Nicolson type schemes of (1.6). By Thomée [11], for  $t_n \in J$ , the basic time-stepping error is bounded by

$$\|\hat{e}^n\| \leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} + Ck^2 \int_0^{t_n} (\|u_{ttt}\| + \|Au_{tt}\|) ds.$$

This time, we shall use the term CN-stability to refer to  $E_k$ -stability. Since the Crank-Nicolson time discretization is strictly accurate of order  $p = 2$ , we obtain immediately from Theorem 3.6 the following result.

**Theorem 5.1.** *Let  $u$  be the solution of the initial-boundary value problem in (1.1), and let  $U_h^n$  be the Crank-Nicolson approximate solution defined by (1.6). Let  $Q_{kh}^n(\varphi)$  be defined by (3.19). Assume that  $B_h$  and  $B_{h,t}$  are dominated by  $A_h$ . If the quadrature rule is CN-stable, and if  $\|u_{0h} - u_0\| \leq Ch^r \|u_0\|_r$ , then we have*

$$\begin{aligned} \|U_h^n - u(t_n)\| &\leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} \\ &\quad + Ck^2 \int_0^{t_n} (\|u_{ttt}\| + \|Au_{tt}\|) ds + C \|Q_{kh}^n(R_h u)\|, \quad t_n \in J. \end{aligned}$$

We shall give two quadrature formulas below; both of them have persistent dominated quadrature weights and hence are CN-stable.

**5.1. A modified trapezoidal rule.** The simplest second-order quadrature formula is the trapezoidal rule. We shall apply the standard trapezoidal rule with

meshsize  $k$  on  $[0, t_{n-1}]$  and the rectangular rule on  $[t_{n-1}, t_{n-1/2}]$  to define a *modified trapezoidal rule*:

$$(5.1) \quad \sigma^n(\varphi) = \frac{k}{2} \sum_{j=1}^{n-1} (\varphi(t_j) + \varphi(t_{j-1})) + \frac{k}{2} \varphi(t_{n-1}) \equiv \sigma_2^n(\varphi) + \sigma_1^n(\varphi).$$

Obviously, the storage requirement for this rule is  $S_{\max} \leq t^0/k = O(k^{-1})$ . In the previous section we have seen that the storage requirement of the backward Euler type scheme using the trapezoidal rule is  $O(k^{-1/2})$ . We shall prove that the Crank-Nicolson type scheme using the trapezoidal rule (5.1) is second-order with respect to time, so that a larger time stepsize may be used for the same overall accuracy. Hence, in this respect the storage requirements for these two schemes are of the same order.

We find immediately that this rule has dominating weights  $\omega_j = k$  and

$$\sum_{j=i+1}^{n-1} |\omega_{j+1,i} - \omega_{ji}| = 0, \quad 0 \leq i \leq n-1, \quad t_n \in J.$$

Hence, the quadrature rule (5.1) has persistent dominated weights, which shows the CN-stability of this rule by Theorem 3.5.

We now give the error estimate for the fully discrete Crank-Nicolson type scheme using the modified trapezoidal rule.

**Theorem 5.2.** *Let  $u$  be the solution of (1.1) and  $U_h^n$  the solution of the Crank-Nicolson type scheme (1.6), using the modified trapezoidal rule (5.1). Assume that  $B_h, B_{h,t}$ , and  $B_{h,s}^{(i)}$ ,  $i = 1, 2$ , are dominated by  $A_h$ . If  $\|u_{0h} - u_0\| \leq Ch^r \|u_0\|_r$ , then, for  $t_n \in J$ ,*

$$\begin{aligned} \|U_h^n - u(t_n)\| &\leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} \\ &\quad + Ck^2 \left\{ \|u_0\| + \int_0^{t_n} (\|u_t\| + \|u_{tt}\| + \|Au_{tt}\|) ds \right\}. \end{aligned}$$

*Proof.* By Theorem 5.1 and Lemma 3.7 we need only prove that

$$\sum_{j=0}^{n-1} \|(q^{j,1/2} - q^{j,-1/2})(\Phi_h u)\| \leq Ck^2 \int_0^{t_n} (\|u(s)\| + \|u_t(s)\| + \|u_{tt}(s)\|) ds.$$

