

## LONG-TIME NUMERICAL SOLUTION OF A PARABOLIC EQUATION WITH MEMORY

VIDAR THOMÉE AND LARS B. WAHLBIN

**ABSTRACT.** Long-time stability and convergence properties of two time-discretization methods for an integro-differential equation of parabolic type are studied. The methods are based on the standard backward Euler and second-order backward differencing methods. The memory term is approximated by a quadrature rule, with emphasis on such rules with reduced computational memory requirements. Discretization of the spatial partial differential operators by the finite element method is also considered.

### 1. INTRODUCTION

In this paper we shall study the long-time approximation of the solution to the initial value problem

$$(1.1) \quad \begin{aligned} u_t + Au &= \int_0^t b(t-s)Bu(s) ds + f(t) \quad \text{for } t \in R_+ = (0, \infty), \\ u(0) &= u_0. \end{aligned}$$

Here,  $A$  is a selfadjoint strictly positive definite linear operator with compact inverse in a real Hilbert space  $H$ ,  $b(t)$  is a scalar function on  $R_+$ ,  $B$  is an operator with  $D(B) \supset D(A)$  such that, with  $(\cdot, \cdot)$  and  $\|\cdot\|$  the inner product and norm in  $H$ , and  $\|v\|_j = \|A^{j/2}v\|$ ,

$$(1.2) \quad |(Bv, w)| \leq c_0 \|v\|_1 \|w\|_1, \quad \text{with } c_0 > 0,$$

and  $f(t)$  is a function from  $R_+$  into  $H$ .

A concrete problem that we have in mind is when  $A$  is an elliptic second-order differential operator with homogeneous Dirichlet boundary conditions and  $B$  a partial differential operator of at most second order. Such problems occur, e.g., in heat conduction with memory. With  $u(x, t)$  denoting temperature, if one assumes the balance law

$$u_t = -c \operatorname{div} q + f$$

and then, for the heat flux  $q$ ,

$$q = -k(x)\nabla u + \int_0^t b(x, t-s)\nabla u(s) ds,$$

---

Received by the editor May 12, 1992.

1991 *Mathematics Subject Classification.* Primary 65M60, 65R20.

Both authors were supported by the National Science Foundation, USA, and the first author also by the Army Research Office through the Mathematical Sciences Institute at Cornell.

an equation essentially of the desired form would follow. Here,  $b \equiv 0$  corresponds to the classical Fourier's Law. We refer to [2, §3] in particular, for further details. Other examples are briefly described in [8], and some further examples referenced in [6].

It is not difficult to see that, under reasonable assumptions on  $u_0$  and  $f$ , problem (1.1) has a unique solution on  $R_+$ . In the general case the relevant a priori estimate obtained from (1.1) by the energy method will depend on Gronwall's lemma, and the bound for  $u(t)$  will grow exponentially with  $t$ .

Here we shall consider the case that  $b$  is exponentially decreasing and that the memory term is dominated by the elliptic term in the sense that, with some  $K$ ,  $\alpha > 0$ ,

$$(1.3) \quad (i) \quad |b(t)| \leq K e^{-\alpha t} \quad \text{and} \quad (ii) \quad \beta = \int_0^\infty |b(s)| ds < 1/c_0.$$

As we shall show in §2, we then have the long-time stability property

$$(1.4) \quad \|u(t)\| \leq e^{-\gamma t} \|u_0\| + \int_0^t e^{-\gamma(t-s)} \|f(s)\| ds \quad \text{for } t \in R_+, \text{ with } \gamma > 0.$$

For the time-discretization of (1.1) we shall consider two basic types of approximations. The first is based on the backward Euler scheme, and the second on the second-order backward difference scheme. In both cases the integral on the right in (1.1) is replaced by a quadrature formula. Restricting our present discussion to the backward Euler case, we introduce a uniform time-step  $k$  and the backward different quotient  $\bar{\partial}U^n = (U^n - U^{n-1})/k$ , where  $U^n$  denotes the approximation to  $u$  at time  $t_n = nk$ . We further introduce a quadrature formula with nonnegative coefficients  $\omega_{nj} = \omega_{nj}(k)$ ,

$$(1.5) \quad \sigma^n(g) = \sum_{j=0}^{n-1} \omega_{nj} g(t_j) \quad \left( \simeq \int_0^{t_n} g(s) ds \right).$$

The backward Euler time discrete version of (1.1) that we shall consider is then, with  $b_n(s) = b(t_n - s)$  and  $f^n = f(t_n)$ ,

$$(1.6) \quad \begin{aligned} \bar{\partial}U^n + AU^n &= \sigma^n(b_n BU) + f^n \\ &= \sum_{j=0}^{n-1} \omega_{nj} b(t_n - t_j) BU^j + f^n \quad \text{for } n \geq 1, \quad U^0 = u_0. \end{aligned}$$

Note that  $U^n$  does not enter on the right of (1.6), which is convenient in implementation.

Error bounds for problems of this nature on finite time intervals have been given in, e.g., [5, 8, and 9], see also [3 and 7] for spatially discrete versions of (1.1) when  $A$  and  $B$  are differential operators. As for (1.1), the required stability estimates are derived in these papers (without assumption (1.3)) by means of a discrete version of Gronwall's lemma, which results in bounds that grow exponentially with  $t_n$  and thus render the corresponding error estimates useless for large  $t_n$ .

Our purpose, therefore, is to now try to derive a long-time stability property analogous to (1.4) for the discrete problem and an associated error estimate. It

turns out that this is considerably easier for the case  $B = A$  than in the general case, owing to the possibility of then using separation of variables. The stability estimate in this case now reads

$$(1.7) \quad \|U^n\| \leq e^{-\gamma t_n} \|u_0\| + k \sum_{j=1}^n e^{-\gamma t_{n-j}} \|f^j\| \quad \text{for } n \geq 0, \quad k \leq k_0, \quad \text{with } \gamma > 0,$$

and the corresponding error estimate is

(1.8)

$$\|U^n - u(t_n)\| \leq Ck \int_0^{t_n} e^{-\delta(t_n-s)} (\|u_{tt}(s)\| + \|u(s)\|_q) ds, \quad \text{for } n \geq 0, \quad \text{with } \delta > 0,$$

where

$$\|v\|_q = \sum_{j=0}^q \left\| \frac{d^j}{dt^j} v \right\|_2 = \sum_{j=0}^q \left\| \frac{d^j}{dt^j} Av \right\|.$$

The integer  $q$  depends on the quadrature scheme used. The proofs are given in §3.

In §4 we analyze similarly discretizations of (1.1) in the special case  $B = A$ , which are based on a second-order backward difference approximation of the time derivative. In this case we need to replace the condition  $\beta < 1$  by the slightly stronger  $\beta\Lambda < 1$ , where  $\Lambda$  is a constant which is shown in an Appendix to be smaller than 1.1.

In §5 we return to the discussion of backward Euler-type schemes, now in the case of a general  $B$ . Here we show that if  $b$  is exponentially decreasing, then the stability result of [5] may be improved by a multiplicative exponentially decreasing factor. In particular, under the appropriate assumptions on the quadrature rule, this will show stability as in (1.7) when  $B$  and  $b$  are sufficiently small. The techniques of proof are similar to those in [5] and do not reproduce (1.7) for the case  $B = A$ .

In §6 we will be concerned with application to the case when  $A$  and  $B$  are finite element approximations of spatial second-order partial differential operators.

A particular aspect of the investigations in [5] and [9] is the systematic construction of quadrature formulas (1.5) in which a large number of the coefficients  $\omega_{nj}$  vanish in such a way that only a small portion of the  $U^j$  need to be saved as the time-stepping progresses. This aspect is taken into account throughout the present work, and a discussion of suitable quadrature schemes is included in §7.

Throughout this article,  $C$  will denote constants subject to change each time they occur.

## 2. ON THE CONTINUOUS PROBLEM

In this section we shall show the stability estimate (1.4).

**Proposition 2.1.** *Assume (1.2), (1.3), and let  $u$  satisfy (1.1). Then*

$$\|u(t)\| \leq e^{-\gamma t} \|u_0\| + \int_0^t e^{-\gamma(t-s)} \|f(s)\| ds \quad \text{for } t \in \mathbf{R}_+, \quad \text{with } \gamma > 0.$$

*Proof.* We consider first the case of the homogeneous equation ( $f = 0$ ). We then have, for  $u_\gamma(t) = e^{\gamma t}u(t)$ ,

$$(2.1) \quad u_{\gamma,t} - \gamma u_\gamma + Au_\gamma = \int_0^t b_\gamma(t-s)Bu_\gamma(s) ds, \quad \text{where } b_\gamma(t) = e^{\gamma t}b(t).$$

Now choose  $\gamma > 0$  so small that

$$(2.2) \quad \beta_\gamma = \int_0^\infty |b_\gamma(s)| ds \leq (1 - 2\lambda_0^{-1}\gamma)^{1/2}/c_0,$$

where  $\lambda_0 > 0$  is the smallest eigenvalue of the operator  $A$ . In light of (1.3)(i), this can be accomplished by the dominated convergence theorem. Forming the inner product of the equation (2.1) by  $2u_\gamma(t)$  and using the Cauchy-Schwarz and arithmetic-geometric means inequalities, we have

$$\begin{aligned} \frac{d}{dt}\|u_\gamma\|^2 - 2\gamma\|u_\gamma\|^2 + 2\|u_\gamma\|_1^2 &= 2 \int_0^t b_\gamma(t-s)(Bu_\gamma(s), u_\gamma(t)) ds \\ &\leq \|u_\gamma(t)\|_1^2 + \beta_\gamma c_0^2 \int_0^t |b_\gamma(t-s)| \|u_\gamma(s)\|_1^2 ds. \end{aligned}$$

