ADAPTIVE FINITE ELEMENT METHODS FOR PARABOLIC PROBLEMS. VI. ANALYTIC SEMIGROUPS

KENNETH ERIKSSON, CLAES JOHNSON, AND STIG LARSSON

ABSTRACT. We continue our work on adaptive finite element methods with a study of time discretization of analytic semigroups. We prove optimal *a priori* and *a posteriori* error estimates for the discontinuous Galerkin method showing, in particular, that analytic semigroups allow long-time integration without error accumulation.

1. INTRODUCTION

This paper is a continuation of the series of papers [1], [2], [3], [4], [5] on adaptive finite element methods for parabolic problems. The method considered is the discontinuous Galerkin method (the dG-method) based on a space-time finite element discretization with piecewise polynomial basis functions that are continuous in space and discontinuous in time. In [1], [2], [3], [4], [5] we proved optimal *a priori* and *a posteriori* error estimates for the dG-method for parabolic problems, typically of the form: find $u : [0, \infty) \to H$ such that

$$\dot{u}(t) + Au(t) = f(t), \ t > 0; \quad u(0) = u_0, \tag{1.1}$$

where $H = L_2(\Omega)$ with Ω a bounded domain in $\mathbf{R}^n$, $Av = -\Delta v$ with domain of definition $H^2(\Omega) \cap H_0^1(\Omega)$, $\dot{u} = \partial u/\partial t$, and $f \in L_1((0,\infty); H)$. The proofs are based on a combination of the orthogonality inherent in Galerkin's method and what we call "strong stability," which is essentially the same as the "smoothing property" characteristic of parabolic problems.

In the semidiscrete case, with discretization only in time, using piecewise polynomials of degree $q \geq 0$ on a mesh $0 = t_0 < t_1 < \cdots < t_N = T$, the *a priori* and *a posteriori* error estimates of [1], [2], [3], [4], [5] essentially take the form

$$\|u - U\|_{L_\infty(J;H)} \leq C^i C_q^s L_N \|k^{q+1} u^{(q+1)}\|_{L_\infty(J;H)}, \tag{1.2}$$

$$\|u - U\|_{L_\infty(J;H)} \leq C^i C^s L_N \Big(\|[U]\|_{L_\infty(J;H)} + \|k^{q+2} f^{(q+1)}\|_{L_\infty(J;H)} \Big), \tag{1.3}$$

respectively, where U is the piecewise polynomial approximate solution, $J = (0, T)$ is the time interval under consideration, $k = k(t)$ is the local time step defined by $k(t) = k_n = t_n - t_{n-1}$ for $t \in I_n = (t_{n-1}, t_n)$, $u^{(r)} = \partial^r u/\partial t^r$, $[U] = [U](t) = U_{n-1}^+ - U_{n-1}^-$ for $t \in I_n$, $U_n^\pm = \lim_{s \to 0^\pm} U(t_n + s)$, and $U_0^- = u_0$. Further, C^s is a stability constant related to the continuous problem (1.1) and C_q^s is a stability constant

---

1991 *Mathematics Subject Classification.* 65M60, 65M15.

Key words and phrases. analytic semigroup, parabolic problem, discontinuous Galerkin method, error estimate, a priori, a posteriori, strong stability.

Published in *SIAM J. Numer. Anal.* **35** (1998), 1315–1325.

`http://epubs.siam.org/sam-bin/dbq/article/31021`.

related to the dG-discretization, both independent of T, u, k, and U. Finally, C^i is an interpolation constant depending only on q, and $L_N = (1 + \log(t_N/k_N))^{1/2}$.

It should be remarked that the results of [1], [2], [3], [4], [5] are formulated for the special cases $q = 0$ and $q = 1$ only, and that they include also the effect of discretization with respect to the spatial variables. Moreover, it is shown that the error at the mesh points t_n is of order $O(k^{2q+1})$.

An important feature of (1.2)–(1.3) is their optimality compared to the error in interpolation, for which we have

$$\|u - \pi u\|_{L_\infty(J;H)} \leq C^i \|k^{q+1} u^{(q+1)}\|_{L_\infty(J;H)}.$$

In particular, the stability constants C^s and C_q^s are independent of the length T of the time interval, which shows that long-time integration without error accumulation (modulo the logarithmic factor L_N) is possible for the class of problems under consideration. (The optimality of (1.3) follows from (1.2), since $\|[U]\|_{L_\infty(J;H)} = \|[U-u]\|_{L_\infty(J;H)} \leq 2\|U-u\|_{L_\infty(J;H)}$.) Moreover, k^{q+1} is combined with the time derivative $u^{(q+1)}$ and not with $u^{(q+2)}$ as in classical error estimates for parabolic problems. The possibility of obtaining the optimal error estimates (1.2)–(1.3) is directly connected to the strong stability expressing the parabolic nature of the problem considered.