We shall start from (4.3). Since this time

$$(q^{j,1/2} - q^{j,-1/2})(\varphi) = k\varphi(t_j) - \int_{t_{j-1/2}}^{t_{j+1/2}} \varphi(s) ds,$$

we obtain

$$\|(q^{j,1/2} - q^{j,-1/2})(\varphi)\| \leq Ck^2 \int_{t_{j-1/2}}^{t_{j+1/2}} \|D_s^2 \varphi(s)\| ds,$$

and hence

$$\|I_1\| \leq Ck^2 \int_{t_{j-1/2}}^{t_{j+1/2}} (\|u(s)\| + \|D_s u(s)\| + \|D_s^2 u(s)\|) ds.$$

We easily obtain

$$\begin{aligned} \|\mathbf{I}_2\| &\leq Ck^3 \int_0^{t_{j-1}} (\|u(s)\| + \|D_s u(s)\| + \|D_s^2 u(s)\|) ds \\ &\quad + Ck^2 \int_{t_{j-1/2}}^{t_{j+1/2}} (\|u(s)\| + \|D_s u(s)\|) ds. \end{aligned}$$

These bounds for  $\mathbf{I}_1$  and  $\mathbf{I}_2$  lead to the desired conclusion.  $\square$

**5.2. A modified Simpson's rule.** The storage requirement for the modified trapezoidal rule defined by (5.1) is  $O(k^{-1})$ . As we did before for the backward Euler type scheme, we may use a quadrature rule with higher-order truncation error on fewer quadrature points to reduce the memory requirement without sacrificing the accuracy. We now present such a quadrature formula based on Simpson's rule.

Let  $m_1 = \lfloor k^{-1/2} \rfloor$  and  $k_1 = m_1 k$ . Define  $j_n$  to be the largest even integer such that  $j_n k_1 < t_n$ . We introduce the quadrature points

$$\bar{t}_j = \bar{t}_{nj} = \begin{cases} jk_1, & 0 \leq j \leq j_n, \\ j_n k_1 + (j - j_n)k, & j_n < j \leq l_n, \end{cases}$$

where  $\bar{t}_{l_n} = t_{n-1}$ . We now apply Simpson's rule with stepsize  $k_1$  on  $[0, \bar{t}_{j_n}]$ , the trapezoidal rule with stepsize  $k$  on  $[\bar{t}_{j_n}, t_{n-1}]$ , and the rectangular rule with stepsize  $k/2$  on  $[t_{n-1}, t_{n-1/2}]$ , i.e., we set

$$\begin{aligned} \sigma^n(\varphi) &= \frac{k_1}{3} \sum_{\substack{j=1 \\ j \text{ odd}}}^{j_n-1} \{\varphi(\bar{t}_{j-1}) + 4\varphi(\bar{t}_j) + \varphi(\bar{t}_{j+1})\} \\ (5.2) \quad &\quad + \frac{k}{2} \sum_{j=j_n}^{l_n-1} (\varphi(\bar{t}_j) + \varphi(\bar{t}_{j+1})) + \frac{k}{2} \varphi(\bar{t}_{l_n}) \\ &\equiv \sigma_3^n(\varphi) + \sigma_2^n(\varphi) + \sigma_1^n(\varphi). \end{aligned}$$

The storage requirement for this rule is

$$S_{\max} \leq t^0/k_1 + k_1/k \leq t^0/m_1 k + m_1 = O(k^{-1/2}).$$

We now show that this quadrature formula has persistent dominated weights, and therefore is CN-stable. It is easy to see that the quadrature rule has dominating weights of the form

$$\omega_j = \begin{cases} Ck_1, & j \equiv 0 \pmod{m_1}, \\ Ck, & \text{otherwise.} \end{cases}$$