Since  $\lambda_0\|u\|^2 \leq \|u\|_1^2$ , we thus obtain, after integrating in  $t$  and interchanging the order in the resulting double integral,

$$\begin{aligned} \|u_\gamma\|^2 + (1 - 2\lambda_0^{-1}\gamma) \int_0^t \|u_\gamma\|_1^2 ds \\ \leq \|u_0\|^2 + \beta_\gamma c_0^2 \int_0^t \int_\sigma^t |b_\gamma(s-\sigma)| ds \|u_\gamma(\sigma)\|_1^2 d\sigma \\ \leq \|u_0\|^2 + \beta_\gamma^2 c_0^2 \int_0^t \|u_\gamma(\sigma)\|_1^2 d\sigma. \end{aligned}$$

By (2.2) we obtain that  $\|u_\gamma(t)\| \leq \|u_0\|$ , which concludes the proof for the homogeneous equation. The general case now follows by the following version of Duhamel's principle, where  $E(t)u_0$  denotes the solution of the homogeneous case ( $f = 0$ ) of (1.1).

**Lemma 2.1.** *The solution  $u(t)$  to (1.1) is given by*

$$u(t) = E(t)u_0 + \int_0^t E(t-s)f(s) ds.$$

*Proof.* By linearity and uniqueness, it suffices to show that the second term above, which we now denote  $\tilde{u}(t)$ , satisfies (1.1) with  $u_0 = 0$ . To see this, we note

$$\begin{aligned} \int_0^t b(t-s)B\tilde{u}(s) ds &= \int_0^t b(t-s)B \int_0^s E(s-\sigma)f(\sigma) d\sigma ds \\ &= \int_0^t \int_\sigma^t b(t-s)BE(s-\sigma)f(\sigma) ds d\sigma \\ &= \int_0^t \int_0^{t-s} b(t-s-\sigma)BE(\sigma)f(s) d\sigma ds. \end{aligned}$$

Hence,

$$\begin{aligned} \tilde{u}_t + A\tilde{u} - \int_0^t b(t-s)A\tilde{u}(s) ds \\ = \int_0^t \left[ E_t(t-s) + AE(t-s) - \int_0^{t-s} b(t-s-\sigma)BE(\sigma)d\sigma \right] f(s) ds + f(t) \\ = f(t), \end{aligned}$$

which shows the claim.  $\square$

We end this section by elucidating the conditions (1.3) in the case when  $B = A$  (and hence  $c_0 = 1$  in (1.2)),  $b(t) = Ke^{-\alpha t}$ , and  $f = 0$ . Considering then a Fourier component  $u_\lambda(t) = (u(t), \varphi_\lambda)$  with respect to an orthonormal eigensystem  $\{\lambda, \varphi_\lambda\}$  for  $A$ , and setting  $v(t) = e^{\alpha t}u_\lambda(t)$ , we have

$$v_t + (\lambda - \alpha)v = K\lambda \int_0^t v(s) ds,$$

or, after differentiation,  $v_{tt} + (\lambda - \alpha)v_t = K\lambda v$ , with initial data  $v(0) = u_\lambda(0)$  and  $v_t(0) = (\alpha - \lambda)u_\lambda(0)$ . This equation has exponential solutions of the form  $e^{r_\pm t}$ , and hence  $u_\lambda$  is a combination of  $e^{(r_\pm - \alpha)t}$ , where

$$r_\pm - \alpha = -\frac{\lambda + \alpha}{2} \pm \left( \left( \frac{\lambda + \alpha}{2} \right)^2 + (K - \alpha)\lambda \right)^{1/2}.$$

In general, with the initial conditions above, both  $r_+$  and  $r_-$  will be present in the solution. The roots are always real, and in order for  $u_\lambda$  to decrease exponentially in time, we must have  $K < \alpha$ , i.e.,  $\beta = \int_0^\infty b(s) ds = K/\alpha < 1$ ; for  $K > \alpha$ , blow-up would occur.

### 3. BACKWARD EULER METHODS

In this section we shall derive the stability estimate (1.7) and the error estimate (1.8). Before doing so, we shall demonstrate why the straightforward generalization of the proof of Proposition 2.1 above presents difficulties in the case of a general  $B$  under our assumptions on the quadrature rules employed, even for  $\gamma = 0$ .

Considering thus the case of a general  $B$  in (1.6), we would be aiming at showing the boundedness of  $\|U^n\|$  in the case of the homogeneous equation ( $f = 0$ ). After multiplication of (1.6) by  $2U^n$ , we have

$$\begin{aligned} \bar{\delta}\|U^n\|^2 + 2\|U^n\|_1^2 &\leq 2 \sum_{j=0}^{n-1} \omega_{nj} b_{n-j}(BU^j, U^n) \\ &\leq \|U^n\|_1^2 + c_0^2 \left( \sum_{j=0}^{n-1} \omega_{nj} |b_{n-j}| \|U^j\|_1 \right)^2. \end{aligned}$$

From this it follows easily after multiplication by  $k$ , summation and use of Cauchy-Schwarz's inequality in the last term, and assuming

$$\sigma^n(|b_n|) = \sum_{j=0}^{n-1} \omega_{nj} |b(t_{n-j})| \leq \tilde{\beta},$$

with  $\beta < \tilde{\beta} < 1/c_0$ , that

$$(3.1) \quad \|U^N\|^2 + k \sum_{j=1}^N \|U^j\|_1^2 \leq \|U^0\|^2 + \tilde{\beta} c_0^2 k \sum_{n=1}^N \left( \sum_{j=0}^n \omega_{nj} |b_{n-j}| \|U^j\|_1^2 \right).$$

Changing the order of summation, we find that the last term equals

$$\tilde{\beta} c_0^2 k \sum_{j=0}^N \tilde{\omega}_j \|U^j\|_1^2,$$

where  $\tilde{\omega}_j = \sum_{n=j}^N \omega_{nj} |b_{n-j}|$ . The analogue of the proof of Proposition 2.1 would be to bound the coefficients in this sum by those in the sum on the left in (3.1). This is possible, e.g., for the rectangle rule, using  $\omega_{nj} = k$  for  $j = 0, \dots, n - 1$ , since we would then have  $\tilde{\omega}_j = k \sum_0^{N-1} |b_n| < \tilde{\beta}$ , for small  $k$ . However, letting  $\sigma^n$  be sparse, with  $\omega_{nj} = 0$  for many  $j$  and correspondingly larger for other  $j$ , the coefficients for some  $j$  would be larger than the  $k$  allowed on the left and thus prohibit cancellation. For instance, in the modified trapezoidal rule (see §7)  $\tilde{\omega}_j$  is of order  $k^{-1/2}$  for certain  $j$ , and thus the coefficient of  $\|U^j\|_1^2$  on the right is of order  $k^{1/2} > k$ .

We now state and prove the stability result for the case  $B = A$  under an assumption which may be considered as a discrete analogue of  $\beta_\gamma = \int_0^\infty |b_\gamma(s)| ds < 1$  ( $= 1/c_0$ ), and which will be discussed in more detail in Lemma 3.1 below.

**Proposition 3.1.** *Assume that there are positive  $\tilde{\beta}, \bar{\gamma}$  and  $\bar{k}$  such that with  $b_\gamma(t) = e^{\gamma t} b(t)$ ,  $b_{\gamma,n}(s) = b_\gamma(t_n - s)$ ,*

$$(3.2) \quad \tilde{\beta}_{\gamma,n} \equiv \sigma^n(|b_{\gamma,n}|) \leq \tilde{\beta} < 1 \quad \text{for } \gamma \leq \bar{\gamma}, n \geq 1, k \leq \bar{k}.$$

*Then there are positive constants  $k_0$  and  $\gamma$  such that, for the solution of (1.6) (with  $B = A$ ),*

$$\|U^n\| \leq e^{-\gamma t_n} \|U^0\| + k \sum_{j=1}^n e^{-\gamma t_{n-j}} \|f^j\| \quad \text{for } n \geq 0, k \leq k_0.$$

*Proof.* By linearity we may write the solution to (1.6) as  $U^n = U_0^n + U_1^n + \dots + U_n^n$ , where

$$(3.3) \quad \bar{\partial} U_i^n + A U_i^n = \sigma^n(b_n A U_i) + \delta_{ni} f^n \quad \text{for } n \geq 0, i \geq 1, \text{ with } U_i^0 = \delta_{i0} U^0.$$

The solution of (3.3) is obviously zero for  $n < i$ , and we may think of (3.3) as a homogeneous equation for  $n > i$  with initial data  $U_i^i = (I + kA)^{-1} k f^i$  for  $n = i$  when  $i \geq 1$ . We shall show that

$$(3.4) \quad \|U_i^n\| \leq e^{-\gamma t_{n-i}} \|U_i^i\| \quad \text{for } n \geq i.$$

In view of the choice of  $U_i^i$  for  $i > 0$ , this completes the proof.