It is now natural to ask: what is the largest class of linear problems of the form (1.1) for which optimal error estimates of the form (1.2)–(1.3) are valid for time discretization by the dG-method, including, in particular, the possibility to integrate over long time without error accumulation? We shall see that an answer to this question may be given as follows. If $-A$ is the infinitesimal generator of a bounded, analytic semigroup, then the dG-method for (1.1) admits optimal *a priori* and *a posteriori* error estimates over arbitrarily long time intervals. Moreover, analytic semigroups seem to be the largest class of linear evolution problems with this property. The reason for this is, as we shall see, the connection between strong stability and the defining property of analytic semigroups.

More precisely, in the proof of the *a posteriori* error estimate (Theorem 1) we use the strong stability of the *continuous* evolution problem, which is directly connected to the defining property of an analytic semigroup. On the other hand, for the *a priori* error estimate (Theorem 4) we use the strong stability of the *discrete* problem with a less obvious connection to the analyticity of the semigroup. The technical novelty of this paper is a proof of the fact that the dG-method for an analytic semigroup satisfies strong stability estimates leading to optimal *a priori* error estimates (Theorem 2).

The magnitudes of the stability constants C^s and C_q^s in the error estimates is obviously crucial. We consider here situations where these constants are of moderate size. When the constants increase, the strong stability (or analyticity) degenerates, and the underlying evolution problem essentially becomes "hyperbolic." For hyperbolic problems we expect (see [7]) that the *a priori* estimate (1.2) is replaced by

$$\|u - U\|_{L_\infty(J;H)} \leq C^i C \|k^{q+1} u^{(q+2)}\|_{L_1(J;H)} \leq C^i C T \|k^{q+1} u^{(q+2)}\|_{L_\infty(J;H)},$$

so that $C_q^s = CT$. The error may thus grow linearly with time, and k^{q+1} is combined with $u^{(q+2)}$ instead of $u^{(q+1)}$. The error estimate of the classical analysis of time discretization has this form also for parabolic problems, thus missing the possible improvement in this case.

To sum up, we may thus roughly classify initial-value problems as "parabolic" if long-time integration without error accumulation is possible, and "hyperbolic" if the error estimates contain constants growing linearly in time. "Parabolic problems" would then correspond to analytic semigroups.

2. Analytic semigroups

We consider the initial value problem (1.1) in the following more general setting. Let H be a Hilbert space with inner product $(\cdot,\cdot)$ and norm $\|\cdot\|$, and let $A : D(A) \subset H \to H$ be a densely defined, closed, linear operator. We recall (see, e.g., [8], [10]) that $-A$ is the infinitesimal generator of an *analytic semigroup* $E(t) = \exp(-tA)$, $t \geq 0$, if and only if it is *sectorial*, i.e., if there are constants $M > 0$, $\theta \in (\pi/2, \pi)$, $\omega \in \mathbf{R}$, such that

$$\|(\lambda I + A)^{-1}\| \leq \frac{M}{|\lambda - \omega|}, \quad \forall \lambda \neq \omega, \ |\arg(\lambda - \omega)| < \theta. \tag{2.1}$$

The analytic semigroup $E(t)$ is *uniformly bounded* with respect to t if and only if $\omega \leq 0$. In this case there is a constant C^s such that

$$\|E(t)u_0\| \leq C^s \|u_0\|, \quad \|AE(t)u_0\| \leq C^s t^{-1} \|u_0\|, \quad \forall u_0 \in H, \ t > 0. \tag{2.2}$$

Throughout this paper we assume that $-A$ is the generator of a bounded analytic semigroup $E(t)$, i.e., we assume (2.2) or, equivalently, (2.1) with $\omega \leq 0$. It is well-known that $v(t) = E(t)\phi$ is the solution of the homogeneous initial value problem

$$\dot{v}(t) + Av(t) = 0, \ t > 0; \quad v(0) = \phi, \tag{2.3}$$

and the solution of (1.1) is then given by

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

under appropriate smoothness assumptions on f. The first inequality in (2.2) reflects the stability of the evolution problem. The second inequality is what we refer to as *strong stability*, which is thus equivalent to the bounded analyticity of the semigroup $E(t)$.

For future reference we note that the adjoint A^* of A is the infinitesimal generator of the dual semigroup $E(t)^* = \exp(-tA^*)$, $t \geq 0$, which satisfies bounds analogous to (2.2) with the same constant C^s. The constant C^s will appear as the stability constant in the *a posteriori* error estimate that we shall prove below. As already remarked, we assume that this constant is of moderate size.