We thus need only prove the validity of (3.17). Consider a fixed  $i = 2lm_1 + i_0$ , where  $0 \leq i_0 < 2m_1$ . By the definition of  $\omega_{ji}$ , when  $j \not\equiv 0 \pmod{2m_1}$ , we have  $\omega_{j+1,i} - \omega_{j,i} = 0$  for  $j > i$ . Now let  $j \equiv 0 \pmod{2m_1}$ . If  $j > 2(l+1)m_1$ , both  $\omega_{j+1,i}$  and  $\omega_{j,i}$  are quadrature weights corresponding to the part of  $\sigma^n(\varphi)$  determined by Simpson's rule, and hence  $\omega_{j+1,i} - \omega_{j,i} = 0$ . Since the only remaining  $j > i$  is  $j = j_0 = 2(l+1)m_1$ , and since the quadrature weights are dominated by  $\omega_i$ , we obtain

$$\sum_{j=i+1}^{n-2} |\omega_{j+1,i} - \omega_{j,i}| = |\omega_{j_0+1,i} - \omega_{j_0,i}| \leq 2\omega_i.$$

Following the outline of the proofs of Theorems 4.3 and 4.4, we may conclude that

$$\sum_{j=0}^{n-1} \|(q^{j,1/2} - q^{j,-1/2})(\Phi u)\| \leq Ck^2 \left\{ \|u_0\| + \int_0^{t_n} \sum_{i=1}^4 \|D_i^j u\| ds \right\},$$

and therefore we obtain:

**Theorem 5.3.** *Let  $u$  be the solution of (1.1) and  $U_h^n$  the solution of the Crank-Nicolson type scheme (1.6) using the modified Simpson's rule (5.2). Assume that  $B_h, B_{h,t}$ , and  $B_{h,s}^{(i)}$ ,  $i = 1, \dots, 4$ , are dominated by  $A_h$ . If  $\|u_{0h} - u_0\| \leq Ch^r \|u_0\|_r$ , then, for  $t_n \in J$ ,*

$$\begin{aligned} \|U_h^n - u(t_n)\| &\leq Ch^r \left\{ \|u_0\|_r + \int_0^{t_n} \|u_t\|_r ds \right\} \\ &\quad + Ck^2 \left\{ \|u_0\| + \int_0^{t_n} \left( \sum_{i=1}^4 \|D_i^j u\| + \|Au_t\| \right) ds \right\}. \end{aligned}$$

### 6. A THIRD-ORDER PADÉ APPROXIMATION

In this section, we consider higher-order Padé approximations; in particular, a third-order case.

**6.1. Higher-order Padé type scheme.** Following common practice, we shall call a time discretization *accurate of order  $p$*  if, in addition to (3.5), we have, for  $l = 0, \dots, p - 1$ ,

$$\sum_{i=1}^m \tau_i^l g_i(\lambda) = \frac{l!}{(-\lambda)^{l+1}} \left( e^{-\lambda} - \sum_{j=0}^l \frac{(-\lambda)^j}{j!} \right) + O(\lambda^{p-l}), \quad \lambda \rightarrow 0.$$

Setting

$$\gamma_l(\lambda) = \frac{l!}{(-\lambda)^{l+1}} \left( r(\lambda) - \sum_{j=0}^l \frac{(-\lambda)^j}{j!} \right) - \sum_{i=1}^m \tau_i^l g_i(\lambda),$$

$$(6.1) \quad l = 0, \dots, p - 1,$$

$$\gamma_p(\lambda) = \frac{p!}{(-\lambda)^{p+1}} \left( r(\lambda) - \sum_{j=0}^p \frac{(-\lambda)^j}{j!} \right),$$

we shall say that it is *strictly accurate of order  $p$*  if

$$(6.2) \quad \begin{aligned} \gamma_l(\lambda) &= 0, & l &= 0, \dots, p - 1, \\ \gamma_p(\lambda) &= O(1), & \lambda &\rightarrow 0. \end{aligned}$$

For instance, the backward Euler and Crank-Nicolson discretizations are strictly accurate of order  $p = 1$  and  $p = 2$ , respectively.