With  $\{\lambda, \varphi_\lambda\}$  a discrete orthonormal eigensystem of  $A$ , we have for  $c^n = (e^{\gamma t_n} U_i^n, \varphi_\lambda)$

$$(1 + k\lambda)c^n = e^{\gamma k} c^{n-1} + k\lambda \sigma^n(b_{\gamma,n} c) \quad \text{for } n > i.$$

We thus find that

$$|c^n| \leq (1 + k\lambda)^{-1} (e^{\gamma k} + \lambda k \tilde{\beta}) \max_{j \leq n-1} |c^j| \leq \max_{j \leq n-1} |c^j|,$$

provided  $\gamma$  and  $k$  are so small that  $e^{\gamma k} - 1 \leq \lambda_0 k(1 - \tilde{\beta})$ , where  $\lambda_0$  is the smallest eigenvalue of  $A$ . Since  $c^j = 0$  for  $j < i$ , we have  $|c^n| \leq |c^i|$  from which (3.4) follows by Parseval's identity and our definition of  $c^n$ .  $\square$

To prove the error estimate we assume that the quadrature rule is such that, for some integer  $q$ ,

$$(3.5) \quad \begin{aligned} |q^n(g)| &\equiv \left| \sigma^n(g) - \int_0^{t_n} g(s) ds \right| \\ &\leq Ck^{j/q} \int_0^{t_n} |g(s)|_j ds \quad \text{for } n \geq 1, \quad j = 1, q, \end{aligned}$$

where  $|g(s)|_j = \sum_{i=0}^j |(d^i/dt^i)g(s)|$ . We further assume that, cf. (1.3)(i),

$$(3.6) \quad |b(t)|_q \leq Ce^{-\alpha t}.$$

We first show that assumption (3.2) is then satisfied.

**Lemma 3.1.** *Assume (1.3), (3.5), and (3.6). Then for any  $\tilde{\beta}$  with  $\beta < \tilde{\beta} < 1$  there are positive  $\bar{\gamma}$  and  $\bar{k}$  such that (3.2) holds.*

*Proof.* Note that  $\frac{d}{dt}|g| = (\frac{d}{dt}g)\text{sgn}g$ . We have, using (3.5) with  $j = 1$ , and setting  $\epsilon = 1/q$ ,

$$\begin{aligned} \tilde{\beta}_{\gamma, n} &\leq |q^n(|b_{\gamma, n}|)| + \int_0^{t_n} |b(s)|e^{\gamma s} ds \\ &\leq Ck^\epsilon \int_0^\infty |b(s)e^{\gamma s}|_1 ds + \int_0^\infty |b(s)|e^{\gamma s} ds \equiv I_1 + I_2. \end{aligned}$$

Here,  $I_1$  is bounded if  $\gamma < \alpha$ , by (3.6), and tends to zero with  $k$ . Further,  $I_2$  approaches  $\beta \equiv \int_0^\infty |b(s)| ds < 1$  as  $\gamma$  tends to zero, by dominated convergence.  $\square$

We shall now prove the error estimate (1.8).

**Theorem 3.1.** *Assume (1.3) (with  $c_0 = 1$ ), (3.5), and (3.6). Then there are positive  $C, \delta$ , and  $k_0$  such that, for the solutions of (1.1) and (1.6) (with  $B = A$ ),*

$$(3.7) \quad \|U^n - u(t_n)\| \leq Ck \int_0^{t_n} e^{-\delta(t_n-s)} (\|u_{tt}(s)\| + \|u(s)\|_q) ds \quad \text{for } n \geq 0, \quad k \leq k_0.$$

*Proof.* Setting  $e^n = U^n - u(t_n)$ , we have

$$(3.8) \quad \bar{\partial}e^n + Ae^n = \sigma^n(b_n Ae) - \tau^n + q^n(b_n Au) \quad \text{for } n \geq 1, \quad \text{with } e^0 = 0,$$

where  $\tau^n = \bar{\partial}u(t_n) - u_t(t_n)$ . It is easily seen that

$$k \sum_{j=1}^n e^{-\gamma t_{n-j}} \|\tau^j\| \leq k \sum_{j=1}^n e^{-\gamma t_{n-j}} \int_{t_{j-1}}^{t_j} \|u_{tt}\| ds \leq ke^{\gamma k} \int_0^{t_n} e^{-\gamma(t_n-s)} \|u_{tt}(s)\| ds.$$

Further, by (3.5) and (3.6),

$$\|q^j(b_j Au)\| \leq Ck \int_0^{t_j} \|b(t_j - s)u(s)\|_q ds \leq Ck \int_0^{t_j} e^{-\alpha(t_j-s)} \|u(s)\|_q ds,$$

and thus, with  $\delta < \delta_0 = \min(\alpha, \gamma)$ ,

$$\begin{aligned}
 k \sum_{j=1}^n e^{-\gamma t_{n-j}} \|q^j(b_j A u)\| &\leq C k^2 \sum_{j=1}^n \int_0^{t_j} e^{-\delta_0(t_n-s)} \|u(s)\|_q ds \\
 &\leq C k \int_0^{t_n} (t_n - s + k) e^{-\delta_0(t_n-s)} \|u(s)\|_q ds \leq C k \int_0^{t_n} e^{-\delta(t_n-s)} \|u(s)\|_q ds.
 \end{aligned}$$

The desired result is now a consequence of Proposition 3.1 applied to (3.8).  $\square$

We remark that the right-hand side of (3.7) may be estimated in terms of data using the techniques of Theorem 2.1 by noting that time derivatives of  $u$  satisfy integro-differential equations of the same type as  $u$ .

#### 4. SECOND-ORDER BACKWARD DIFFERENCING METHODS

In this section we shall treat second-order backward differencing time-discretizations of the type

$$(4.1) \quad \left(\bar{\partial} + \frac{k}{2} \bar{\partial}^2\right) U^n + A U^n = \sigma^n(b_n A U) + f^n \quad \text{for } n \geq 2,$$

with start-up values  $U^0$  and  $U^1$  given by

$$(4.2) \quad U^0 = u_0, \quad \bar{\partial} U^1 + A U^1 = \sigma^1(b_1 A U) + f^1.$$

We shall first give a stability result similar to Proposition 3.1. For this we introduce a quantity  $\Lambda$ , which enters naturally in the proof below, by

$$(4.3) \quad \Lambda = \sup_{\mu > 0} \Lambda(\mu), \quad \Lambda(\mu) = \frac{2\mu}{3 + 2\mu} \sum_{j=1}^{\infty} \left| \frac{r_+^j - r_-^j}{r_+ - r_-} \right|,$$

where  $r_{\pm} = r_{\pm}(\mu) = (2 \pm \sqrt{1 - 2\mu}) / (3 + 2\mu)$  are the zeros of the polynomial

$$(4.4) \quad \psi(r) = r^2 - \frac{4}{3 + 2\mu} r + \frac{1}{3 + 2\mu} = (r - r_+)(r - r_-).$$

As will follow from our Appendix,  $\Lambda < 1.1$ .

**Proposition 4.1.** *Assume that (1.3) (with  $c_0 = 1$ ) and (3.2) hold, and also that  $\Lambda \hat{\beta} < 1$ . Then there exist positive  $C$ ,  $\gamma$ , and  $k_0$  such that, for the solution of (4.1),*

$$\|U^n\| \leq C e^{-\gamma t_n} (\|U^0\| + \|U^1\|) + C k \sum_{j=2}^n e^{-\gamma t_{n-j}} \|f^j\| \quad \text{for } n \geq 0, \quad k \leq k_0.$$

*Proof.* We shall proceed along the lines of §3, and write  $U^n = U_0^n + U_1^n + \dots + U_n^n$ , where

$$\left(\bar{\partial} + \frac{k}{2} \bar{\partial}^2\right) U_i^n + A U_i^n = \sigma^n(b_n A U_i) + \delta_{ni} f^n \quad \text{for } n \geq 2, \quad U_i^n = \delta_{ni} U^n, \quad n = 0, 1.$$

We wish to show that

$$(4.5) \quad \|U_i^n\| \leq C e^{-\gamma t_{n-i}} \|U_i^i\| \quad \text{for } n \geq i.$$

Since  $U_i^i = (\frac{3}{2}I + kA)^{-1} k f^i$  when  $i \geq 2$ , this will complete the proof.