Remark. We do not assume that $\omega < 0$ in (2.1), in which case the bounds in (2.2) would decay exponentially with t. Thus, A may have spectrum at $\lambda = 0$ or arbitrarily close to $\lambda = 0$. Our results hold, e.g., for $A = -\Delta$ (the Laplacian) with homogeneous Dirichlet boundary conditions ($\omega < 0$), and for $A = -\Delta$ with homogeneous Neumann boundary conditions ($\omega = 0$).

3. The discontinuous Garlerkin method

For the discretization of (1.1) we consider the dG(q)-method, which is a Galerkin method with discontinuous piecewise polynomials of degree $q \geq 0$ defined as follows. Let $0 = t_0 < \cdots < t_{n-1} < t_n < \cdots$ be a mesh with corresponding time steps $k_n = t_n - t_{n-1}$ and time intervals $I_n = (t_{n-1}, t_n)$. Let $P_q(H) = \{v : v(t) = \sum_{j=0}^q v_j t^j, \ v_j \in H\}$ and define the function space $V_q = \{v = v(t) : v|_{I_n} \in$

$P_q(H),\ n = 1,2,\dots\}$. We note that the functions in V_q may be discontinuous at the mesh points t_n and we define $v_n^{\pm} = \lim_{s\to 0^{\pm}} v(t_n + s)$.

The dG(q)-method for (1.1) defines an approximate solution $U \in V_q$ recursively on I_n, for $n = 1,2,\dots$, by setting $U_0^- = u_0$, and then

$$\int_{I_n} (\dot U + AU, v)\,dt + (U_{n-1}^+, v_{n-1}^+) = (U_{n-1}^-, v_{n-1}^+) + \int_{I_n} (f, v)\,dt, \quad \forall v \in P_q(H). \tag{3.1}$$

In Theorem 3 below we show that (3.1) has a unique solution $U \in P_q(H)$ for any $U_{n-1}^- \in H$ and $f \in L_1(I_n; H)$.

For the analysis of this method we introduce some more notation. Let $[v]_n = v_n^+ - v_n^-$ and define the bilinear form

$$B_N(v,w) = \sum_{n=1}^{N} \int_{I_n} (\dot v + Av, w)\,dt + \sum_{n=1}^{N-1} ([v]_n, w_n^+) + (v_0^+, w_0^+) \tag{3.2}$$

$$= \sum_{n=1}^{N} \int_{I_n} (v, -\dot w + A^* w)\,dt - \sum_{n=1}^{N-1} (v_n^-, [w]_n) + (v_N^-, w_N^-), \tag{3.3}$$

where the second line is obtained by integration by parts. By summation of (3.1) it follows that U satisfies

$$B_N(U, v) = (u_0, v_0^+) + \int_0^{t_N} (f, v)\,dt, \quad \forall v \in V_q,\ N = 1,2,\dots. \tag{3.4}$$

In view of (1.1) we then have, for the error $e = u - U$,

$$B_N(e, v) = 0, \quad \forall v \in V_q,\ N = 1,2,\dots. \tag{3.5}$$

This "orthogonality relation" is a basic ingredient in our error analysis.

4. The *a posteriori* error estimate

We shall now prove an *a posteriori* error estimate of the form (1.3) for the dG(q)-method (3.1) applied to (1.1), with the logarithmic factor modified to $L_N = 1 + \log(t_N/k_N)$. We recall that the piecewise constant functions $k = k(t)$ and $[U] = [U](t)$ are defined by

$$k(t) = k_n = t_n - t_{n-1}, \quad [U](t) = [U]_{n-1} = U_{n-1}^+ - U_{n-1}^-, \quad \text{for } t \in I_n.$$

Theorem 1. *Let u and U be the solutions of (1.1) and (3.4), respectively. Then we have, for $N = 1,2,\dots$,*

$$\|u(t_N) - U_N^-\| \le C^i C^s L_N \Big(\|[U]\|_{L_\infty(J_N;H)} + \|k^{q+2} f^{(q+1)}\|_{L_\infty(J_N;H)} \Big),$$

where $J_N = (0, t_N)$, the constant C^i depends only on q, C^s is the constant in (2.2), and $L_N = 1 + \log(t_N/k_N)$.