Let us consider the subdiagonal Padé discretization that is strictly accurate of order  $p$ ,  $p \geq 1$ . By Brenner, Crouzeix, and Thomée [3], if  $\|u_{0h} - u_0\| \leq$

$Ch^r \|u_0\|_r$ , then we have the error estimate

$$\begin{aligned} \|\hat{e}^n\| \leq Ch^r & \left\{ \|u_0\|_r + t_n \sup_{0 \leq s \leq t_n} \|u_t(s)\|_r \right\} \\ & + Ck^p \left\{ t_n \|u_t^{(p)}(0)\|_2 + (1 + t_n) \int_0^{t_n} \|u_t^{(p+1)}(s)\|_2 ds \right\}. \end{aligned}$$

Therefore, for a scheme using an  $E_{kh}$ -stable quadrature formula, we have by Theorem 3.6 that

$$\|U^n - u(t_n)\| \leq C(u)\{h^r + k^p\} + C\|Q_{kh}^n(R_h u)\|.$$

**6.2. Third-order (1, 2) Padé approximation.** For simplicity, we shall only give a scheme based on the third-order (1, 2) Padé approximation

$$r(z) = -\frac{2z - 6}{z^2 + 4z + 6}.$$

When  $n \leq 2$ , we shall chose  $\tau_1 = 0$ ,  $\tau_2 = 1/2$ , and  $\tau_3 = 1$ , and further

$$g_1(\lambda) = \frac{1}{\lambda^2 + 4\lambda + 6}, \quad g_2(\lambda) = \frac{4}{\lambda^2 + 4\lambda + 6}, \quad g_3(\lambda) = \frac{\lambda + 1}{\lambda^2 + 4\lambda + 6}.$$

When  $n \geq 3$ , we set  $\tau_i = i$ ,  $1 \leq i \leq 3$ , and let

$$g_1(\lambda) = \frac{\lambda + 23/2}{\lambda^2 + 4\lambda + 6}, \quad g_2(\lambda) = \frac{-8}{\lambda^2 + 4\lambda + 6}, \quad g_3 = \frac{5/2}{\lambda^2 + 4\lambda + 6}.$$

Thus, we obtain a time-stepping procedure strictly accurate of order 3.

Since the subdiagonal Padé approximation has the smoothing property, quadrature formulas with dominated weights are  $E_k$ -stable with respect to the above time discretization. Now we shall construct a third-order quadrature formula by means of Simpson's rule. More specifically, denote the largest even integer less than or equal to  $n$  by  $j_n$ . When  $n < 3$ , we use the rectangular rule on  $[0, t_{n-\tau_i}]$ , i.e.,

$$(6.3a) \quad \sigma^{n, \tau_i}(\varphi) = k(n - \tau_i)\varphi(0).$$

When  $n \geq 3$ , we apply Simpson's rule on  $[0, t_{j_n-i}]$  and the trapezoidal rule on  $[t_{j_n-i}, t_{n-i}]$ , i.e.,

$$\begin{aligned} (6.3b) \quad \sigma^{n, \tau_i}(\varphi) = & \frac{k}{3} \sum_{\substack{j=0 \\ j \text{ even}}}^{j_n-i-2} (\varphi(t_j) + 4\varphi(t_{j+1}) + \varphi(t_{j+2})) \\ & + \frac{k}{2}(n - i - j_n)(\varphi(t_{j_n}) + \varphi(t_{n-i})). \end{aligned}$$

Since, clearly, the quadrature weights of (6.3) are dominated by  $Ck$ , the quadrature formula is  $E_k$ -stable. Furthermore, we have

$$\|q_i^{n, \tau_j}(\varphi)\| \leq \begin{cases} C(\varphi)k^2, & n \leq 2, \\ C(\varphi)k^3, & n \geq 3, \end{cases}$$

and, by the stability of  $E_{kh}$ , we have

$$\|Q_{kh}^n(\varphi)\| \leq C \max_{1 \leq i \leq m} \left\{ k \max_{1 \leq j \leq 2} \|q_i^{n, \tau_j}(B_h \varphi)\| + \max_{3 \leq j \leq n} \|q_i^{n, \tau_j}(B_h \varphi)\| \right\}.$$

By  $R_h = A_h^{-1}A$ , and since  $B_h$  and its derivatives are dominated by  $A_h$ , we further obtain

$$\|Q_{kh}^n(R_h u)\| \leq C(u)k^3, \quad t_n \in J.$$

This leads to the following theorem.