We shall first establish (4.5) with  $\gamma = 0$ . As in the proof of Proposition 3.1, it suffices to consider a Fourier coefficient  $c^n = c_i^n = (U_i^n, \varphi_\lambda)$  which now satisfies, with  $\mu = k\lambda$ ,

$$(4.6) \quad \left(\frac{3}{2} + \mu\right)c^n - 2c^{n-1} + \frac{1}{2}c^{n-2} = \mu\sigma^n(b_n c) \quad \text{for } n \geq \max(i+1, 2),$$

where  $c^i = c_i^i$ ,  $c^{i-1} = c_i^{i-1}$  (with  $c_i^{i-1} = 0$  for  $i \geq 1$  and  $c_0^1 = 0$ ) are given. Since (4.6) has  $\psi$  as its characteristic polynomial, its solution may be represented as

$$(4.7) \quad c^n = \frac{r_+^{n-i+1} - r_-^{n-i+1}}{r_+ - r_-} c^i + \frac{2\mu}{3+2\mu} \sum_{j=i+1}^n \frac{r_+^{n+1-j} - r_-^{n+1-j}}{r_+ - r_-} \sigma^j(b_j c) \equiv J_1 + J_2,$$

for  $n \geq i$ , with the first term on the right replaced by

$$-r_+ r_- (r_+^{n-1} - r_-^{n-1}) / (r_+ - r_-) c^0$$

and the summation extended from  $j = 2$  when  $i = 0$ , and with the obvious modification when  $r_+ = r_-$ , i.e., when  $\mu = 1/2$ . As is easily seen,  $|r_+| < 1$  and  $|r_-| \leq 1/2$ , and hence

$$\left| \frac{r_+^j - r_-^j}{r_+ - r_-} \right| = |r_+^{j-1} + r_+^{j-2} r_- + \dots + r_-^{j-1}| \leq 2.$$

Consequently,  $|J_1| \leq 2|c^i|$ . For  $J_2$  we have

$$|J_2| \leq \Lambda(\mu) \max_{j \leq n} |\sigma^j(b_j c)| \leq \Lambda\tilde{\beta} \max_{j \leq n-1} |c^j|,$$

and hence

$$|c^n| \leq 2|c^i| + \Lambda\tilde{\beta} \max_{j \leq n-1} |c^j| \quad \text{for } n \geq \max(i+1, 2).$$

If  $\Lambda\tilde{\beta} < 1$ , we conclude  $|c^n| \leq C|c^i|$ , from which (4.5) with  $\gamma = 0$  follows by Parseval's identity.

Note that for  $\mu < 1/2$  the roots of (4.4) are real. In particular,  $r_+ > r_-$  and hence

$$(4.8) \quad \begin{aligned} \Lambda(\mu) &= \frac{2\mu}{3+2\mu} \sum_{j=1}^{\infty} \frac{r_+^j - r_-^j}{r_+ - r_-} = \frac{2\mu}{3+2\mu} \frac{1}{r_+ - r_-} \left( \frac{r_+}{1-r_+} - \frac{r_-}{1-r_-} \right) \\ &= \frac{2\mu}{3+2\mu} \frac{1}{\psi(1)} = 1. \end{aligned}$$

We remark that it follows that if we restrict  $k$  so that  $\lambda k \leq 1/2$  for all  $\lambda$  in the spectrum of  $A$  (which is possible, e.g., when  $A$  is a spatially discrete analogue of an elliptic operator), then it suffices to have  $\tilde{\beta} < 1$  in Proposition 4.1.

We proceed to show (4.5) for some positive  $\gamma$ . With  $c^n$  now equaling a Fourier coefficient of  $e^{\gamma t_n} U_i^n$ ,  $c^n = (e^{\gamma t_n} U_i^n, \varphi_\lambda)$ , we have

$$c^n - \frac{4}{3+2\mu} e^{\gamma k} c^{n-1} + \frac{1}{3+2\mu} e^{2\gamma k} c^{n-2} = \frac{2\mu}{3+2\mu} \sigma^n(b_{\gamma, n} c) \quad \text{for } n \geq \max(i+1, 2),$$

where  $b_{\gamma, n}(s) = e^{\gamma(t_n-s)} b(t_n - s)$ . The analogue of (4.7) follows with  $r_\pm$  replaced by  $\rho_\pm = e^{\gamma k} r_\pm$ . Since the minimal eigenvalue  $\lambda_0$  of  $A$  is positive, and since  $|r_+(\mu)| \leq |r_+(k\lambda_0)| \leq e^{-ck\lambda_0}$  for  $k$  sufficiently small, we have

$|\rho_+| \leq |e^{-k(c\lambda_0 - \gamma)}| \leq 1$  for  $\gamma \leq c\lambda_0$ . Similarly,  $1/(1 - |\rho_-|) \leq C$ , and hence  $|J_1| \leq C|c^i|$ .

It remains to investigate the quantities  $\Lambda_\gamma(\mu)$  obtained by replacing  $r_\pm$  by  $\rho_\pm$  in (4.3). In the case  $\mu \leq 1/2$ , a calculation as in (4.8), and using (4.4), establishes for small  $\gamma k$

$$\begin{aligned} \Lambda_\gamma(\mu) &= \frac{2\mu}{3 + 2\mu} \cdot \frac{1}{(1 - \rho_+)(1 - \rho_-)} = \frac{2\mu}{3 + 2\mu} \cdot \frac{e^{-2\gamma k}}{\psi(e^{-\gamma k})} \\ &= \frac{2\mu}{2\mu + 3 - 4e^{\gamma k} + e^{2\gamma k}} \leq 1 + C \frac{\gamma k}{\mu}. \end{aligned}$$

Since  $\mu = \lambda k$ , we see that for  $\lambda \geq \lambda_0$  and  $\gamma$  and  $k_0$  small enough,  $\Lambda_\gamma(\mu) \leq 1 + C\gamma = \Lambda(\mu) + C\gamma$ , uniformly in  $k \leq k_0$  and  $\mu \leq 1/2$ . For  $\mu > 1/2$ , since  $r = (3 + 2\mu)^{-1/2} \leq 1/2$ , we have

$$\begin{aligned} \Lambda_\gamma(\mu) - \Lambda(\mu) &= \frac{2\mu}{3 + 2\mu} \sum_{j=1}^{\infty} (e^{(j-1)\gamma k} - 1) \left| \frac{r_+^j - r_-^j}{r_+ - r_-} \right| \\ &\leq \gamma k \sum_{j=1}^{\infty} j^2 (e^{\gamma k/2})^{j-1} \leq C\gamma k, \end{aligned}$$

for  $k$  small enough. Thus altogether, for  $k$  and  $\gamma$  small,

$$|J_2| \leq \Lambda\tilde{\beta} \max_{j \leq n-1} |c^n|$$

and therefore, as before,  $|c^n| \leq C|c^i|$ . By our present definition of  $c^n$ , we now obtain (4.5) with a positive  $\gamma$ . This completes the proof of the proposition.  $\square$

In order to derive an error estimate, we further assume that the quadrature scheme used is second-order accurate, now with an error bound as (4.9)

$$|q^n(g)| = \left| \sigma^n(g) - \int_0^{t_n} g(s) ds \right| \leq Ck^2 \int_0^{t_n} |g(s)|_q ds + Ck \int_{t_{n-1}}^{t_n} |g(s)|_q ds,$$

and we also assume (3.5) and (3.6). The rather special form of this error bound is motivated by our examples in §7, where the piece  $\int_{t_{n-1}}^{t_n} g(s) ds$  of the integral will always be approximated by the left rectangle rule as  $kg(t_{n-1})$ .

**Theorem 4.1.** *Assume that (1.3) (with  $c_0 = 1$ ), (3.5), (3.6), and (4.9) hold, and further that  $\Lambda\tilde{\beta} < 1$ . Then there exist positive  $C$ ,  $\delta$ , and  $k_0$  such that for the solutions of (1.1) (with  $B = A$ ), and (4.1), (4.2), we have, for  $n \geq 0$ ,  $k \leq k_0$ ,*

$$\|U^n - u(t_n)\| \leq Ck^2 \left( e^{-\delta t_n} \|u_{tt}(0)\| + \int_0^{t_n} e^{-\delta(t_n-s)} (\|u_{ttt}(s)\| + \|u(s)\|)_q ds \right).$$

*Proof.* Letting  $e^n = U^n - u(t_n)$ , we have

$$\left( \bar{\partial} + \frac{k}{2} \bar{\partial}^2 \right) e^n + Ae^n = \sigma^n(b_n Ae) - \tau^n + q^n(b_n Au) \quad \text{for } n \geq 2,$$

where  $\tau^n = (\bar{\partial} + \frac{k}{2}\bar{\partial}^2)u(t_n) - u_t(t_n)$ . Here,  $e^0 = 0$  and, with  $U^1$  given in (4.2), the proof of Theorem 3.1 immediately gives

$$\begin{aligned} \|e^1\| &\leq Ck^2 \left( \max_{0 \leq s \leq k} \|u_{tt}(s)\| + \int_0^k \| \|u\| \|_q ds \right) \\ &\leq Ck^2 \left( \|u_{tt}(0)\| + \int_0^k (\|u_{tt}\| + \| \|u\| \|_q) ds \right). \end{aligned}$$

Also, essentially as in Theorem 3.1 but now using (4.9), with  $\delta < \min(\alpha, \gamma)$ ,

$$\begin{aligned} &k \sum_{j=2}^n e^{-\gamma t_{n-j}} \|q^j(b_j Au)\| \\ &\leq Ck^2 \int_0^{t_n} e^{-\delta(t_n-s)} \| \|u(s)\| \|_q ds + Ck^2 \sum_{j=2}^n e^{-\gamma t_{n-j}} \int_{t_{j-1}}^{t_j} \| \|u(s)\| \|_q ds \\ &\leq Ck^2 \int_0^{t_n} e^{-\delta(t_n-s)} \| \|u(s)\| \|_q ds. \end{aligned}$$

Finally, by expanding  $\tau^n$  in a Taylor series around  $t_n$ , employing the integral form of the remainder term, we have

$$k \sum_{j=2}^n e^{-\gamma t_{n-j}} \|\tau^n\| \leq k^2 e^{\gamma k} \int_0^{t_n} e^{-\gamma(t_n-s)} \|u_{ttt}(s)\| ds.$$

By Lemma 3.1 we may apply Proposition 4.1, and the theorem follows.  $\square$

### 5. THE CASE OF A MORE GENERAL MEMORY FORM

In this section we shall treat backward Euler schemes (1.4) in the case that the operator occurring in the memory term in (1.1) is different from the positive definite operator  $A$  on the left side of the equation. Our purpose now is to derive a discrete stability estimate for (1.4) of the form of Proposition 3.1 in the present more general situation.