Proof. Let $z(t) = E(t_N - t)^* e_N^-$ for $t < t_N$, where $e_N^- = u(t_N) - U_N^-$. In view of (2.2) we have

$$\|z(t)\| \le C^s \|e_N^-\|, \quad \|A^* z(t)\| \le C^s (t_N - t)^{-1} \|e_N^-\|, \quad \text{for } t < t_N. \tag{4.1}$$

Since z is the solution of the backward evolution problem

$$-\dot z(t) + A^* z(t) = 0,\ t < t_N; \quad z(t_N) = \phi, \tag{4.2}$$

with data $\phi = e_N^-$, which is the dual of the problem (2.3), it is clear from (3.3) that

$$B_N(w,z) = (w_N^-, e_N^-),$$

for any piecewise smooth function w. In particular, with $w = e = u - U$, in view of (3.5) and (3.2), we have, for any $v \in V_q$,

$$\begin{aligned}
\|e_N^-\|^2 &= B_N(e,z) = B_N(e, z - v)\\
&= \sum_{n=1}^{N} \Big(\int_{I_n} (\dot{e} + Ae, z - v)\, dt + ([e]_{n-1}, (z-v)_{n-1}^+) \Big)\\
&= \sum_{n=1}^{N} \Big(\int_{I_n} (f - \dot{U} - AU, z - v)\, dt - ([U]_{n-1}, (z-v)_{n-1}^+) \Big),
\end{aligned}$$

where we also used the facts that $e_0^- = u_0 - U_0^- = 0$, $\dot{u} + Au = f$, and $[u]_{n-1} = 0$. Choosing $v = Pz$, where $P : L_2(J_N; H) \to V_q$ is the orthogonal projection defined by

$$\int_{I_n} (Pz, w)\, dt = \int_{I_n} (z, w)\, dt, \quad \forall w \in P_q(H), \quad n = 1, 2, \ldots, \tag{4.3}$$

we get

$$\|e_N^-\|^2 = \sum_{n=1}^{N} \Big(\int_{I_n} (f - Pf, z - Pz)\, dt - ([U]_{n-1}, (z - Pz)_{n-1}^+) \Big).$$

Estimating $z - Pz$ here by

$$\|z - Pz\|_{L_\infty(I_n;H)} \le C^i \min \Big(\|\dot{z}\|_{L_1(I_n;H)}, \|z\|_{L_\infty(I_n;H)} \Big),$$

and recalling (4.1), we get

$$\begin{aligned}
\|e_N^-\|^2 &\le C^i \sum_{n=1}^{N-1} \Big(k_n \|f - Pf\|_{L_\infty(I_n;H)} + \|[U]_{n-1}\| \Big) \|\dot{z}\|_{L_1(I_n;H)}\\
&\quad + C^i \Big(k_N \|f - Pf\|_{L_\infty(I_N;H)} + \|[U]_{N-1}\| \Big) \|z\|_{L_\infty(I_N;H)}\\
&\le C^i \Big(\|k(f - Pf)\|_{L_\infty(J_N;H)} + \|[U]\|_{L_\infty(J_N;H)} \Big)\\
&\quad \times \Big(\|\dot{z}\|_{L_1(J_{N-1};H)} + \|z\|_{L_\infty(I_N;H)} \Big)\\
&\le C^i C^s L_N \Big(\|k(f - Pf)\|_{L_\infty(J_N;H)} + \|[U]\|_{L_\infty(J_N;H)} \Big) \|e_N^-\|.
\end{aligned}$$

Finally, we use the fact that

$$\|f - Pf\|_{L_\infty(I_n;H)} \le C^i k_n^{q+1} \|f^{(q+1)}\|_{L_\infty(I_n;H)}, \tag{4.4}$$

and the proof is complete. □

5. The *a priori* error estimate

We now prove an *a priori* error estimate of the form (1.2). We begin by showing that the uniform-in-time strong stability (2.2) for the bounded analytic semigroup $E(t)$ carries over to the solution operator of the dG(q)-method.

In order to formulate this, we let $\phi \in H$ be given and let $Z \in V_q$ be the solution of the equation

$$B_N(v, Z) = (v_N^-, \phi), \quad \forall v \in V_q, \tag{5.1}$$

which is the dG(q)-discretization of the backward problem (4.2) and the dual of (3.4) with $f \equiv 0$. The following result is the main technical innovation of the present work. It provides the strong stability estimate needed in the *a priori* error estimate. The corresponding result for the case when A is self-adjoint was proved in [1] under the mesh condition $k_n \leq ck_{n+1}$, which we have removed here.