**Theorem 6.1.** *Let  $u$  be the solution of the initial-boundary value problem in (1.1) and  $U_h^n$  the approximate solution generated by the third-order (1, 2) subdiagonal Padé-type scheme described above using the modified Simpson's rule given by (6.3). Assume that  $B_{h,t}$  and  $D_s^i B_h(t, s)$ ,  $0 \leq i \leq 3$ , are dominated by  $A_h$ . If  $\|u_{0h} - u_0\| \leq Ch^r \|u_0\|_r$ , then*

$$\|U_h^n - u(t_n)\| \leq C(u)\{h^r + k^3\}, \quad t_n \in J.$$

#### ACKNOWLEDGMENT

This paper is partly based on my Ph.D. dissertation [17]. I thank Professor Vidar Thomée for the guidance and support that he gave to me during my studies at Göteborg. Also I am grateful to Dr. Stig Larsson for numerous stimulating discussions on this work.

#### BIBLIOGRAPHY

1. P. Acquistapace and B. Terreni, *Existence and sharp regularity for linear parabolic nonautonomous integro-differential equations*, Israel J. Math. **53** (1986), 257–303.
2. G. A. Baker, J. H. Bramble, and V. Thomée, *Single step Galerkin approximations for parabolic problems*, Math. Comp. **31** (1977), 818–847.
3. P. Brenner, M. Crouzeix, and V. Thomée, *Single step methods for inhomogeneous linear differential equations in Banach space*, RAIRO Anal. Numér. **16** (1982), 5–26.
4. J. R. Cannon and Y. P. Lin, *A priori  $L^2$  error estimates for Galerkin methods for nonlinear parabolic integrodifferential equations*, manuscript, 1987.
5. M. Crouzeix and V. Thomée, *On the discretization in time of semilinear parabolic equations with nonsmooth initial data*, Math. Comp. **49** (1987), 35–93.
6. J. Douglas, Jr. and B. F. Jones, Jr., *Numerical methods for integro-differential equations of parabolic and hyperbolic types*, Numer. Math. **4** (1962), 96–102.
7. M.-N. Le Roux and V. Thomée, *Numerical solution of semilinear integro-differential equations of parabolic type with nonsmooth data*, SIAM J. Numer. Anal. **26** (1989), 1291–1309.
8. Y. P. Lin, V. Thomée, and L. B. Wahlbin, *Ritz-Galerkin projections to finite element spaces and applications to integro-differential and related equations*, Technical Report 89-40, Mathematical Sciences Institute, Cornell University.
9. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, 1983.
10. I. H. Sloan and V. Thomée, *Time discretization of an integrodifferential equation of parabolic type*, SIAM J. Numer. Anal. **23** (1986), 1052–1061.
11. V. Thomée, *Galerkin finite element methods for parabolic problems*, Lecture Notes in Math., vol. 1054, Springer-Verlag, 1984.
12. —, *On the numerical solution of integro-differential equations of parabolic type*, in *Numerical Mathematics Singapore 1988* (R. P. Agarwal, Y. M. Chow and S. J. Wilson, eds.), Intern. Ser. Numer. Math. **86**, Birkhäuser Verlag, Basel, 1988, pp. 477–493.
13. —, *Numerical solution of integro-differential equations of parabolic type*, Institut Mathématique de Rennes, Université de Rennes I, 1990.
14. V. Thomée and N.-Y. Zhang, *Error estimates for semidiscrete finite element methods for parabolic integro-differential equations*, Math. Comp. **53** (1989), 121–139.

15. M. F. Wheeler, *A priori  $L_2$  error estimates for Galerkin approximations to parabolic partial differential equations*, SIAM J. Numer. Anal. **10** (1973), 723–759.
16. E. G. Yanik and G. Fairweather, *Finite element methods for parabolic and hyperbolic partial integro-differential equations*, Nonlinear Anal. **12** (1988), 785–809.
17. N.-Y. Zhang, *On the discretization in time and space of parabolic integro-differential equations*, Dissertation, Chalmers University of Technology and the University of Göteborg, 1990.

CENTER FOR RESEARCH ON PARALLEL COMPUTATION, RICE UNIVERSITY, P. O. BOX 1892, HOUSTON, TEXAS 77251

*E-mail address*: nyzhang@rice.edu