In [5] such an estimate was obtained, but only on a finite interval in time. This result may be formulated as follows: Assume that  $|b(t)| \leq K$ , that (instead of (1.2))

$$(5.1) \quad \|A^{-1}B\| \leq c_1,$$

and further that the quadrature weights are dominated in the sense that

$$(5.2) \quad \omega_{js} \leq \omega_s \quad \text{for } j \geq s + 1.$$

Then

$$\|U^n\| \leq e^{Kc_1 \tilde{\omega}_{n0}} \left( \|U^0\| + k \sum_{j=1}^n \|f^j\| \right) \quad \text{for } n \geq 0, \quad \text{where } \tilde{\omega}_{ni} = \sum_{s=i}^{n-1} \omega_s.$$

We shall now sharpen the techniques of [5] to improve this result under the further assumption that  $b$  is exponentially decreasing. Letting  $\lambda_0$  denote the minimal eigenvalue of  $A$ , we then have the following:

**Proposition 5.1.** *Assume that (1.3)(i), (5.1), and (5.2) hold, and let  $\gamma = \frac{2}{5} \min(\alpha, \lambda_0)$ . There exists  $k_0 > 0$  such that, for the solution of (1.4),*

$$(5.3) \quad \|U^n\| \leq e^{Kc_1\hat{\omega}_{n0}-\gamma t_n} \|U^0\| + k \sum_{j=1}^n e^{Kc_1\hat{\omega}_{nj}-\gamma t_{n-j}} \|f^j\| \quad \text{for } n \geq 0, \quad k \leq k_0.$$

We remark that, for some quadrature schemes considered in §7, (5.2) will hold with  $\hat{\omega}_{nj} \leq c_2 t_{n-j} + c_3$ . If so, (5.3) shows long-time stability provided  $Kc_1$  is sufficiently small compared to  $\alpha$  and  $\lambda_0$  (and provided we have appropriate bounds on  $\|f^j\|$ ).

*Proof of Proposition 5.1.* With the notation and argument of the proof of Proposition 3.1 it suffices to show that now

$$(5.4) \quad \|U_i^n\| \leq e^{Kc_1\hat{\omega}_{ni}-\gamma t_{n-i}} \|U_i^i\| \quad \text{for } n \geq i.$$

With  $E_k = (I + kA)^{-1}$  we have

$$(5.5) \quad U_i^n = E_k^{n-i} U_i^i + k \sum_{j=i+1}^n E_k^{n-j+1} \sigma^j(b_j B U_i) \quad \text{for } n \geq i.$$

Here,

$$\begin{aligned} k \sum_{j=i+1}^n E_k^{n-j+1} \sigma^j(b_j B U_i) &= k \sum_{j=i+1}^n E_k^{n-j+1} \sum_{s=0}^{j-1} \omega_{js} b_{j-s} B U_i^s \\ &= k \sum_{s=i}^{n-1} \left[ \sum_{j=s+1}^n \omega_{js} b_{j-s} E_k^{n-j+1} A \right] A^{-1} B U_i^s. \end{aligned}$$

We shall show below that

$$(5.6) \quad \left\| k \sum_{j=s+1}^n \omega_{js} b_{j-s} E_k^{n-j+1} A \right\| \leq K \omega_s e^{-\gamma t_{n-s}} \quad \text{for } n > s.$$

Assuming this for the moment, we have from (5.5) and (5.1), for  $k$  small enough,

$$\|U_i^n\| \leq e^{-\lambda_0 t_{n-i}/2} \|U_i^i\| + Kc_1 \sum_{s=i}^{n-1} \omega_s e^{-\gamma t_{n-s}} \|U_i^s\|.$$

Multiplying by  $e^{\gamma t_{n-i}}$ , we find with  $\phi^n = e^{\gamma t_n} \|U^n\|$ , for  $\gamma \leq \lambda_0/2$ ,

$$\phi^n \leq \phi^i + Kc_1 \sum_{s=i}^{n-1} \omega_s \phi^s \quad \text{for } n \geq i,$$

from which (5.4) follows by a discrete version of Gronwall’s lemma (cf. [9, Lemma 3.3]).

It remains to prove (5.6). By spectral representation this reduces to showing

$$\left| k \sum_{j=s+1}^n \omega_{js} b_{j-s} r(k\lambda)^{n-j+1} \lambda \right| \leq K \omega_s e^{-\gamma t_{n-s}}, \quad \text{where } r(\mu) = (1 + \mu)^{-1},$$

or, using (5.2) and (1.3)(i),

$$k\lambda \sum_{j=s+1}^n r(k\lambda)^{n-j+1} e^{-\alpha t_{j-s}} \leq e^{-\gamma t_{n-s}},$$

i.e., after changes of variables,

$$(5.7) \quad Q = Q(\lambda, k, n) \equiv k\lambda \sum_{j=1}^n r(k\lambda)^j e^{-\alpha t_{n-j+1}} \leq e^{-\gamma t_n}.$$

We consider first  $\lambda \geq \alpha$  and want to show that (5.7) holds with  $\gamma = .4\alpha$ . Replacing  $\alpha k$  by  $k$  and  $\lambda$  by  $\lambda/\alpha$ ,  $\gamma$  by  $\gamma/\alpha$ , we may assume that  $\alpha = 1$ . We begin then by showing that for any  $\bar{\lambda} < .6e$  ( $\approx 1.63$ ) there is a  $k_0$  such that (5.7) holds with  $\gamma = .4$  for  $1 \leq \lambda \leq \bar{\lambda}$ ,  $k \leq k_0$ . In fact, since  $r(k\lambda)$  is decreasing, we have for these  $\lambda$ , with any  $\delta < 1$  and  $k$  small,  $r(k\lambda) \leq r(k) \leq e^{-k\delta}$ , and hence, using  $xe^{-px} \leq (pe)^{-1}$  for  $x \geq 0$ ,

$$Q \leq \bar{\lambda} t_n e^{-\delta t_n} \leq \bar{\lambda} (\delta - .4)e^{-1} e^{-.4t_n} \leq e^{-\gamma t_n},$$

if  $\bar{\lambda} \leq (\delta - .4)e$ , which holds for  $\delta$  suitably close to 1.

We next show that (5.7) holds also for  $\lambda \geq \bar{\lambda}$ . For this we write via summation of the finite power series,

$$Q = \frac{\lambda k}{e^k - (1 + \lambda k)} (r^n - e^{-t_n}),$$

or

$$e^{\gamma t_n} Q = \frac{\lambda k}{1 + \lambda k - e^k} (e^{-.6t_n} - (e^{.4k} r)^n) = \frac{\lambda k}{1 + \lambda k - e^k} (e^{-.6t_n} - e^{-(\zeta - .4)t_n}),$$

where we have set  $r = e^{-\zeta k}$  with  $\zeta = \zeta(\lambda, k) = -\ln r(\lambda k)/k$ . Note that  $\zeta(\lambda, k)$  is monotone increasing in  $\lambda$  and, since  $1 + a \leq e^a$ , we have  $\lambda \geq \zeta(\lambda, k)$ . Further,  $\zeta(\bar{\lambda}, k) = \bar{\lambda} + O(k)$ . An elementary investigation shows that the function  $|e^{-at} - e^{-bt}|$  with  $0 \leq t < \infty$ ,  $a, b > 0$ , has a maximum which equals

$$|b - a| a^{\frac{a}{b-a}} b^{\frac{b}{a-b}} = \left| \frac{b - a}{a} \right| \left( \frac{b}{a} \right)^{\frac{b}{a-b}}.$$

With  $a = \zeta - .4$  and  $b = .6$  we hence have

$$e^{\gamma t_n} Q \leq \left[ \frac{\lambda k}{1 + \lambda k - e^k} \cdot \frac{\zeta - 1}{\zeta - .4} \right] \times \left[ \left( \frac{.6}{\zeta - .4} \right)^{\frac{.6}{\zeta - 1}} \right] \equiv I_1 \times I_2.$$

For  $k$  small enough, since  $\zeta \leq \lambda$ ,

$$I_1 \leq \frac{\lambda}{\lambda - 1 + O(k)} \cdot \frac{\lambda - 1}{\lambda - .4} = 1 + \frac{.4}{\lambda - .4} + \frac{O(k)}{\lambda - 1} \leq 1 + \frac{1}{\zeta - 1} (.4 + O(k)),$$

and thus

$$\ln I_1 \leq \frac{.4 + O(k)}{\zeta - 1} = \frac{.6}{\zeta - 1} \left( \frac{2}{3} + O(k) \right).$$

Further, since  $\zeta(\lambda, k) \geq \zeta(\bar{\lambda}, k) = \bar{\lambda} + O(k)$ ,

$$\ln I_2 \leq -\frac{.6}{\zeta - 1} \cdot \ln \left( \frac{\zeta - .4}{.6} \right) \leq -\frac{.6}{\zeta - 1} \cdot \ln \left( \frac{\bar{\lambda} - .4}{.6} + O(k) \right),$$

and hence

$$\ln(e^{\gamma t_n} Q) \leq \frac{.6}{\zeta - 1} \left[ \frac{2}{3} + O(k) - \ln \left( \frac{5\bar{\lambda}}{3} - \frac{2}{3} + O(k) \right) \right].$$

Since  $5\bar{\lambda}/3$  is arbitrarily close to  $e$  and  $\frac{2}{3} - \ln(e - \frac{2}{3}) \approx -.05 < 0$ , we may first choose  $\bar{\lambda}$  and then  $k_0$  so that the right-hand side is negative for  $k \leq k_0$ , which shows (5.7).