Theorem 2. *There is a constant C_q^s depending only on C^s and q such that the solution Z of (5.1) satisfies, for $N = 1, 2, \dots,$ with $J_N = (0, t_N)$ and $L_N = 1 + \log(t_N/k_N)$,*

$$\|Z\|_{L_\infty(J_N;H)} \leq C_q^s \|\phi\|, \tag{5.2}$$

$$\|A^*Z\|_{L_\infty(I_n;H)} \leq C_q^s (t_N - t_{n-1})^{-1} \|\phi\|, \quad n = 1, \dots, N, \tag{5.3}$$

$$\|A^*Z\|_{L_1(J_N;H)} \leq C_q^s L_N \|\phi\|. \tag{5.4}$$

Proof. We first note that (5.4) follows from (5.3). For the proofs of (5.2) and (5.3) we note that, by the change of variable $t \to t_N - t$, it is equivalent to estimate the solution of the forward evolution problem (3.4) with A replaced by A^*, $f = 0$, and a reversed mesh. Since $\|E(t)\| = \|E(t)^*\|$ and $\|AE(t)\| = \|A^*E(t)^*\|$, we conclude that A^* is sectorial with the same constants as A; see (2.1). For convenience of notation, we shall therefore estimate the solution U of (3.4) with $f = 0$ and show that

$$\|U\|_{L_\infty(I_n;H)} + t_n \|AU\|_{L_\infty(I_n;H)} \leq C\|u_0\|, \quad n = 1, 2, \dots. \tag{5.5}$$

Here, and in the following, C denotes various constants that depend only on q and on the constants in (2.1) or, equivalently, on the constant C^s in (2.2), but not on t_n or the mesh.

1. Estimates at the nodes. It is known, see [6], that the nodal values of U are given by $U_n^- = \prod_{l=1}^n r(k_l A) u_0$, where $r(\lambda) = p_1(\lambda)/p_2(\lambda)$ is a rational function, namely, the Padé approximation of $e^{-\lambda}$ of order $O(\lambda^{2q+1})$, with p_1 and p_2 polynomials of degrees q and $q+1$, respectively. For example, $r(\lambda) = 1/(1+\lambda)$ for $q = 0$ (the implicit Euler method), and $r(\lambda) = (1 - \lambda/3)/(1 + 2\lambda/3 + \lambda^2/6)$ for $q = 1$.

The estimate corresponding to (5.2) at the nodes, i.e.,

$$\|U_n^-\| = \Big\| \prod_{l=1}^n r(k_l A) u_0 \Big\| \leq C\|u_0\|, \quad n = 1, 2, \dots, \tag{5.6}$$

was proved in [9]. We shall show that

$$\|AU_n^-\| \leq Ct_n^{-1} \|u_0\|, \quad n = 1, 2, \dots, \tag{5.7}$$

corresponding to (5.3); for transparency, we begin with the special case $q = 0$.

Let $k_{\max} = \max_{1 \leq j \leq n} k_j$ and consider first the case $k_{\max} \leq t_n/2$. We use the Dunford-Taylor formula

$$AU_n^- = \frac{1}{2\pi i} \int_\Gamma \prod_{l=1}^n r(k_l \lambda) AR(\lambda, A)\, d\lambda\, u_0, \tag{5.8}$$

where $R(\lambda, A) = (\lambda I - A)^{-1}$ is the resolvent operator, and $\Gamma = \{\lambda \in \mathbf{C} : \lambda = c|x| - ix, \ -\infty < x < \infty\}$ with $c > 0$ so small that Γ is contained in the resolvent set of A, so that (2.1) (with $\omega = 0$) can be applied with $-\lambda \in \Gamma$. Then $\|AR(\lambda, A)\| \le C$ on Γ (see [10, Theorem 5.2]), and, using the fact that, since $q = 0$, $|r(k_l\lambda)| = 1/|1 + k_l\lambda| \le 1/(1 + ck_l|x|)$ for $\lambda = c|x| - ix$, we obtain

$$\|AU_n^-\| \le C \int_0^\infty \prod_{l=1}^n \frac{1}{1 + ck_l x}\, dx\, \|u_0\|.$$

Here, for $x \ge 0$,

$$\prod_{l=1}^n (1 + k_l x) \ge 1 + \left(\sum_{l=1}^n k_l\right) x + \tfrac{1}{2}\left(\sum_{l,j=1,\ l\ne j}^n k_l k_j\right) x^2,$$

where $\sum_{l=1}^n k_l = t_n$, and

$$t_n^2 = \Big(\sum_{l=1}^n k_l\Big)^2 = \sum_l k_l^2 + \sum_{l\ne j} k_l k_j \le k_{\max} t_n + \sum_{l\ne j} k_l k_j,$$

so that, under our present assumption $k_{\max} \le t_n/2$, we have $\sum_{l,j=1,l\ne j}^n k_l k_j \ge t_n^2/2$. Hence

$$\prod_{l=1}^n (1 + k_l x) \ge 1 + t_n x + \tfrac{1}{4} t_n^2 x^2 \ge C(1 + t_n^2 x^2), \quad \text{for } x \ge 0,$$

and consequently, with $s = ct_n x$,

$$\|AU_n^-\| \le C t_n^{-1} \int_0^\infty \frac{ds}{1 + s^2}\, \|u_0\| \le C t_n^{-1} \|u_0\|.$$

We now consider the case when $k_{\max} \ge t_n/2$. For some m with $1 \le m \le n$, we have $k_m = k_{\max}$. By application of (5.6) to the discrete counterpart of the equation $\dot{w} + Aw = 0$ with $w = Au$, we also have

$$\|AU_n^-\| \le C \|AU_m^-\|.$$

The desired estimate (5.7) now follows at once from the fact that, since $q = 0$,

$$AU_m^- = k_m^{-1}(U_{m-1}^- - U_m^-).$$

More precisely,

$$\|AU_m^-\| \le k_m^{-1}(\|U_{m-1}^-\| + \|U_m^-\|) \le C k_m^{-1} \|U_{m-1}^-\|,$$

since $\|r(k_m A)\| = k_m^{-1}\|(k_m^{-1} + A)^{-1}\| \le C$ according to the resolvent estimate (2.1). This completes the proof of (5.7) in the special case $q = 0$.

In the general case, $q \ge 0$, we still have (5.8) but with $r(\lambda) = p_1(\lambda)/p_2(\lambda)$ as described above. Again, because $|r(\lambda)| \le 1$ in the right half-plane (see, e.g., [11]), and p_2 is of higher degree than p_1, we have that $|r(k_l\lambda)| \le 1/(1 + ck_l|x|)$ on Γ for some $c > 0$, and we thus obtain (5.7) when $k_{\max} \le t_n/2$ as before.

In the case $k_{\max} \ge t_n/2$ it suffices to verify that

$$\|AU_m^-\| \le C k_m^{-1} \|U_{m-1}^-\|.$$

To see that this estimate holds also for $q > 0$ we write $r(\lambda) = r_1(\lambda) + r_2(\lambda)$, where $r_1(\lambda) = c/(\lambda + d)$, with $-d$ being a root of p_2, so that d is in the right half-plane, and $r_2(\lambda) = p_3(\lambda)/p_2(\lambda)$ with c chosen so that $p_3(\lambda) = p_1(\lambda) - cp_2(\lambda)/(\lambda + d)$ is of degree

$\leq q-1$. Correspondingly, we write $U_m^- = U_m^1 + U_m^2$, where $U_m^j = r_j(k_m A)U_{m-1}^-$, $j=1,2$. Similarly to the case $q=0$, we have

$$\|AU_m^1\| \leq Ck_m^{-1}(\|U_{m-1}^-\| + \|U_m^1\|) \leq Ck_m^{-1}\|U_{m-1}^-\|,$$

where in the last step we used the fact that $\|r_1(k_m A)\| = |c|k_m^{-1}\|(dk_m^{-1}+A)^{-1}\| \leq C$ according to the resolvent estimate (2.1). For the other part we write

$$AU_m^2 = \frac{1}{2\pi i}\int_\Gamma \frac{p_3(k_m\lambda)}{p_2(k_m\lambda)} AR(\lambda, A)\, d\lambda\, U_{m-1}^-,$$

again leading to the estimate, with $s = ck_m x$,

$$\|AU_m^2\| \leq Ck_m^{-1}\int_0^\infty \frac{ds}{1+s^2}\,\|U_{m-1}^-\| \leq Ck_m^{-1}\|U_{m-1}^-\|,$$

because p_2 has no roots on Γ and is of degree $q+1$, while p_3 is of degree $q-1$.

2. Interior estimates. So far we have estimated U and AU at the discrete time levels t_n. In order to obtain the desired estimates for all $t \in I_n$, we write

$$U(t) = \sum_{l=0}^q U_{n,l}\psi_l((t-t_{n-1})/k_n), \quad \text{for } t \in I_n,$$

where ψ_l are the polynomials of degree $\leq q$ determined by

$$\psi_l(j/q) = \delta_{jl}, \quad j,l = 0,\dots,q \quad (\delta_{jl} \text{ is Kronecker's delta}).$$