It remains to consider the case  $\lambda_0 \leq \lambda \leq \alpha$ . Presenting the argument “modulo  $O(k)$ -terms”, now  $\zeta(\lambda, k)$  behaves uniformly like  $\lambda$ , and we consider instead

$$\tilde{Q}(\lambda, k, n) = \lambda k \sum_{j=1}^n e^{-\lambda t_j} e^{-\alpha t_{n-j+1}} = \frac{\lambda}{\alpha} \left[ \alpha k \sum_{i=1}^n e^{-\alpha t_i} e^{-\lambda t_{n-i+1}} \right].$$

Comparing the expression inside the square brackets with the previous  $Q(\lambda, k, n)$ , we see that the roles of  $\lambda$  and  $\alpha$  are reversed, and since  $\lambda/\alpha \leq 1$ , our analysis above clearly applies with the appropriate modifications, taking now  $\gamma = .4\lambda_0$ .

This completes the proof of the proposition.  $\square$

In the same way as earlier, the stability implies an error estimate. We shall not insist on the details.

### 6. APPLICATION TO FINITE ELEMENT DISCRETIZATIONS IN SPACE

In this section we shall briefly discuss the application of the results above to spatial discretization by finite elements of an integro-differential equation of the form (1.1). Thus, let  $A$  denote a second-order selfadjoint elliptic operator with vanishing Dirichlet boundary conditions in a bounded domain  $\Omega \subset \mathbb{R}^d$  with smooth boundary, corresponding to a coercive bilinear form

$$A(u, u) = \int_{\Omega} \left( \sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} + a_0 u^2 \right) dx \geq c \|u\|_{H^1}^2, \text{ for } u \in H_0^1, \text{ with } c > 0.$$

Here we denote as usual by  $H^r = H^r(\Omega)$  the Sobolev space of functions with derivatives of order at most  $r$  in  $L_2(\Omega)$ , with the standard norm, and by  $H_0^1$  the functions in  $H^1$  which vanish on  $\partial\Omega$ . Our basic Hilbert space is now  $L_2(\Omega)$ .

Letting  $S_h, 0 < h \leq h_0$ , be a family of finite-dimensional subspaces of  $H_0^1$ , and introducing  $A_h : S_h \rightarrow S_h$  by  $(A_h \psi, \chi) = A(\psi, \chi)$  for  $\psi, \chi \in S_h$ , and the orthogonal projection  $P_h : L_2 \rightarrow S_h$  by  $(P_h u, \chi) = (u, \chi)$  for  $\chi \in S_h$ , we consider the semidiscrete analogue of (1.1) with  $B = A$  in  $S_h$  to be

$$u_{h,t} + A_h u_h = \int_0^t b(t-s) A_h u(s) ds + P_h f(t) \text{ for } t > 0, u_h(0) = u_{0,h}.$$

The backward Euler method applied to this determines  $U^n \in S_h$  from

$$(6.1) \quad \bar{\partial} U^n + A_h U^n = \sigma^n (b_n A_h U) + P_h f^n, \text{ for } n \geq 1, U^0 = u_{0,h}.$$

We shall use the stability result of Proposition 3.1 to derive an error estimate. Checking the proof of that proposition, one sees that the properties of

the operator  $A$  enter only in the condition  $e^{\gamma k} - 1 \leq \lambda_0 k(1 - \tilde{\beta})$  for its lowest eigenvalue. Since the lowest eigenvalue  $\lambda_{0,h}$  of  $A_h$  is bounded below by the smallest eigenvalue  $\lambda_0$  of  $A$ , we may apply Proposition 3.1 to (6.1) with quantities that are independent of  $h$ .

We next introduce the Ritz projection  $R_h: H_0^1 \rightarrow S_h$  by  $A(R_h v, \chi) = A(v, \chi)$ , for  $\chi \in S_h$ . We assume that the  $S_h$  are such that the error estimate

$$(6.2) \quad \|R_h v - v\| \leq Ch^r \|v\|_{H^r} \quad \text{for } v \in H_0^1 \cap H^r$$

holds, where  $r$  is a positive integer. Writing

$$U^n - u(t_n) = (U^n - R_h u(t_n)) + (R_h u(t_n) - u(t_n)) = \theta^n + \rho^n,$$

we may use (6.2) in the obvious way to bound  $\rho^n$ . For  $\theta^n \in S_h$  we find by our definitions, using that  $A_h R_h v = P_h A v$  for  $v \in H^2 \cap H_0^1$ ,

$$\bar{\partial} \theta^n + A_h \theta^n = \sigma^n (b_n A_h \theta) + P_h (\bar{\partial} \rho^n + \tau^n + q^n (b_n A u)) \quad \text{for } n \geq 1,$$

and, assuming for simplicity that  $u_{0,h} = R_h u_0$ , we have  $\theta^0 = 0$ . Here,  $\tau^n = \bar{\partial} u(t_n) - u_t(t_n)$ . As remarked above, we may apply Proposition 3.1 (with  $H = S_h$  equipped with the  $L_2$  inner product) so that, since  $\|P_h v\| \leq \|v\|$ ,

$$(6.3) \quad \|\theta^n\| \leq k \sum_{j=1}^n e^{-\gamma t_{n-j}} (\|\bar{\partial} \rho^j\| + \|\tau^j\| + \|q^j (b_j A u)\|).$$

By (6.2) and the definition of  $\tau^j$ ,

$$\|\bar{\partial} \rho^j\| \leq C \frac{h^r}{k} \int_{t_{j-1}}^{t_j} \|u_t\|_{H^r} ds, \quad \|\tau^j\| \leq \int_{t_{j-1}}^{t_j} \|u_{tt}\| ds.$$

Using this in (6.3) and treating the term in the quadrature error as in the proof of Theorem 3.1, we find the following.

**Theorem 6.1.** *Assume that (1.3) (with  $c_0 = 1$ ), (3.5), (3.6), and (6.2) hold, and let  $u_{0,h} = R_h u_0$ . Then there exist positive  $C$ ,  $\delta$ , and  $k_0$  such that for the solutions of (1.1) (with  $B = A$ ) and (6.1),*

$$\begin{aligned} \|U^n - u(t_n)\| &\leq Ch^r \left( \|u(t_n)\|_{H^r} + \int_0^{t_n} e^{-\delta(t_n-s)} \|u_t\|_{H^r} ds \right) \\ &\quad + Ck \int_0^{t_n} e^{-\delta(t_n-s)} (\|u_{tt}\| + \|u\|_q) ds \quad \text{for } n \geq 0, k \leq k_0. \end{aligned}$$

A similar result holds also for the second-order backward difference methods, with  $k$  replaced by  $k^2$ . (Again, it is easy to check that the quantities in Proposition 4.1 depend on  $A$  only through a lower bound on its lowest eigenvalue.)

As for the case of  $B \neq A$  treated in §5, where now  $B$  is any spatial partial differential operator with smooth coefficients of at most second order, an application of Proposition 5.1 would require a bound for  $\|A_h^{-1} B_h\|$ , where  $(B_h v, \chi) = B(v, \chi)$ , for  $\chi \in S_h$ , with  $B(\cdot, \cdot)$  the natural bilinear form on  $H_0^1 \times H_0^1$  associated with  $B$ . This may be obtained by a standard error estimate for the associated elliptic problem, together with an inverse assumption (cf. [9, §3]). In some cases, such as when  $B = g(x)A + \tilde{B}$  with  $\tilde{B}$  a first-order operator, the inverse assumption is not necessary for this (cf. also [4, §4]).

Finally, to obtain an error bound in the case  $B \neq A$ , rather than comparing  $U^n$  to  $R_h u(t_n)$ , one may conveniently compare it to a so-called Ritz-Volterra projection of  $u$ . We shall not pursue this and refer to [1, 3, and 4] for details.

### 7. ON QUADRATURE RULES

Following the philosophy of [5] and [9], we shall now briefly discuss some quadrature rules that are suitable for application in the situations described above.

In the case of the backward Euler scheme treated in §3 when  $B = A$ , the only requirement on  $\sigma^n(g)$  for the stability and error estimates to hold is that the accuracy condition (3.5) be satisfied for some  $q$ . The simplest choice is then the left-side evaluation composite rectangle rule,

$$(7.1) \quad \sigma^n(g) = k \sum_{j=0}^{n-1} g(t_j),$$

for which (3.5) is valid with  $q = 1$ . A disadvantage of this rule is that all the preceding values  $U^j$  need to be stored as the computation progresses in time.

We shall now give a first-order rule based on the use of the composite trapezoidal formula on intervals of length  $O(k^{1/2})$ , and then somewhat modified at the right end of the interval of integration. Let  $\mu = [k^{-1/2}]$ , where  $[x]$  denotes the integral part of  $x$ . Set  $k_1 = \mu k$  and  $\bar{t}_j = j k_1$ , and let  $j_n$  be the largest integer such that  $\bar{t}_{j_n} < t_n$ . In approximating the integral on  $[0, t_n]$ , we shall now apply the composite trapezoidal rule with stepsize  $k_1$  on  $[0, \bar{t}_{j_n}]$ , then the one-interval trapezoidal rule on  $[\bar{t}_j, t_{n-1}]$  (which may be void) and, finally, the left-side rectangle rule on the remaining interval  $[t_{n-1}, t_n]$ . Thus,

$$(7.2) \quad \sigma^n(g) = \frac{k_1}{2} \sum_{j=1}^{j_n} (g(\bar{t}_j) + g(\bar{t}_{j-1})) + \frac{1}{2} (t_{n-1} - \bar{t}_{j_n}) (g(t_{n-1}) + g(\bar{t}_{j_n})) + k g(t_{n-1}).$$

Since the rule is second-order in  $k_1$  on  $(0, \bar{t}_{j_n})$  and on  $(\bar{t}_{j_n}, t_{n-1})$ , and first-order in  $k$  on  $(t_{n-1}, t_n)$ , we have

$$|q^n(g)| \leq C k_1^j \int_0^{t_{n-1}} |g|_j ds + C k \int_{t_{n-1}}^{t_n} |g|_1 ds \leq C k^{j/2} \int_0^{t_n} |g|_j ds \quad \text{for } j = 1, 2,$$

so that (3.5) holds with  $q = 2$ . We note that the storage requirement is now  $O(k^{-1/2})$  time levels on a unit length time interval, as opposed to the  $O(k^{-1})$  bound for the composite left rectangle rule.

Going further with this storage-saving idea, one may set  $\mu = [k^{-1/4}]$  and  $k_2 = \mu^3 k = O(k^{1/4})$ . The new quadrature rule then uses Simpson's formula on as many intervals of length  $2k_2$  that can be fitted, starting from the left, into  $[0, t_{n-1}]$ . On the remaining interval, which is of length at most  $O(k^{1/4})$ , it uses the composite trapezoidal rule on as many intervals of length  $k_1 = \mu^2 k = O(k^{1/2})$  that fit in, thus reaching  $\bar{t}_{j_n}$ , then the one-interval variable-length trapezoidal rule on the interval  $[\bar{t}_{j_n}, t_{n-1}]$  and finally the left rectangle rule for  $[t_{n-1}, t_n]$ . It is clear that the number of time-levels that need to be stored per unit time is  $O(k_2^{-1}) + O(k_2 k_1^{-1}) + 1 = O(k^{-1/4})$ . Similarly to above, the method satisfies (3.5) with  $q = 4$ .

We now turn to some quadrature rules which satisfy the second-order estimate (4.9) and thus are appropriate for use with second-order backward differencing. The simplest would be to use the trapezoidal rule on each time interval of length  $k$  in  $[0, t_{n-1}]$  with the left-side rectangle rule on  $[t_{n-1}, t_n]$ , i.e.,

$$\sigma^n(g) = \frac{1}{2}k g(0) + k \sum_{j=1}^{n-2} g(t_j) + \frac{3}{2}k g(t_{n-1}).$$

This rule satisfies (3.5) with  $q = 1$  and, since its error is bounded by

$$Ck^2 \int_0^{t_{n-1}} |g(s)|_2 ds + Ck \int_{t_{n-1}}^{t_n} |g(s)|_1 ds,$$

also (4.9) with  $q = 2$ .

In order to save storage in the second-order context, we introduce a method based on Simpson's rule on longer time intervals. For this, let  $\mu = [k^{-1/2}]$ ,  $k_1 = \mu k = O(k^{1/2})$ , and  $\bar{t}_j = j k_1$ , and let  $j_n$  be the largest even integer with  $\bar{t}_{j_n} \leq t_{n-1}$ . We then apply Simpson's rule on as many intervals of length  $2k_1$  as possible in  $[0, t_{n-1}]$ , starting from the left, then the trapezoidal rule on most of the remaining small intervals of length  $k$ , and the left rectangle rule on  $[t_{n-1}, t_n]$ . Thus,

$$\begin{aligned} \sigma^n(g) &= \frac{k_1}{3} \sum_{j=1}^{j_n/2} [g(\bar{t}_{2j}) + 4g(\bar{t}_{2j-1}) + g(\bar{t}_{2j-2})] \\ &\quad + \frac{k}{2} \sum_{j=j_n\mu+1}^{n-1} (g(t_j) + g(t_{j-1})) + k g(t_{n-1}). \end{aligned}$$

This rule satisfies (3.5) with  $q = 2$  and (4.9) with  $q = 4$ , and its storage requirement is  $O(k^{-1/2})$  time-levels per unit time.

For further storage reduction in the second-order case, an obvious idea would be to use, e.g., the eighth-order formula of Newton-Cotes type, often referred to as the six-strip approximation,

$$\int_a^b g(s) ds = \frac{b-a}{840} (41g_0 + 216g_1 + 27g_2 + 272g_3 + 27g_4 + 216g_5 + 41g_6)$$

(where  $g_i$  has the obvious meaning). It could be used in a composite fashion on as many intervals of length  $6[k^{-3/4}]k$  that fit into  $[0, t_{n-1}]$  starting from the left, followed by Simpson's rule on intervals of length first  $2[k^{-1/2}]k$  and then  $2[k^{-1/4}]k$ , then by the trapezoidal rule on intervals of length  $k$ , and finally by the left side rectangle rule on  $(t_{n-1}, t_n)$ . Now (3.5) holds with  $q = 4$  and (4.9) with  $q = 8$ . The storage requirement is then  $O(k^{-1/4})$  per unit time.

We next consider quadrature rules suitable for application of the analysis in §5 of the backward Euler methods when  $B \neq A$ . Here, in addition to the accuracy requirements, we assume that the quadrature coefficients are dominated as in (5.2), and it is desirable for the stability estimate that  $\tilde{\omega}_{nj} = \sum_{s=i}^{n-1} \omega_s \leq c_2 t_{n-j} + c_3$ . For the rectangle rule (7.1) one may choose  $\omega_s = k$  and thus  $\tilde{\omega}_{nj} = t_{n-j}$ , so that this condition is satisfied.

Turning now to the composite trapezoidal rule (7.2), we see that  $\omega_{n,n-1} = (t_{n-1}/2 - \bar{t}_{j_n}) + k$  for every  $n$ . Since  $t_{n-1} - \bar{t}_{j_n} \geq k_1/2$  about half of the

time, it is clear that  $\tilde{\omega}_{nj} \gtrsim k_1(n-j)/2 \simeq t_{n-j}k^{-1/2}/2$ , which renders the result of Proposition 5.1 rather useless. The following slight modification may now be made. Instead of approximating the integral on the next-to-last interval  $[\bar{t}_{j_n}, t_{n-1}]$ , which is of variable length, by the one-step trapezoidal rule, we employ the composite rectangle (or trapezoidal) rule with all subintervals of length  $k$  involved. It is then clear that  $\tilde{\omega}_{nj} \leq 2t_{n-j} + ck_1$  (cf. [5]). This modification has a slightly higher storage requirement, but this requirement is still within the previous  $O(k^{-1/2})$  bound per unit time interval.

Since by (1.3)(i) our memory is fading at an exponential rate, the contribution at large  $t$  from the solution on  $[0, T]$  is of order  $O(e^{-\gamma(t-T)})$ . Since this is eventually bounded by the discretization error  $O(k^m)$ ,  $m = 1$  or  $2$ , when this is uniform for  $t$  positive, one may then set  $\omega_{nj} = 0$  for  $n-j \geq Mk^{-1} \ln(k^{-1})$ , with a suitable  $M$ . When the storage requirement per unit time is  $O(k^{-1/p})$ , the total storage requirement now becomes  $O(k^{-1/p}) \ln(k^{-1})$ .

We conclude by remarking that even in the case that  $u(t)$  approaches a limit  $u_\infty$  "exponentially fast" as  $t \rightarrow \infty$ , Theorem 3.1 (and its analogue for the backward difference formula) merely asserts that  $U^n$  is uniformly within  $O(k)$  (and  $O(k^2)$ , respectively) of  $u(t_n)$ . In fact, no assertion can be made in general that  $U^n$  has a limit as  $n \rightarrow \infty$  for a fixed  $k$ . For instance, in the cases of the modified Simpson's rules, from computer experiments it appears that  $U^n$  approaches a periodic limit cycle with period determined by the periodic changes in the quadrature formulas. It seems, though, that the periodic changes are very small compared to the main part of the error.