With matrices $L = (L_{jl})$ and $K = (K_{jl})$ defined by $L_{jl} = \int_0^1 \psi_l'(s)\psi_j(s)\,ds + \delta_{l0}\delta_{j0}$ and $K_{jl} = \int_0^1 \psi_l(s)\psi_j(s)\,ds$, and vectors $\hat{U}_n = (U_{n,j})$ and $F = (F_j)$ with $F_0 = U_{n-1}^-$, $F_j = 0$, $j = 1,\dots,q$, we can write (3.1) with $f=0$ as

$$(IL + k_n AK)\hat{U}_n = F,$$

where $I : H \to H$ is the identity operator. Solving for the components of $\hat{U}_n$ using Cramer's rule, we find that

$$U_{n,j} = r_j(k_n A)U_{n-1}^-,$$

where $r_j(\lambda) = p_{1,j}(\lambda)/p_2(\lambda)$, p_2 is the same polynomial of degree $q+1$ as above, with no roots in the right half-plane, and the $p_{1,j}$ are polynomials of degree q. Using again the Dunford-Taylor formula, we have

$$A^l U_{n,j} = \frac{1}{2\pi i}\int_{\Gamma_0} r_j(k_n\lambda)R(\lambda, A)\,d\lambda\, A^l U_{n-1}^-, \quad l = 0,1,\ j = 0,\dots,q,$$

where the contour Γ_0 is obtained from Γ by passing to the left of the origin on a circular arc of radius ϵk_n^{-1} with ϵ so small that the $r_j(\lambda)$ have no poles inside the circle of radius ϵ. Recalling from (2.1) that $\|R(\lambda, A)\| \leq M|\lambda|^{-1}$ for $\lambda \in \Gamma_0$, we obtain by a standard argument

$$\|A^l U_{n,j}\| \leq C\|A^l U_{n-1}^-\|, \quad l = 0,1,\ j = 0,\dots,q.$$

Hence, using also our previous estimate (5.6) for U_n^-,

$$\|U\|_{L_\infty(I_n;H)} \leq C\max_j \|U_{n,j}\| \leq C\|U_n^-\| \leq C\|u_0\|, \quad n = 1,2,\dots, \tag{5.9}$$

which is the desired estimate for U in (5.5).

In order to estimate AU we first assume that $k_n \le t_n/2$. We then obtain in a similar way (note that $n > 1$ and $t_n \le 2t_{n-1}$ in this case)

$$\|AU\|_{L_\infty(I_n;H)} \le C \max_j \|AU_{n,j}\| \le C\|AU_{n-1}^-\| \le \frac{C}{t_{n-1}}\|u_0\| \le \frac{C}{t_n}\|u_0\|.$$

It remains to consider the case when $k_n \ge t_n/2$. Then the identity

$$k_n A\hat{U}_n = K^{-1}(F - IL\hat{U}_n),$$

together with our previous estimate (5.6) for U_n^-, gives

$$\|AU\|_{L_\infty(I_n;H)} \le C \max_j \|AU_{n,j}\| \le Ck_n^{-1}\|U_{n-1}^-\| \le \frac{C}{t_n}\|U_{n-1}^-\| \le \frac{C}{t_n}\|u_0\|.$$

This proves the estimate for AU in (5.5). □

The following existence result for equation (3.1) is proved by the techniques used in the second step of the previous proof. We omit the details.

Theorem 3. *Equation (3.1) has a a unique solution $U \in P_q(H)$ for any $U_{n-1}^- \in H$ and $f \in L_1(I_n;H)$. The solution satisfies $U(t) \in D(A)$ and*

$$\|U\|_{L_\infty(I_n;H)} + k_n\|AU\|_{L_\infty(I_n;H)} \le C\big(\|U_{n-1}^-\| + \|f\|_{L_1(I_n;H)}\big),$$

where C depends only on q.

Finally we prove the *a priori* error estimate.

Theorem 4. *Let u and U be the solutions of (1.1) and (3.4), respectively. Then we have, for $N = 1, 2, \ldots,$*

$$\|u - U\|_{L_\infty(J_N;H)} \le C^i C_q^s\big(\max_{1\le n\le N} L_n\big)\|k^{q+1}u^{(q+1)}\|_{L_\infty(J_N;H)},$$

where $J_N = (0, t_N)$, the constant C^i depends only on q, C_q^s is the constant in (5.4), and $L_n = 1 + \log(t_n/k_n)$.