APPENDIX. ESTIMATION OF THE CONSTANT  $\Lambda$  IN PROPOSITION 4.1

Our objective is to show that the quantity  $\Lambda$  defined in (4.3) satisfies  $\Lambda < 1.1$ . In view of (4.8), it suffices to show that for  $\mu \geq 1/2$ ,  $\Lambda(\mu) < 1.1$ . Writing  $r_\pm = (2 \pm i\sqrt{2\mu - 1})/(3 + 2\mu) = re^{\pm i\theta}$ , with  $r = (3 + 2\mu)^{-1/2}$ ,  $\theta = \arctan((2\mu - 1)^{1/2}/2)$ , we have

$$\Lambda(\mu) = \frac{2\mu}{3 + 2\mu} \sum_{j=1}^{\infty} \left| \frac{r_+^j - r_-^j}{r_+ - r_-} \right| = \frac{2\mu}{3 + 2\mu} \sum_{j=1}^{\infty} r^{j-1} \frac{|\sin j\theta|}{\sin \theta}.$$

We begin our proof by noting that since (cf. (4.8))

$$\frac{2\mu}{3 + 2\mu} \sum_{j=1}^{\infty} r^{j-1} \frac{\sin j\theta}{\sin \theta} = \frac{2\mu}{3 + 2\mu} \sum_{j=1}^{\infty} \frac{r_+^j - r_-^j}{r_+ - r_-} = 1,$$

we have

$$(A.1) \quad \Lambda(\mu) - 1 = -\frac{4\mu}{3 + 2\mu} \sum_{\sin j\theta < 0} r^{j-1} \frac{\sin j\theta}{\sin \theta}.$$

The range of  $\theta$  involved is  $0 \leq \theta < \pi/2$ ; our proof will now proceed by considering different subranges of  $\theta$ .

We start by treating the case of  $\theta \in I_8 = (0, \pi/8)$ . Then  $\sin j\theta > 0$  for  $j \leq 8$ , and since  $|\sin j\theta| \leq j|\sin \theta|$ , we have by (A.1),

$$(A.2) \quad \Lambda(\mu) - 1 \leq \frac{4\mu}{3 + 2\mu} \sum_{j=9}^{\infty} jr^{j-1} = \frac{4\mu}{3 + 2\mu} \left( 9 \frac{r^8}{1-r} + \frac{r^9}{(1-r)^2} \right).$$

The right-hand side is increasing in  $\mu$  and  $r$ , and since  $r \leq 1/2$  and  $\mu = \frac{1}{2} + 2 \tan^2 \theta < \frac{1}{2} + 2 \tan^2(\pi/8) < .85$ , (A.2) shows  $\Lambda(\mu) \leq 1.06$  in  $I_8$ .

We next treat the cases  $\theta \in I_q = (\pi/(q + 1), \pi/q)$  for  $2 \leq q \leq 7$ . We note that  $\sin j\theta \leq 0$  for  $j = q + 1, \dots, 2q$ , and hence

$$\Lambda(\mu) - 1 = \left[ \frac{-4\mu}{3 + 2\mu} \sum_{j=q+1}^{2q} r^{j-1} \frac{\sin j\theta}{\sin \theta} \right] - \left[ \frac{4\mu}{3 + 2\mu} \sum_{\substack{j \geq 2q+1 \\ \sin j\theta < 0}} r^{j-1} \frac{\sin j\theta}{\sin \theta} \right] \equiv J_1^q(\theta) + J_2^q(\theta).$$

The interval  $I_2 = (\pi/3, \pi/2)$  will be further divided into  $I'_2 = (\pi/3, 2\pi/5)$  and  $I''_2 = (2\pi/5, \pi/2)$ . The following table shows upper bounds for  $J_1^q(\theta)$  and  $J_2^q(\theta)$  in the respective interval, which will be shown below.

	$I_7$	$I_6$	$I_5$	$I_4$	$I_3$	$I'_2$	$I''_2$
$J_1^q(\theta) <$	.010	.017	.032	.056	.049	.045	.050
$J_2^q(\theta) <$	.001	.002	.006	.015	.038	.052	.008
$\Lambda(\mu) <$	1.011	1.019	1.036	1.071	1.087	1.097	1.058

Writing  $r^{j-1} \sin j\theta / \sin \theta = (r_+^j - r_-^j) / (r_+ - r_-)$ , and using that

$$(1 - r_+)(1 - r_-) = 2\mu / (3 + 2\mu),$$

cf. (4.4), we easily find by summing finite geometric series that

$$J_1^q(\theta) = \frac{-2r^q}{\sin \theta} [\sin(q + 1)\theta - r \sin q\theta - r^q(\sin(2q + 1)\theta - r \sin 2q\theta)].$$

Next, using that  $\sin(m + 1)\theta = \sin m\theta \cos \theta + \cos m\theta \sin \theta$  with  $m = q$  and  $2q$ , and also the fact that  $\cot \theta = 2/\sqrt{2\mu - 1}$ , this leads to (after rearranging terms),

$$(A.3) \quad J_1^q(\theta) = 2r^q \left[ -\cos q\theta + r^q \cos 2q\theta - \frac{\sin q\theta}{\sqrt{2\mu - 1}} + \frac{r^q \sin 2q\theta}{\sqrt{2\mu - 1}} \right].$$

We now note that, for  $\pi/(q + 1) \leq \theta < \pi/q$ , we have  $\sin q\theta > 0$  and  $\sin 2q\theta \leq 0$ , and thus

$$(A.4) \quad J_1^q(\theta) \leq 2r^q(1 + r^q).$$

For  $J_2^q(\theta)$ , we have as in (A.2)

$$(A.5) \quad J_2^q(\theta) \leq \frac{4\mu}{3 + 2\mu} \left( (2q + 1) \frac{r^{2q}}{1 - r} + \frac{r^{2q+1}}{(1 - r)^2} \right).$$

Bounding  $r = \cos \theta/2$  and  $\mu = \frac{1}{2} + 2 \tan^2 \theta$  in (A.4) and (A.5) by their values at the left and right endpoints of the respective intervals yields the bounds of the table except those for  $J_1^3(\theta)$  and the  $J_1^2(\theta)$ 's.

Consider thus  $J_1^3(\theta)$  on  $I_3$ . We have from (A.3), now throwing away only the last term on the right, and using standard trigonometric formulas,

$$J_1^3(\theta) \leq 2r^3 \left[ -\cos 3\theta - \frac{\sin 3\theta}{\sqrt{2\mu - 1}} \right] + 2r^6 = 2 \frac{14\mu - 27}{(3 + 2\mu)^3} + 2r^6.$$

Here,  $2.5 \leq \mu < 6.5$ , and an elementary calculation shows that the term involving  $\mu$  has its maximum at  $\mu = 204/56$  ( $\approx 3.64$ ) and is bounded by .045. Since  $2r^6 \leq 2(\cos(\pi/4)/2)^6 < .004$ , this shows the bound of the table.

Similarly we have on  $I_2 = (\pi/3, \pi/2)$

$$J_1^2(\theta) \leq 2r^2 \left[ -\cos 2\theta - \frac{\sin 2\theta}{\sqrt{2\mu - 1}} \right] + 2r^4 \cos 4\theta = 2 \frac{2\mu - 9}{(3 + 2\mu)^2} + 2r^4 \cos 4\theta.$$

The maximum value of the term involving  $\mu$  now occurs at  $\mu = 10.5$  and is bounded by .042. Since  $4\pi/3 \leq 4\theta \leq 8\pi/5$  on  $I_2$ , the last term is bounded by  $2(1/4)^4 \cos 8\pi/5 < .003$  on  $I_2'$  and, trivially, by  $2(1/4)^4 < .008$  on  $I_2''$ .

Together, these estimates show the remaining bounds of the table and thus complete the proof of our claim.

We conclude by remarking that computer experimentation suggests that  $1.04 < \Lambda < 1.05$ .

#### BIBLIOGRAPHY

1. J. R. Cannon and Y. Lin, *Nonclassical  $H^1$  projection and Galerkin methods for nonlinear parabolic integro-differential equations*, *Calcolo* **25** (1988), 187–201.
2. B. R. Coleman and M. E. Gurtin, *Equipresence and constitutive equations for rigid heat conductors*, *Z. Angew. Math.* **18** (1967), 199–208.
3. Y. Lin, V. Thomée, and L. B. Wahlbin, *Ritz-Volterra projections to finite element spaces and applications to integro-differential and related equations*, *SIAM J. Numer. Anal.* **28** (1991), 1047–1070.
4. A. K. Pani, V. Thomée, and L. B. Wahlbin, *Numerical methods for hyperbolic and parabolic integro-differential equations*, *J. Integral Equations Appl.* **4** (1992), 533–583.
5. I. H. Sloan and V. Thomée, *Time discretization of an integro-differential equation of parabolic type*, *SIAM J. Numer. Anal.* **23** (1986), 1052–1061.
6. V. Thomée, *On the numerical solution of integro-differential equations of parabolic type*, *Internat. Ser. Numer. Math.* **86** (1988), 477–493.
7. V. Thomée and N.-Y. Zhang, *Error estimates for semidiscrete finite element methods for parabolic integro-differential equations*, *Math. Comp.* **53** (1989), 121–139.
8. E. G. Yanik and G. Fairweather, *Finite element methods for parabolic and hyperbolic partial integro-differential equations*, *Nonlinear Anal.* **12** (1988), 785–809.
9. N.-Y. Zhang, *On fully discrete Galerkin approximations for partial integro-differential equations of parabolic type*, *Math. Comp.* **60** (1993), 133–166.

DEPARTMENT OF MATHEMATICS, CHALMERS UNIVERSITY OF TECHNOLOGY, S-41296 GÖTEBORG, SWEDEN

E-mail address: thomee@math.chalmers.se

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NEW YORK 14853

E-mail address: wahlbin@math.cornell.edu