Proof. Let $Q : C(J_N;H) \to V_q$ be the interpolation operator defined by

$$(Qu)_n^- = u_n^-; \quad \int_{I_n} (Qu)v\,ds = \int_{I_n} uv\,ds, \quad \forall v \in P_{q-1}(H)$$

(the second condition is not used when $q = 0$). Choosing $\phi = e_N^-$, $v = Qu - U$ in (5.1), we get

$$\begin{aligned}
\|e_N^-\|^2 &= ((Qu - U)_N^-, e_N^-) = B_N(Qu - U, Z)\\
&= B_N(Qu - u, Z) + B_N(e, Z) = B_N(Qu - u, Z)\\
&= \sum_{n=1}^{N} \int_{I_n} (Qu - u, -\dot{Z} + A^*Z)\,dt\\
&\quad - \sum_{n=1}^{N-1}((Qu - u)_n^-, [Z]_n) + ((Qu - u)_N^-, Z_N^-)\\
&= \sum_{n=1}^{N} \int_{I_n} (Qu - u, A^*Z)\,dt,
\end{aligned}$$

where we used the orthogonality relation (3.5), (3.3), and the definition of Q. The interpolation error estimate

$$\|u - Qu\|_{L_\infty(I_n;H)} \le C^i k_n^{q+1} \|u^{(q+1)}\|_{L_\infty(I_n;H)},$$

and (5.4) of Theorem 2 now yield

$$\|e_N^-\| \le C^i C_q^s L_N \|k^{q+1} u^{(q+1)}\|_{L_\infty(J_N;H)}, \quad N = 1, 2, \ldots . \tag{5.10}$$

In order to estimate $\|e\|_{L_\infty(I_n;H)}$ we write $e = u - U = (u - Pu) + (Pu - U) \equiv \eta + \theta$, where P is the orthogonal projection defined in (4.3). In view of (4.4) it suffices to estimate $\|\theta\|_{L_\infty(I_n;H)}$. Substituting $e = \eta + \theta$ in the local version of the error equation (3.5), and integrating by parts on $\dot\eta$, we arrive at

$$\begin{aligned}\int_{I_n} (\dot\theta + A\theta, v)\, dt + (\theta_{n-1}^+, v_{n-1}^+) &= (e_{n-1}^-, v_{n-1}^+) - (\eta_n^-, v_n^-) - \int_{I_n} (\eta, -\dot v + A^* v)\, dt \\ &= (e_{n-1}^-, v_{n-1}^+) - (\eta_n^-, v_n^-) \quad \forall v \in P_q(H).\end{aligned}$$

By the same argument as in the proof of (5.9) this leads to

$$\|\theta\|_{L_\infty(I_n;H)} \le C\big(\|e_{n-1}^-\| + \|\eta_n^-\|\big).$$

Hence

$$\|e\|_{L_\infty(J_N;H)} \le C\big(\max_{1\le n\le N} \|e_{n-1}^-\| + \|\eta\|_{L_\infty(J_N;H)}\big),$$

and the required result follows by (5.10) and (4.4). □

References

[1] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems.* I. *A linear model problem*, SIAM J. Numer. Anal. **28** (1991), 43–77.

[2] ———, *Adaptive finite element methods for parabolic problems.* II. *Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal. **32** (1995), 706–740.

[3] ———, *Adaptive finite element methods for parabolic problems.* III. *Time steps variable in space*, in preparation.

[4] ———, *Adaptive finite element methods for parabolic problems.* IV. *Nonlinear problems*, SIAM J. Numer. Anal. **32** (1995), 1729–1749.

[5] ———, *Adaptive finite element methods for parabolic problems.* V. *Long-time integration*, SIAM J. Numer. Anal. **32** (1995), 1750–1763.

[6] K. Eriksson, C. Johnson, and V. Thomée, *Time discretization of parabolic problems by the discontinuous Galerkin method*, RAIRO Modél. Math. Anal. Numér. **19** (1985), 611–643.

[7] C. Johnson, R. Rannacher, and M. Boman, *Numerics and hydrodynamic stability: toward error control in computational fluid dynamics*, SIAM J. Numer. Anal. **32** (1995), 1058–1079.

[8] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Birkhäuser, Basel, 1995.

[9] C. Palencia, *On the stability of variable stepsize rational methods for holomorphic semigroups*, Math. Comp. **62** (1994), 93–103.

[10] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.

[11] G. Wanner, E. Hairer and S. P. Nørsett, *Order stars and stability theorems*, BIT **18** (1978), 475–489.

Kenneth Eriksson, Department of Mathematics, Chalmers University of Technology and Göteborg University, S–412 96 Göteborg, Sweden
E-mail address: kenneth@math.chalmers.se

Claes Johnson, Department of Mathematics, Chalmers University of Technology and Göteborg University, S–412 96 Göteborg, Sweden
E-mail address: claes@math.chalmers.se

Stig Larsson, Department of Mathematics, Chalmers University of Technology and Göteborg University, S–412 96 Göteborg, Sweden
E-mail address: stig@math.chalmers.se